ELSEVIER

Biophysical Chemistry 50 (1994) 169–181

Biophysical Chemistry

Effects of DNA topology in the interaction with histone octamers and DNA topoisomerase I

Rodolfo Negri, Giovanna Costanzo, Memmo Buttinelli, Sabrina Venditti, Ernesto Di Mauro

Centro di Studio per gli Acidi Nucleici, CNR, Roma, Italy, Fondazione "Istituto Pasteur – Fondazione Cenci Bolognetti", c/o Dipartimento di Genetica e Biologia Molecolare, Università "La Sapienza", Roma, Italy

(Received 5 January 1994)

Abstract

Several simple proteins and complex protein systems exist which do not recognize a defined sequence but – rather – a specific DNA conformation. We describe experiments and principles for two of these systems: nucleosomes and eukaryotic DNA topoisomerase I. Evidences are summarized that describe the effects of negative DNA supercoiling on nucleosome formation and the influence of DNA intrinsic curvature on their localization. The function of the DNA rotational information in nucleosome positioning and in the selection of multiple alternative positions on the same helical phase are described. This function suggests a novel genetic regulatory mechanism, based on nucleosome mobility and on the correlation between *in vitro* and *in vivo* positions. We observe that the same rules that determine the *in vitro* localization apply to the *in vivo* nucleosome positioning, as determined by a technique that relies on the use of nystatin and on the import of active enzymes in living yeast cells. The sensitivity of DNA topoisomerase I to the topological condition of the DNA substrate is reviewed and discussed taking into account recent experiments that describe the effect of the DNA tridimensional context on the reaction. These topics are discussed in the following order: (i) Proteins that look for a consensus DNA conformation; (ii) Nucleosomes; (iii) Negative supercoiling and nucleosomes; (iv) DNA curvature/bending and nucleosomes; (v) Multiple positioning; (vi) Multiple nucleosomes offer a contribution to the solution of the linking number paradox; (vii) Rotational versus translational information; (viii) A regulatory mechanism; (ix) DNA topoisomerase I; (x) DNA topoisomerase I and DNA supercoiling: a regulation by topological feedback; (xi) DNA topoisomerase I and DNA curvature; (xii) The in-and-out problem in the accessibility of DNA information; (xiii) The integrating function of the free energy of supercoiling.

Key words: DNA topoisomerase I; DNA topology; Promoter activation; *in vitro* and *in vivo* nucleosomes; Rotational and translational information

SSDI 0301-4622(94)00018-F

1. Proteins that look for a consensus DNA conformation

Proteins which interact with DNA search the sequence for a specific site, looking along the strands for information. In many instances sequence information is provided by relatively simple base-pair combinations. A restriction enzyme may be content enough with a CCGG, or similar. Other proteins are much more fastidious, to the point that defining the rules which organize the information they are looking for has proven very difficult.

In general, the wider the relevance of the function of the proteins under consideration, the more wide-spread in the genome are their interaction sites, and the more difficult is the solution to the problem of defining the structure of the interaction sites.

The proteins which react in many different points of the genome may be separated in two major classes: the proteins which interact with a specific sequence, and those which use more subtle and/or complex discriminatory principles. The reaction of the first class only requires an active protein and a sterically available sequence (i.e. the restriction enzymes or, at a more complex level of interaction, the TATA binding factors). The interaction may be a solitary one or may require an interplay with ancillary proteins or factors.

The reaction of a protein of the second class is more mysterious and often paradoxical: how can proteins which interact on specific sites all over the genome find their precise interaction sites along sequences that by definition differ (that is: sequences that differ for the order of the nucleotides that compose them)?

The histone octamers that are part of the nucleosomes are a paradigmatic example of this type of proteins. Many precisely localized nucleosomes have been described [1–5], yet the DNA sequences that interact in these multiprotein complexes differ. Given that nucleosomes cover genomes whose complexity reaches the order of 10^9 bp, the ability to precisely locate a reaction site along an almost endless string of DNA cannot reside in modifications (genetic or epigenetic) of the histones. For this purpose, histones can be considered as a homogeneous component of the selective binding reaction, even though this is certainly a simplification of the problem. The simplification is justified by the extreme complexity of the other reagent: the DNA.

A similar problem characterizes the function of DNA topoisomerase I. Its reaction with DNA is certainly regulated and its interaction sites are wide-spread over the genome; yet, topoisomerization does not occur on a consensus sequence.

For both nucleosomes and DNA topoisomerase I, the solution of this paradox must be looked for in the structural and topological properties of DNA. We will discuss in the last paragraph that this solution probably applies also to other DNA–protein systems.

2. Nucleosomes

The nucleosome subunit of chromatin consists of DNA folded around a histone core as a 1.8 turn left-handed solenoid (for a recent review, see ref. [6]. We use here the currently adopted terminology that refers to nucleosomal core particles as to the particles typically formed by the interaction of 2 × H2A, H2B, H3 and H4 (histone octamers), distinguished from the nucleosomal particles formed by the additional presence of an H1 or H5 molecule. A corollary of the wrapping of DNA around the protein core is that the formation of a nucleosome should be favoured either by preexisting negative supercoiling, or by preformed DNA curves in the same direction.

Evidences that prove the validity of this corollary providing information relevant to a more detailed understanding of the DNA–histone core interaction are detailed below [7–9].

The local conformation of DNA determines the selectivity of the interaction of histone octamers as shown by Klug, Travers, Drew and co-workers [10–13]. Sequence-dependent bendability (or anisotropic flexibility [14–17]) has been indicated as the major cause of a rotationally related positioning of octamers, as shown by DNase I footprinting of nucleosomes on unique DNA sequences [10] and by the specific positions of $(A)_n$ and $(T)_n$ runs relative to the octamer [11].

