Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Multiscale simulation of DNA

Pablo D Dans[1,2], Jürgen Walther[1,2], Hansel Gómez[1,2] and Modesto Orozco[1,2,3]

DNA is not only among the most important molecules in life, but a meeting point for biology, physics and chemistry, being studied by numerous techniques. Theoretical methods can help in gaining a detailed understanding of DNA structure and function, but their practical use is hampered by the multiscale nature of this molecule. In this regard, the study of DNA covers a broad range of different topics, from sub-Angstrom details of the electronic distributions of nucleobases, to the mechanical properties of millimeter-long chromatin fibers. Some of the biological processes involving DNA occur in femtoseconds, while others require years. In this review, we describe the most recent theoretical methods that have been considered to study DNA, from the electron to the chromosome, enriching our knowledge on this fascinating molecule.

Addresses
[1] Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain
[2] Joint BSC-IRB Research Program in Computational Biology, Baldiri Reixac 10-12, 08028 Barcelona, Spain
[3] Department of Biochemistry and Molecular Biology, University of Barcelona, 08028 Barcelona, Spain

Corresponding author: Orozco, Modesto (modesto.orozco@irbbarcelona.org)

Current Opinion in Structural Biology 2016, **37**:29–45

This review comes from a themed issue on **Theory and simulation**

Edited by **Modesto Orozco** and **Narayanaswamy Srinivasan**

For a complete overview see the Issue and the Editorial

Available online 19th December 2015

http://dx.doi.org/10.1016/j.sbi.2015.11.011

Introduction

DNA is a long, flexible, and structurally polymorphic molecule, and its theoretical description is challenged by its intrinsic multiscale nature. DNA is thus a complex multi-resolution molecule whose theoretical study requires moving in an extremely wide range of sizes and time scales (Figure 1). If extended, the nuclear DNA existing in a human cell would measure one meter, while the distance between base pairs (bp) is in the Å-scale (10^{-10} m). Some changes in DNA, like those aging-related occur in the year time-scale (10^{8}–10^{10} s), others, like the chromatin reorganization along cell cycle, happen in the day time-scale (10^{5} s); the local breathing of nucleobases occurs in the millisecond range (10^{-3} s), while electronic rearrangements take place in the sub-femtosecond time-scale ($<10^{-15}$ s).

During the last years we have witnessed the development of a wide repertoire of theoretical methods that aimed to reproduce the properties of DNA, either isolated or protein bound. Even if primitive, these methods allow researchers to consider the DNA at different resolution levels, and provide information of great value on the structure, dynamics, and interactions of this fascinating molecule. We will briefly summarize some of these most recent theoretical approaches, focusing our analysis on the contributions of the last three years, when the field has experienced a significant improvement.