These pioneering studies based on the analysis of preferred sequences have shown that the 3′ end of homo $(A)_n$ or $(T)_n$ is preferentially positioned towards the histone octamer [12] and have established the principle of the rotational relationship between different possible locations of the histone octamer. This conclusion, based on the analysis of a large number of cloned DNA fragments from chicken erythrocyte nucleosomes, has established that the preferential formation of nucleosomes had occurred *in vivo* on the analysed DNA segments and has defined the sequence-related conformational properties of these DNA segments. In subsequent studies [7] it was observed that one of two segments analyzed had properties of intrinsic curvature.

3. Negative supercoiling and nucleosomes

Removal of nucleosomes from closed DNA domains releases supercoiling (as originally observed by Germond et al. [18]. The availability of local free energy is expected to facilitate interaction with several proteins including the ones that have generated it, the histones themselves. Sensitivity of histones to DNA topology has actually been reported by Goulet et al. [19] and negative supercoiling does indeed allow formation of nucleosomal core particles in topologically stringent conditions, as shown by Negri et al. [20].

This result was obtained in a DNA system programmed in such a way as to mimic the topological situation that is likely to occur *in vivo* in chromatin when a single individual nucleosome dissociates and (possibly) reassociates. In chromatin single nucleosomes come off a chain of similar particles and leave behind a topologically defined space. Therefore, an *in vitro* experimental system for nucleosome reconstitution in a closed DNA domain was established, into which only one individual nucleosome could fit. For this purpose, microcircles of various linking number were produced by ligation of a 185 bp long DNA segment in differential conditions. This size roughly corresponds to the average internucleosomal spacing (146 bp + a short variable linker) and is so small that it disfavours the formation of titratable superhelical turns (as defined by Wang et al. [21], and Shore and Baldwin [22,23]). The conditions of ligation were such that molecules with three different linking numbers could be produced: Lk = 18, 17 and 16 (or, in differential terms: $\Delta Lk = 0, -1$ and -2). In this way it is possible to analyze the efficiency of topologically closed microdomains in forming nucleosomes avoiding the complications intrinsic to larger DNA domains (namely: topological heterogeneity and complexity [19,24].

DNA in these microcircular topoisomers is topologically highly constrained, as supposedly is a 180–200 bp long DNA tract on the chromatin, from which a nucleosome has come off and tends to go back.

In this system the promoting effect of free energy on histone octamers reconstitution could be studied on isolated topoisomers. The results showed that the efficiency of formation of histone octamers is strictly dependent upon linkage reduction, thus pointing to the close relationship between local free energy and protein interaction in this system.

If the nucleosomes surrounding a nucleosomal particle removed from a promoter behave as a barrier that impedes the transmission of free energy along the chromatin fiber, the superhelical density could accumulate locally (on average: 1 superhelical turn/200 bp, that is: $-\sigma = 0.05$) and be available for the activation of the correlated functions of DNA topoisomerase I and of RNA polymerase II and for nucleosome reformation.

The complete impediment of transmission of free energy through nucleosomes has been suggested by the suppression of the thermal untwisting of DNA by histone octamers (described by Keller [25] and by Morse and Cantor [26,27]. The partial impediment of such transmission has been discussed [28,29]. We shall come back to the possible integrating function of free energy after having described a few other effects of DNA topology on nucleosomes.

4. DNA curvature / bending and nucleosomes

Let us distinguish curvature (an intrinsic property) and bending (a property induced by physico-chemical conditions and/or by protein

binding). Nucleosomes are interested in both. Nucleosomes preferentially form on intrinsically curved DNA sequences and – upon formation – induce or increase the curvature to the fixed and well characterized value of 1.8 superhelical turns. These considerations have ultimately led to the development of algorithms for prediction of histone octamers binding sites [30–32].

Keeping the basic equation of DNA topology ($Lk = T + W$) in mind, we observe that nucleosomes formation is sensitive to the variation of both twisting (T) and writhing (W). We have seen in the previous section that linkage reduction favors the association of histone octamers on microcircles so small that they cannot undergo writhing (that is: octamer formation is favored by variations of the twist value). The fact that DNA wraps around histone cores has led to formulate the easy prediction that DNA curvature should favor histone octamer formation (evidencing that their formation is favored also by the other term of the equation, the writhing). The possibility that the strain of bending DNA about the histone core might be relieved if the DNA contained a sequence-directed curvature has been proposed long ago by Trifonov and co-workers [14,15].

The relevance of local anisotropy in positioning nucleosomes *in vivo* and histone octamers *in vitro* has been observed in several instances. Preferential localization of histone octamers on large DNAs containing curved inserts has been reported [33], showing the preference for curved segments in selective reconstitution conditions and the localization on both curved and adjacent regions under non-selective conditions. In reconstitutions performed on the *Crithidia fasciculata* kinetoplast DNA (a left-handed intrinsically and highly curved molecule) it was observed by Costanzo et al. [8] that the curved segment preferentially attracts and phases histone octamers. In particular, the results *in vitro* show that: (a) the intrinsic periodic anisotropy of the curved DNA favors the reconstitution of octamers relative to mixed-sequence DNA by an order of magnitude; (b) the octamers form precisely in phase with the phased anisotropy; (c) octamers bound to a segment of curved DNA influence the position of neighbouring particles.

Taken together, the points (a) and (b) indicate that octamer formation occurs by preferential selection of sequences located on a defined rotational phase. This is equivalent to say that: the *C. fasciculata* DNA is programmed to locate multiple nucleosomes on a given side, or "face". Is this DNA an exception or does it represent a frequent or even universal case? Based on our experience (detailed below) the *C. fasciculata* intrinsically curved kinetoplast DNA segment is not an exception.

The mechanisms that can cause or influence nucleosome positioning have been discussed [1,3–5,34–38]. For a recent review, see Freeman and Garrard [6].

A single location or a set of defined positions could be determined by one of the following causes (or by a combination thereof): (a) boundary effects and size of the domain; (b) sequentiality of the formation; (c) higher affinity for a defined sequence-related conformational property. The logic of the case (a) is intuitive: in a limited space only a defined number of nucleosomes can fit, and the space available defines the positions. In this case, the quality of the underlying DNA sequences plays only a minor role. The borders of the domain could be set by special DNA structures (endowed with extremely high or extremely low affinity for nucleosomes). Irrespective of their nature, their role in this case is represented by the difference with the average affinity for DNA of the other histone octamers. The same task could be acquitted by preexisting nucleosomes or other proteins. This takes us into model (b): the position of the second nucleosome could be defined by the first one, and so on. Cooperativity of the process could be a relevant component of the mechanism of nucleosome deposition. Strangely enough, this possibility has not been explored in detail and cooperativity mechanisms in histone octamers have been reported only recently [39,40]. Evidence that the position of a given histone octamer is affected by the neighbouring one has been obtained [8].

Cases (a) and (b) are somewhat similar in their nature; (b) takes the history of the local organization into a greater consideration. Case (c) is based on the experimental observation that specific

combinations of sequences have been identified in the tracts of DNA engaged *in vivo* in nucleosomal particles (as discussed in Travers [41] and in the references therein). These studies have revealed that certain sequences are favored in defined positions of the 146 bp unit, thus establishing the relevance of the local DNA conformation in the positioning of nucleosomes.

According to these analyses, the sequence-dependent positioning of nucleosomes is explained by the differential flexibility of different sequences, thus allowing localization on the basis of the local lower resistance to protein-induced flexures relative to the surrounding DNA.

5. Multiple positioning

In the case of *C. fasciculata* intrinsically curved DNA segment, histone octamers formed in multiple positions and evidence was obtained (a) for this DNA being characterized by anisotropic flexibility, (b) that the DNA curvature generated by this anisotropic property depends on the distribution of differently flexible sequences and (c) that DNA sequences adopt *in vitro* defined locations with respect to the histone octamers. This set of properties defines preferentiality over bulk DNA and solves the paradox of the specific localization obtained on variable sequences. In conclusion: the relevant parameter reside in the structure of the sequences, not in the sequences themselves.

In this type of interaction, the rotational constraints are strong, are regularly repeated and define on what side of the DNA the interaction will occur.

At this point, if one general problem is solved, another one requires attention: how can the specific localization be obtained? Do translational signals exist on DNA that can define single positions among the several ones which are possible on the same side of the double strand? The existence of asymmetrically located translational signals has been proposed and discussed [42,43]. Their nature remains ill-defined and elusive; this is probably the major problem today in understanding the factors that cause the localization of single specifically located nucleosomes.

We suggest that rotational and translational informations are strictly related. The evidence on which this suggestion relies has been obtained on other multiple nucleosome positioning/bent DNA systems. Before discussing this point in more detail, we devote the following section to the description of the information provided by multiple positioned nucleosomes on the actual helical period of the DNA on the nucleosomal surface.

6. Multiple nucleosomes offer a contribution to the solution of the linking number paradox

Why the variation of linkage influences the coiling of DNA around the histone octamer (as discussed in section 3)? The answer is probably in the variations of its helical twist. This is also the solution of the so-called "linking number paradox" of nucleosomal DNA.

Measurements of the linking number (Lk) of closed circular DNAs extracted from animal viruses [18,44] indicated a ΔLk near to 1 per nucleosome, in apparent contradiction with the X-ray crystal analysis that showed that in a nucleosome the double helix is wound about 1.8 turns of superhelix [45]. One would therefore expect that the Lk change is 1.8 and not the observed 1.0. Explanation of the paradox requires that the helical period of the DNA on the nucleosome and when free in solution differ [45,46]. Upon nucleosome formation, the helical repeat h (the number of base-pairs per unit winding number $h = N/\Phi$, where N is the total number of base-pairs involved and Φ is the winding number) changes from 10.6 per turn (47–49) to 10.17 (10). This value has been calculated with an external probe and is consistent with the observed change in linking number. Additional precise measurements have been obtained by the hydroxylradical footprinting technique showing an average helical periodicity of 10.18 ± 0.05 basepairs per turn for a nucleosome reconstituted on a 5S rRNA gene of *Xenopus borealis* [50]. The same analysis was extended to show the heterogeneity of the distribution of the helical periodicity inside the nucleo-

some (10.7 basepairs/turn in the central three turns, 10.0 in the flanking regions) [51].

The solution of the apparent topological paradox thus requires that the helical repeat changes, upon nucleosome formation, from the value in solution (10.6) to a value closer to 10.0 bp per turn [52–54]. The possibility that the twist of DNA does not actually change in DNA wrapped around a histone core has been raised [24,28].

We have reported above that on certain DNAs (as the *C. fasciculata* kinetoplast DNA (ref. [8]) or the *S. cerevisiae* 5S-gene ref. [9]) the histone octamers formation is highly favored by the phased repetition of defined sequence motifs.

This is predicted by the observation that certain motifs, i.e. $(A)_n$ and $(T)_n$ runs, assume specific positions relative to the octamer surface [11]. Correlating the periodical structural property that provides the rotational signals for the octamers localization on these DNAs with the strong preference of localization on these signals, information on the helical periodicity optimal for histone octamers formation is obtained. What is of interest here is the comparison of the period of the helical phase of the analysed DNAs with the experimentally observed period of the deposited octamers. A best fit analysis of the position of the octamers with the positions of the $(A)_n$ and $(T)_n$ runs (which preferentially attract octamers) has shown a periodicity of 10.26 ± 0.04. This calculation supports the interpretation that the linking number paradox is solved by the reduction of the helical repeat of DNA in the nucleosome [8].

The fact that a particular helical periodicity characterizes DNA on the nucleosomal surface, is connected with the definition of rotational and translational information, described in the following sections.