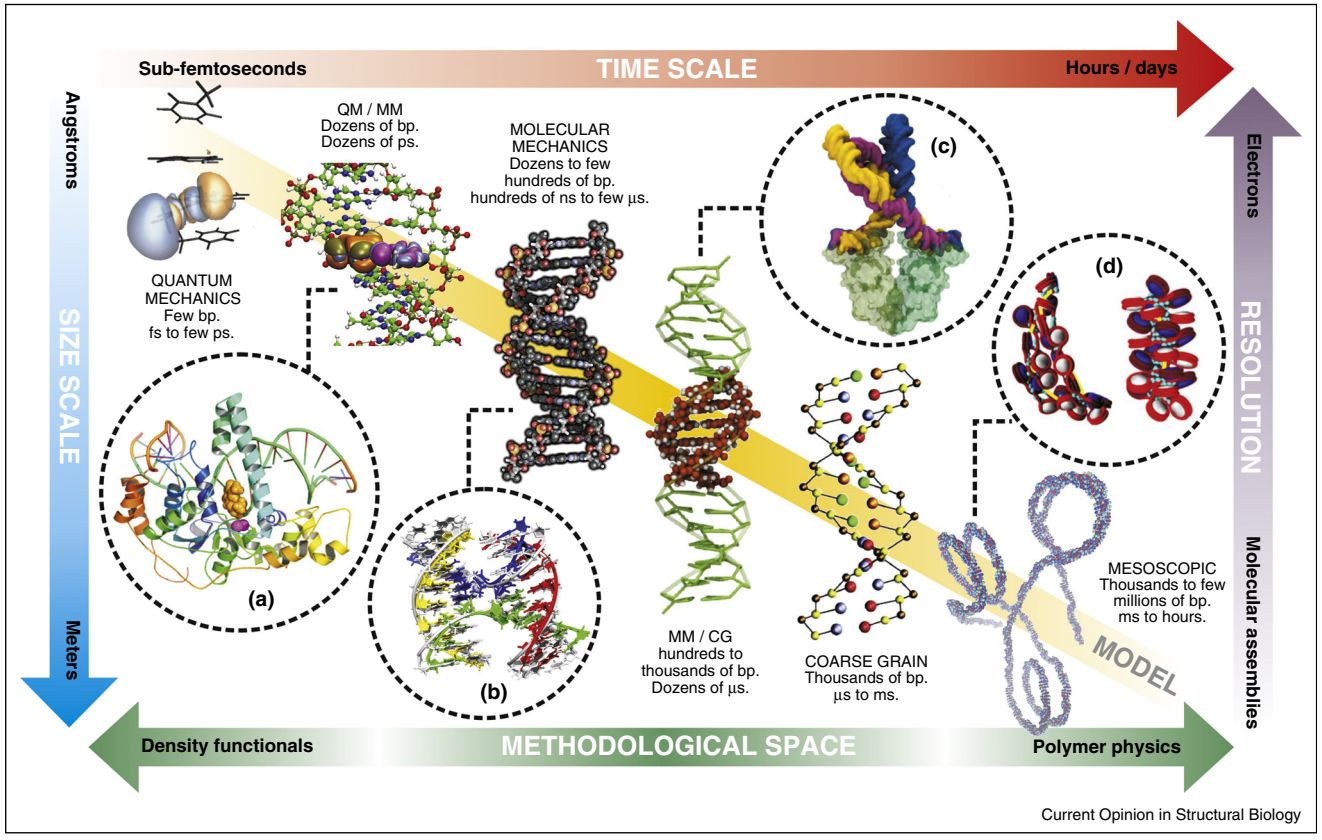
For the sake of simplicity, throughout this manuscript we will classify theoretical methods in four groups, according to their level of resolution (Figure 1): firstly, electronic, secondly, atomistic, thirdly, coarse grained, and lastly, mesoscopic. It is worth noting that moving in the resolution space means moving also in the methodological space, since the basic physical models underlying the different approaches vary considering the resolution level, from quantum mechanical calculations when dealing with electronic problems, to ideal fiber models when studying chromatin (Figure 1).

Electronic studies

Quantum Mechanics (QM) provides a theoretical framework where in principle, high quality results can be obtained for any system, without any *ad hoc* parameterization. Unfortunately, QM methods are very costly, even when the most efficient programs like SIESTA [1] or BIGDFT [2,3] and the fastest supercomputers are available. Use of QM methods in the nucleic acids world is then quite limited to the study of small model systems, where QM calculations are feasible, and to its combination with classical atomistic MM (Molecular Mechanics) methods to study processes of quantum nature involving macro-molecules, which cannot be studied only at the MM level (see Section ''Atomistic studies'').

Hobza's and Šponer's groups have worked for many years on the use of high level QM to describe the basic nature of nucleotide interactions using model systems [4,5]. Their work provided good understanding of the nucleic acids interactions, and benchmarks datasets for force fields validations. Recent examples of this type of works

Figure 1



Global scheme illustrating the intrinsic multiscale nature of DNA. The models and applications discussed thorough this work are sorted in this scheme according to five dimensions: firstly, the time scale that each model is able to sample; secondly, the size of the systems; thirdly, the methodological space; fourthly, the resolution of the models available at each level; and, depicted in the diagonal, lastly, a representation of the models used to tackle DNA properties at different levels. Each representation is accompanied by a legend showing an approximate range of applicability, and limitations. Four applications are highlighted; **(a)** the combined QM/MM work from the Magistrato's group, were a protein–DNA complex with a DNA lesion is studied in detail (adapted with permission from [18]). **(b)** Holliday-junction simulated with the new parmBSC1 refined force-field for DNA simulations, adapted from [44••]. **(c)** LacI-DNA dynamics by multiscale simulations using the SIRAH force-field from Pantano's group (adapted with permission from [125••]). **(d)** The model from Schlick and coworkers was used to study the chromatin fiber dynamics. Chromatin fibers in the canonical and hairpin-like conformations are depicted (adapted with permission from [156]).