7. Rotational versus translational information

The effects on nucleosome formation of rotationally phased bendability signals due to particular combinations of base-pairs [8,10,14,16] are not confined to peculiar protozoan sequences. Multiple nucleosome positions have been observed on synthetic [16], natural [55–57] and engineered repetitive [58] DNAs. Analysis of the histone octamer formation *in vitro* and *in vivo* on *S. cerevisiae* 5S rRNA gene [9] has revealed that octamers *in vitro* and nucleosomes *in vivo* completely cover this genetic segment, occupying alternative positions characterized by a unique rotational helical phase. This result was obtained *in vitro* by mapping the reconstituted particles with ExoIII, Micrococcal nuclease-restriction, DNAse I and hydroxylradical digestion; and it was extended *in vivo* by the use of a novel technique that exploits the permeabilization of yeast to macromolecules by nystatin, allowing the import in living yeast of active enzymes [59]. Subsequent PCR-mediated analysis of the cleaved regions allows nucleotide-level *in vivo* nucleosome mapping [9].

The results support the notion that the rotational signals are a major determinant of nucleosome localization and establish that *in vivo* and *in vitro* (at least on this gene system) localization occurs following the same rules.

Multiple localizations 10 basepairs apart are observed and imply that each individual histone octamer can select among several quasi-isoenergetic positions; along these positions histone octamers *in vitro* (and nucleosomes *in vivo*) could enjoy facilitated rotational displacements. When the occupancy is multiple the free-energy variance between different positions is small (< 1 kcal/mol) [16,60], suggesting that changes in nucleosome positioning are obtained by shifting from one permissive position to the next along the same rotational phase. The expected consequence of these rotationally limited movements is the absence of randomly positioned nucleosomes, exactly what is observed [7–9,60]. Actually mobile nucleosomes have been described [7,61] and the possibility that the described displacements could represent a general behavior has been discussed [9,61]. Minor positions of nucleosomes on other systems have been reported [60,62], thus supporting this model.

Nucleosomes have been described which localize on a single defined position (reviewed by Grunstein [63], and references therein; see also Simpson [64], and references therein). Boundary effects, ancillary proteins, translational messages

intrinsic to the DNA sequence are usually invoked to explain these unique localizations.

The fact that multiple positions on a unique rotational phase have been observed suggests that for each of these positions the translational message is nothing but the equilibrium between the rotational messages present on each side of the center of symmetry on the histone octamer. In other words, the frequency of formation of a histone octamer on a given position on a DNA segment containing a long series of flexibility signals is determined by the algebraic sum of flexibility potentialities, relative to a dyad central position.

8. A regulatory mechanism based on nucleosome dynamics

Let us now consider together several observations described in the preceding sections, namely: (a) both the average repetion of the flexible DNA elements and the repetition of the phased histone octamers occurs every 10.26 bp, thus showing a cause–effect relationship between curvature and preferential deposition of octamers; (b) supercoiling (writhing) favors nucleosome formation in large, small or micro domains; (c) intrinsic curvature characterizes the DNAs on which multiple phased positions have been observed; (d) the equilibrium distribution of nucleosomes on intrinsic or supercoiling-induced curvatures may undergo low-energy-requiring displacements.

All these facts support the idea that nucleosome dynamics follows simple rotationally determined rules. In particular, on intrinsically curved sequences multiple locations among quasi-isoenergetic positions are observed.

The possibility of facilitated nucleosome displacement represents an important genetic regulatory possibility.

9. DNA topoisomerase I

In eukaryotes, DNA topoisomerase I is a ubiquitous and abundant enzyme, the enzymology of its reaction is known in great detail, but the biological function of its reaction (its phenotypes) is much less clear. DNA topoisomerase I binds to DNA, cleaves one phosphodiester bond, allows the revolution of one strand around the other (typically leading to relaxation by removal of negative supercoils), then reseals the cleavage and leaves the DNA in a topological state different from the one it started on [65]. The biological function of the enzyme is supposedly related to this relaxing activity and is probably a relevant one, given its abundance and the fact that supercoiling is a basic factor in the interaction with proteins (i.e. nucleosomes, TATA binding factor, the DNA topoisomerase I itself, DNA topoisomerase II, histone H1). However, mutants completely devoid of this enzyme have been characterized which are quite vital and lack clear phenotypical alterations.

A study of the regulation of this enzyme can help in defining its phenotypes and its function. In addition, the consideration that this enzyme does not act at random on DNA but reacts on sites which are defined but are devoid of a clear sequence consensus [65–70], makes it a good candidate for the study of the relationship between DNA conformation and protein interaction.

10. DNA topoisomerase I and DNA supercoiling: a regulation by topological feedback

We have analyzed the correlation of each step of the reaction of eukaryotic DNA topoisomerase I with the topology of its substrate. It was observed that:

(i) topoisomerization by the purified enzyme (from yeast, calf thymus, wheat germ) has a kinetic preference for nicking and closing torsionally stressed over relaxed DNA [71,72]. Unstressed DNA is a poor substrate for DNA topoisomerase I: it becomes reactive to the enzyme only at high enzyme/DNA ratios. These analyses showed that the perfectly relaxed state is the topological condition of lowest reactivity and that kinetic differences among different topological forms are observed between DNA molecules differing only for 1 linking number [73].

(ii) Also the cleavage step by DNA topoisomerase I was observed to be strongly supercoiling dependent [73]. The analysis, performed on several types of DNA substrates and with DNA topoisomerase I from various sources, revealed that a distinct topological condition, the complete absence of torsional stress, impedes the function of DNA topoisomerase I.

(iii) The two other steps of the reaction were analyzed as a function of the topology of the substrate: the binding of the enzyme to DNA and the catalytic constants (K_Cs) of topoisomerization [74]. It was observed that (a) the enzyme binds supercoiled DNA with slower kinetics than the relaxed form. Therefore, the preferential topoisomerization of supercoiled DNA is not due to the binding step. (b) The K_Cs of the topoisomerization reaction do not vary as a function of the topology of the substrate nor of its size (*ibidem*).

All these data lead to the conclusion that the preferential topoisomerization of torsionally stressed DNA is due to the higher availability on this topological form of DNA sites that allow the onset of the reaction, possibly by resident but inactive enzyme [74,75].

This view is coherent with the observation that expression of transfected DNA in eukaryotes depends upon DNA topology [76] and with the fact that a topological swivel is needed in transcription [77]. The set of data on the correlation DNA-topology/activity of DNA topoisomerase I supports the following overall mechanism for the reaction of eukaryotic DNA topoisomerase I: when a molecule of enzyme binds to topologically strained supercoiled DNA, fast topoisomerization occurs. The resulting relaxed DNA becomes unreactive; the enzyme remains stably bound on potentially active but actually inert DNA sites [72–74,78]. Activation occurs when the DNA molecule undergoes even minor changes of its conformation: variation of writhing as small as those that correspond to 1 ΔLk for a DNA molecule larger than 1 kb are sufficient to activate topoisomerization [73].

The series of topology-dependent steps defines a *topological feedback mechanism*: if cleavage and topoisomerization do not occur on relaxed DNA, it means that DNA topoisomerase I induces self-arrest on the topological form which is the final product of the reaction. This topological feedback mechanism provides a role for the enzyme and describes its regulation: every time DNA undergoes topological strain, DNA topoisomerase I is activated by the consequent conformational alteration and brings the DNA back to the unreactive relaxed state.

11. DNA topoisomerase I and DNA curvature

The observations reported above point to the relevance of the conformation of the reactive DNA sites and have suggested that DNA curvature might play a role in the interaction with DNA topoisomerase I.

Direct proof of the role of DNA conformation was obtained when each one of the steps of the reaction were analyzed on a DNA domain containing an intrinsically curved DNA sequence. Binding, cleavage and topoisomerization occur with high preferentiality on the curved segment relative to the rest of the DNA domain [75]. These studies have shown the importance of the local DNA conformation in the reaction of eukaryotic DNA topoisomerase I and have provided an experimental starting point for the identification of a DNA conformational consensus for reactivity.

The consensus motif was identified for the strongest DNA reactive sites, the ones which undergo cleavage even in the absence of superhelical strain, defined as "constitutive" sites, [74,78]. These sites can be described [78] by a set of Eulerian angular values (determined by Bolshoy et al. [79] that define the axial path of DNA (helical twist, deflection angle, direction) and of the orthogonal components of wedge (roll and tilt) and are abundant on intrinsically curved DNA (as observed on synthetic DNAs, on the kinetoplast curved segment from *C. fasciculata*, and in the curve-containing domain B of *Saccharomyces cerevisiae* ARS1 [78]. More specifically: the conformational consensus for the preferential cleavage sites is made of

(a) common features at the 5′ side of the cleaved sites; absence of common features at the 3′ side.

(b) The common features at 5′ are: a "V"-shaped set of twist angles; a series of negatively-directed deflection (wedge) angles; a "reverse-V"-shaped set of roll angles; no common motif for tilt angles.

(c) Sites with similar consensus reveal a hierarchy of reaction intensity, indicating a sequence context effect for DNA topoisomerase I reactivity.

The fact that a conformational consensus for preferential cleavage sites (that is: a defined alteration of the twist, a negative deflection, an alteration of the roll) can be defined [78], argues that DNA topoisomerase I recognizes a local, precisely defined DNA conformation and not an actual curvature or a potential bendability of the whole molecule.

12. The in-and-out problem in the accessibility of DNA information

In the case of the DNA–protein interactions which are not completely and exclusively determined by sequence information, the role of the local chemical environment is important. This is exemplified by eukaryotic DNA topoisomerases I and II whose recognition sequences in DNA are not strictly defined and are only partially related to a consensus (i.e. the *Drosophila* topoisomerase II [80] and DNA topoisomerase I, see above) and may even be relevant for certain restriction enzymes, like Eco RI*. In this case [81] the DNA conformational alteration induced by a neighbouring oligopurine tract on GAATTA enables an otherwise prohibited cleavage.

The definition of "chemical environment" for DNA is so complex that it results vague in the absence of further limitatory parameters. In closed DNA systems, a chemical property like the addition or subtraction of a nucleotide becomes a topological property. In fact, in a closed circle, the positions which will look outwards or inwards depend upon the exact geometry of the system, and therefore depends upon its topological state.

This is not a purely theoretical problem, given that proteins exist which interact with a defined face of the double strand (i.e. in the nucleosomes: the "$(A)_n$ look inwards" argument; the rotational information as a major determinant of the localization, etc. In the DNA topoisomerase I system: the preference for curved DNA and, in the curves, the selection of specific conformationally determined positions. The ubiquity of these two protein systems provides genetic relevance to this geometrical topic).

The nucleosomes. We have mentioned above the existence of natural DNA systems (*C. fasciculata* kinetoplast [8] and *S. cerevisiae* 5S-rRNA gene [9]) in which multiple nucleosome positioning with a unique rotational setting occur *in vitro* and, in *S. cerevisiae*, also *in vivo*. These systems provide an experimental opportunity for directly testing the validity of the "in-or-out" concept for DNA sequences in the nucleosomal complex formation. The approach is simple: the rotational information is by definition provided by a repetition of a given property as a function of the DNA helical period. This function needs not to be the simplest one (i.e. a repetition every helical turn, etc.) but could be complicated by scaling numerical factors in order to favor spirality, or could be palyndromic, or could be constituted by repeated groups of ordered phases separated (punctuated) by irregularities; according to this view, the dyad would be an internal punctuation in nucleosomes. We have mentioned above that the position of the dyad in nucleosomes could be determined not by an asymmetry but by the reciprocal weight (the equilibrium) of the two groups of rotational signals at the two sides of the geometrical center of the histone octamer.

Testing the relevance of the rotational information and the function of its composing elements in the deposition and localization of histone octamers is simple: it is sufficient to alter the information by an appropriate sequence insert. The insert should be characterized by the following properties: (a) size: $n + \frac{1}{2}$ helical turns; (b) conformational properties: absence of bendability signals in-phase respect to each other and/or to the external signals; (c) the position of its insertion: optimally, in the middle of a homogeneous and defined series of multiply positioned nucleosomes characterized by a unique rotational setting. This type of insertion was made on the

yeast 5S gene and the effect on nucleosomes of a part of the shift of the rotational signals from a phased to a new counter-phased positioning was analyzed (Buttinelli, Negri, Di Marcotullio and Di Mauro, submitted). An alteration of the pattern of the *in vitro* histone octamers and of the *in vivo* nucleosomes was observed that was clearly caused by the variation of the phasing signals. The major effect of an "$n + \frac{1}{2}$/turn" insert was the destabilization of a group of nucleosomes, namely the nucleosomes which bear the insert in central position. In these cases, it was shown that the protein particle is biased between two equally strong groups of conflicting signals (in-phase versus out-of-phase) on its two sides and cannot form. Interestingly, in the case of histone octamers on which the insert occupies a paracentral position, alternative positions are observed and differential patterns of nucleosome strings can be generated both *in vivo* and *in vitro*.

The positioning of the functional signals in-phase or in counter-phase is therefore relevant. In nucleosomes the DNA phase is located always on the same side of DNA, and the repetitive signal is always in or out (unless mutations are programmed to change this order). If the genetic function of the tract of DNA under consideration requires it, a specific position (i.e. a TATA sequence) could be conveniently located and be exposed or hidden.

DNA topoisomerase I. The very nature of the rotational information is the spatial distribution of certain groups of sequences. Thus, the variation (i.e. the rotational displacement) of this distribution actually changes the information. We have seen in the previous section that putting in counter-phase part of the rotational signals used in the localization of histone octamers changes the pattern of distribution of these protein particles. The relevance of the tridimensional organization of the substrate DNA molecule is also emphasized by the following observations on the DNA topoisomerase I reaction [82].

We have reported (section 10) that both supercoiling and a consensus local conformation related to DNA curvature activate topoisomerization. These properties are probably correlated. We have observed (*ibidem*) that the cleavage sites of eukaryotic DNA topoisomerase I along various curved and non-curved DNAs map on the same side of the double strand: the cleaved sites localize infact in an angular sector narrower than 90° (on non-curved tracts cleavages are dispersed and localize in any position encompassed between 0° and 360°).

Using a series of constructs bearing DNA inserts that have the same amount of curvature but that shift the planes of curvature by 72°, 144°, 216°, 288° and 360° (relative to the point of insertion in the recipient DNA) [83], we have analysed the effects of the insertion of a curved sequence and of the serial variation of the position of the curved tracts both on the reactivity of the cleavage sites pre-existing on the recipient DNA and on the topoisomerization rates. The results have shown that the insertion of a curve markedly changes the reaction properties of surrounding sequences (both for cleavage and topoisomerization) and correlates the variation of reactivity with the variation of the tridimensional context [82].

In particular, we have observed that the insertion of a curved tract perturbs the reactivity over at least nine helical turns both within the insert and in the flanking DNA, in correlation with the *angular distance* from the plane of curvature. These results demonstrate that the reaction is affected by geometrical parameters wider than the local structure (helical twist, deflection angle, etc.) of the cleaved site. This effect is explained as follows: on a straight linear DNA, no definition of "ins-or-outs" can be applied. On a curved DNA molecule, at the contrary, the amplitude of the curve determines a differential accessibility that in principle decreases increasingly with increasing curvature on the inner side of the double strand. The direction of the curvature defines which side of the double strand will be *in* or *out*. Therefore, the presence and the position of a curved tract in a sequence determines which nucleotides in the closely surrounding sequences will face inwards or outwards. The localization of a curve in the primary sequence potentially determines which elements of the rest of the molecule will be located "in-or-out" (that is: which sequences will face inwards or outwards) along the

writhed path of the DNA supercoiled molecules or along the relaxed path of the open ones.

In conclusion, the analysis of the DNA topoisomerase I reaction on curved DNAs embedded in various DNA tridimensional contexts, has shown the relevance of the differential accessibility and of the in-or-out problem for substrate DNA sequences [82] even in the absence of higher orders of structural organization.

13. The integrating function of the free energy of supercoiling

If a nucleosome releases free energy upon its removal, the local conformational effects of the free energy so made available are a promoting factor for the formation of a replacing nucleosome, as described in large [18,84], small [19,24] or micro [20] DNA domains. A microdomain is defined by the fact that its size is so small that it does not allow onset of writhing (*ibidem*); (in sequences devoid of intrinsic curvatures, DNA segments of ≈ 180 bp do not undergo writhing and the linkage reduction only induces variation of the helical period). It was shown [20] that this variation is necessary for formation of histone octamers in the topologically highly constrained micro domains. The relevance of this phenomenon is in the fact that also on chromatin fibers *in vivo* the DNA substrate for formation of histone octamers is supposedly constrained by the surrounding particles.

Thus, a simple scheme that describes the fate of the free energy made available by the removal of a nucleosome is: a particle leaves, a particle enters. According to the basic equation of DNA topology ($Lk = T + W$), a variation of the linking number is distributed in variations of the twist and of the writhe values.

Nucleosomes writhe DNA around themselves; therefore, the interest of nucleosomes in this equation is devoted mostly to the W term.

DNA topoisomerase I opens and closes DNA, and changes its linking number; even if it is responsive to the local twist and to the DNA curves (section 11) its major effect on the equation is on the first term: Lk.

RNA polymerase II. The twin domain model of transcription [77] describes the topological modifications induced by transcription in closed DNA domains. *In vitro* transcription studies using RNA polymerase II and purified components [85,86] have shown both the activation due to supercoiling of the DNA template and the localization of conformational alterations on the promoter regions. In addition it was observed that in the yeast and animal promoters analysed, conformational modifications induced by supercoiling localize in upstream activator sequences, in the TATA sequence and in sites immediately upstream of the RNA initiation sites [87–89]. The conformation of functionally relevant sites may change as a function of sequence mutations that have taken place elsewhere, showing that the conformational behaviour of the whole promoter is concerted and indicating the transmission *in cis* of topological effects in RNA polymerase II promoters [89]. Evidence for a cause–effect relationship between topology of closed DNA domains and *in vitro* activation of transcription has been obtained [89–91].

Thus, topology of closed domains is strictly linked to transcription. At initiation, RNA polymerase enters the double strand and locally dramatically changes the remaining term of the equation: T.

When a nucleosome comes off a promoter, the locally available free energy may thus be used by three major elements of the transcription machinery: RNA polymerase II, nucleosomes, DNA topoisomerase I; possibly by others. RNA polymerase locally changes the twist, nucleosomes change the writhe, DNA topoisomerase I changes the linking number. Free energy can be used in any possible way: it will be interesting to test which part of this process undergoes kinetic competition and which part, on the contrary, occurs in cooperative terms.

Acknowledgement

This work was supported by Piani Finalizzati Ingegneria Genetica e Biotecnologie (CNR, Italy).

References

[1] A. Almer and W. Hörz, EMBO J. 5 (1986) 2681.
[2] R.H. Almer, A. Hinnen and W. Hörz, EMBO J. 5 (1986) 2689.
[3] H. Richard-Foy and G.L. Hager, EMBO J. 6 (1987) 2321.
[4] B. Piña, D. Barettino, M. Truss and M. Beato, J. Mol. Biol. 216 (1990) 975.
[5] B. Piña, V. Brüggemeier and M. Beato, Cell 60 (1990) 719.
[6] L.A. Freeman and W.T. Garrard, Crit. Rev. Eukaryotic Gene Expres. 2 (1992) 165.
[7] S. Pennings, G. Meersseman and E.M. Bradbury, J. Mol. Biol. 220 (1991) 101.
[8] G. Costanzo, E. Di Mauro, G. Salina and R. Negri, J. Mol. Biol. 216 (1990) 363.
[9] M. Buttinelli, E. Di Mauro and R. Negri, Proc. Natl. Acad. Sci. USA 90 (1993) 9315.
[10] H.R. Drew and A.A. Travers, J. Mol. Biol. 186 (1985) 773.
[11] S.C. Satchwell, H.R. Drew and A.A. Travers, J. Mol. Biol. 191 (1986) 659.
[12] A.A. Travers and A. Klug, Phil. Trans Roy. Soc. London 317 (1987) 537.
[13] A.A. Travers, Ann. Rev. Biochem. 58 (1989) 427.
[14] E.N. Trifonov and J.C. Sussman, Proc. Natl. Acad. Sci. USA 77 (1980) 3816.
[15] E.N. Trifonov, CRC Crit. Rev. Biochem. 19 (1985) 89.
[16] T.E. Shrader and D.M. Crothers, Proc. Natl. Acad. Sci. USA 86 (1989) 7418.
[17] T.E. Shrader and D.M. Crothers, J. Mol. Biol. 216 (1990) 69.
[18] J.E. Germond, B. Hirt, P. Oudet, M. Gross Bellard and P. Chambon, Proc. Natl. Acad. Sci. USA 72 (1975) 1843.
[19] I. Goulet, Y. Zivanovic, A. Prunell and B. Revet, J. Mol. Biol. 200 (1988) 253.
[20] R. Negri, G. Costanzo, S. Venditti and E. Di Mauro, J. Mol. Biol. 207 (1989) 615.
[21] J.C. Wang, L.J. Peck and K. Becherer, Cold Spring Harbor Simp. Quant. Biol. 47 (1983) 85.
[22] D. Shore and R.L. Baldwin, J. Mol. Biol. 170 (1983) 975.
[23] D. Shore and R.L. Baldwin, J. Mol. Biol. 170 (1983) 983.
[24] Y. Zivanovic, I. Goulet, B. Revet, M. Le Bret and A. Prunell, J. Mol. Biol. 200 (1988) 267.
[25] W. Keller, Proc. Natl. Acad. Sci. USA 72 (1975) 4876.
[26] R.H. Morse and C.R. Cantor, Proc. Natl. Acad. Sci. 82 (1985) 4653.
[27] R.H. Morse and C.R. Cantor, Nucl. Ac. Res. 14 (1986) 3243.
[28] R.H. Morse and R.T. Simpson, Cell 54 (1988) 285.
[29] A. Hamiche and A. Prunell, J. Mol. Biol. 228 (1992) 327.
[30] W.G. Turnell and A.A. Travers, Meth. Enzymol. 212 (1992) 387.
[31] D. Boffelli, P. De Santis, A. Palleschi and M. Savino, Biophys Chem. 39 (1991) 127.
[32] B. Piña, M. Truss, H. Ohlenbusch, J. Postma and M. Beato, Nucl. Ac. Res. 18 (1990) 6981.
[33] C.H. Hsieh and J.D. Griffith, Cell 52 (1988) 535.
[34] R. Kornberg, Nature 292 (1981) 579.
[35] T.E. Palen and T.R. Cech, Cell 36 (1984) 933.
[36] F. Strauss and A. Varshavsky, Cell 37 (1984) 889.
[37] R. Benezra, C.R. Cantor and R. Axel, Cell 44 (1986) 697.
[38] F. Thoma and M. Zatchej, Cell 55 (1988) 945.
[39] P. Forte, L. Leoni, B. Sampaolese and M. Savino, Nucl. Ac. Res. 17 (1989) 8683.
[40] M. Buttinelli, L. Leoni, B. Sampaolese and M. Savino, Nucl. Ac. Res. 19 (1991) 4543.
[41] A.A. Travers, Cell 60 (1990) 177.
[42] M.E. Hogan, T.F. Rooney and R.H. Austin, Nature 328 (1987) 554.
[43] S.C. Satchwell and A.A. Travers, EMBO J. 8 (1989) 229.
[44] R.T. Simpson, F. Thoma and J.M. Brubaker, Cell 42 (1985) 799.
[45] J.T. Finch, L.C. Lutter, D. Rhodes, R.S. Brown, B. Rushton, M. Leavitt and A. Klug, Nature 269 (1977) 29.
[46] A. Klug and L.C. Lutter, Nucl. Ac. Res 9 (1981) 267.
[47] J.C. Wang, Proc. Natl. Acad. Sci. USA 76 (1979) 200.
[48] D. Rhodes and A. Klug, Nature 286 (1980) 573.
[49] L.J. Peck and J.C. Wang, Nature 292 (1981) 375.
[50] J.J. Hayes, T.D. Tullius and A.P. Wolffe, Proc. Natl. Acad. Sci. 87 (1990) 7405.
[51] J.J. Hayes, D.J. Clark and A.P. Wolffe, Proc. Natl. Acad. Sci. 88 (1991) 6829.
[52] J.H. White and W.M. Bauer, Proc. Natl. Acad. Sci. USA 85 (1986) 772.
[53] J.H. White and W.M. Bauer, Cell 56 (1989) 9.
[54] A. Klug and A.A. Travers, Cell 56 (1989) 10.
[55] X.Y. Zhang and W. Hörz, J. Mol. Biol. 176 (1984) 105.
[56] W. Linxweiler and W. Hörz, Cell 42 (1985) 281.
[57] M.F. Clarke, P.C. Fitzgerald, J.M. Brubaker and R.T. Simpson, J. Biol. Chem. 260 (1985) 12394.
[58] J.C. Hansen, J. Ausio, V.H. Stanik and K.E. van Holde, Biochemistry 28 (1989) 9129.
[59] S. Venditti and G. Camilloni, Mol. Gen. Genet., 242 (1994) 100.
[60] F. Dong, J.C. Hansen and K.E. van Holde, Proc. Natl. Acad. Sci. USA 87 (1990) 5724.
[61] G. Meersseman, S. Pennings and E.M. Bradbury, EMBO J. 11 (1992) 2951.
[62] J.J. Hayes and A.P. Wolffe, Proc. Natl. Acad. Sci. 90 (1993) 6415.
[63] M. Grunstein, Ann. Rev. Cell. Biol. 6 (1990) 643.
[64] R.T. Simpson, Progr. Nucl. Ac. Mol. Biol. 40 (1991) 143.
[65] J.C. Wang, Ann. Rev. Biochem. 54 (1985) 665.
[66] K.A. Edwards, B.D. Halligan, J.L. Davis, N.L. Nivera and F.L. Liu, Nucl. Ac. Res. 10 (1982) 2565.
[67] C.C. Shen and C.K.J. Shen, J. Mol. Biol. 212 (1990) 67.
[68] M.D. Been, R.R. Burgess and J.J. Champoux, Nucl. Ac. Res. 12 (1984) 3097.
[69] B.J. Bonven, E. Gocke and O. Westergaard, Cell 41 (1985) 541.

[70] H. Busk, B. Thomsen, B. Bonven, E. Kjeldsen, O.F. Nielsen and O. Westergaard, Nature 327 (1987) 638.
[71] M.T. Muller, Biochim. Biophys. Acta 824 (1985) 263.
[72] G. Camilloni, E. Di Martino, M. Caserta and E. Di Mauro, Nucl. Ac. Res. 16 (1988) 7071.
[73] G. Camilloni, E. Di Martino, E. Di Mauro and M. Caserta, Proc. Natl. Acad. Sci. USA 86 (1989) 3080.
[74] M. Caserta, A. Amadei, G. Camilloni and E. Di Mauro, Biochemistry 29 (1990) 8152.
[75] M. Caserta, A. Amadei, E. Di Mauro and G. Camilloni, Nucl. Ac. Res. 17 (1989) 8521
[76] H. Weintraub, P.F. Cheng and K. Conrad, Cell 46 (1986) 115.
[77] L.F. Liu and J.C. Wang, Proc. Natl. Acad. Sci. 84 (1987) 7024.
[78] G. Camilloni, M. Caserta, A. Amadei and E. Di Mauro, Biochim. Biophys. Acta 1129 (1991) 73.
[79] A. Bolshoy, P. McNamara, R.E. Harrington and E.N. Trifonov, Proc. Natl. Acad. Sci. USA 88 (1991) 2321.
[80] M.T. Howard, M.T. Lee, T. Hsieh and J.D. Griffith, J. Mol. Biol. 217 (1991) 53.
[81] S. Venditti and R.D. Wells, J. Biol. Chem. 266 (1991) 16786.
[82] R. Perini, M. Caserta and E. Di Mauro, J. Mol. Biol. 231 (1993) 634.
[83] C.M. Collis, P.L. Molloy, G.W. Both and H.R. Drew, Nucl. Ac. Res. 17 (1989) 9447.
[84] H.G. Patterton and C. van Holte, J. Mol. Biol. 229 (1993) 623.
[85] F. Carnevali, M. Caserta and E. Di Mauro, Nucl. Ac. Res. 10 (1982) 3195.
[86] F. Carnevali, M. Caserta and E. Di Mauro, J. Mol. Biol. 165 (1983) 59.
[87] G. Camilloni, F. Della Seta, R. Negri and E. Di Mauro, J. Biol. Chem. 261 (1986) 6145.
[88] S. Venditti, M. Caserta, E. Di Mauro and G. Camilloni, Biochim. Biophys. Acta 951 (1988) 139.
[89] F. Della Seta, G. Camilloni, S. Venditti and E. Di Mauro, J. Biol. Chem. 263 (1988) 15888.
[90] G. Camilloni, F. Della Seta, A.G. Ficca and E. Di Mauro, Mol. Gen. Genet. 204 (1986) 249.
[91] G. Camilloni, F. Della Seta, R. Negri, A.G. Ficca and E. Di Mauro, EMBO J. 5 (1986) 763.