involved detailed analysis of backbone rotamers in DNA [6] and RNA [7] and the **impact of ion polarization** on the stabilization of certain quadruplexes, which are very difficult to represent by means of classical force-fields [8]. The same groups used also QM to study another complicate DNA motif: the complex of two G-DNA quartets with a monovalent cation. Calculations revealed important differences between MM and QM descriptions of the system, and predicted the 5′-anti-anti-3′ GpG dinucleotide step to be the most stable one, closely followed by the 5′-syn-anti-3′ step, in agreement with the experiments [9•]. However, the study also illustrated the problems of using ultra-reduced systems in these model QM calculations. Similar conclusions on the strength and limitations of QM theory applied to nucleic acids were obtained by the same group in their study of the Sarcin–Ricin internal loop [10]. Following similar ideas and approaches other groups have recently explored specific details of nucleobase interactions in certain types of DNA. For example, Phan's group characterized the guanine base stacking in G-quadruplex nucleic acids [11], and Parker and coworkers described the nature of π–π stacking of nucleobases using symmetry adapted perturbation theory (SAPT), finding good predictive power, but detecting again the limitations implicit to the reduced size of the model systems [12••]. Nawort and coworkers used DFT theory to analyze the impact of the presence of 2′thiouridine and degradation products in tRNA on the fidelity of the translation process [13], and Brovarets and Hovorun used Bader's theory and DFT or MP2 calculations to characterize the probability of occurrence of **ground-state tautomerization** of the G-C Watson–Crick base pair by a double proton transfer (DPT) [14]. The latter is a process

that was suggested as a source of spontaneous mutations in DNA, but these accurate QM calculations [14] provided convincing evidence that in reality G·C double proton transfer is too rare to have any important role as a source of point mutations. All these studies (selected among many others not cited here due to space limitations) have shown the potential of high-level QM calculations as a source of detailed information on nucleotide conformation, and on specific interactions involving nucleobases, but we are still far from the time when high-level QM methods could be used to study the dynamic properties of long pieces of solvated DNA.

When the QM level of theory is reduced, it can be possible to introduce **entire (small) nucleic acids** and describe them at the dynamic level. Thus, Arcella *et al.* [15] used *ab initio* Car-Parrinello Molecular Dynamics (CPMD) with Grimme's corrections to study the changes in covalent topology in a small DNA hairpin when subjected to electrospray ionization mass spectrometry (MS-ESI). The authors used an extended (up to 100 μs) classical MD sampling to select hundreds of snapshots, which were next subjected to CPMD for several picoseconds (total QM sampling in the multi-nanosecond regime). Very encouraging, during the QM dynamics they sampled a significant number of proton transfer processes, refusing then the 'dogma' that covalent structure is unaltered in a MS-ESI experiment.

Even the fastest QM methods are inefficient to study long fragments of DNA, but QM is needed if we are interested in the study of processes where significant rearrangement of the electron distribution occurs, which generates a strong limitation in our current research capabilities. Fortunately, in those cases where the electron redistribution can be localized in a small portion of the DNA, we can use **hybrid QM/MM approaches**, where a small part of the system is treated at the QM level, while the rest is represented classically. A clear example of use of QM/MM methods in the DNA world is found in the study of reactivity. For example in a recent work Molina *et al.* [16], combined QM and QM/MM calculations with X-ray time-course data to describe the complex catalytic mechanism of a restriction enzyme [16]. The authors found a new reaction paradigm for nucleases, where the nucleolytic reaction proceed in two steps, with the attacking water molecule being activated by the targeted phosphate group — the rate limiting step — to later protonate the O3′ atom and break the phosphodiester bond.

QM/MM calculations require of an efficient QM method, which explains the popularity of semiempirical Hamiltonians (often recalibrated to study reactions involving nucleic acids [17]). Recent successful example of **semiempirical QM/MM calculation** applied to nucleic acids include the study of the catalytic mechanism of the Human Flap Endonuclease (hFEN1) [18], or the excellent work by Tuñon's group [19] on the mechanism of action of N6-adenosine methyltransferase, for which they described an ordered stepwise mechanism: methylation followed by proton abstraction of the targeted N6 atom of the adenine. However, it should be mentioned that even if the semiempirical methods are convenient in terms of sampling, and they are able to reproduce kinetic properties, they might significantly distort the calculated potential and free energy surfaces, as recently found by Mlýnský *et al.*, in their description of the catalytic mechanism of Hairpin ribozyme [20]. Clearly, carefully calibration of semiempirical methods is required in those cases where these methods are the only alternative in QM/MM MD studies.

Car-Parrinello molecular dynamics (CPMD) has been also implemented in the context of QM/MM studies of nucleic acids. An example is the study of intra-strand oxidative crosslink lesions in DNA [21,22]. We cannot ignore, however, that CPMD has also known caveats that can limit its accuracy, making it desirable to move to higher levels of QM theory for the inner part of the QM/MM calculation. A few examples in this direction have been recently published. For example, Zhang and coworkers used DFT (Density Functional Theory) Born–Oppenheimer molecular dynamics to study the catalytic mechanism of the nucleotidyl transfer reaction in human DNA polymerase kappa. They described the activation of the 3′-OH primer terminus, the following associative nucleotidyl transfer reaction and accordingly, were able to explain the bypass of major benzo(a)pyrene-derived dG lesion by the enzyme [23].

Another case where the QM level of theory is required, and where coupling between MM and QM description is needed, is the study of charge transfer (CT) in DNA. It is widely known that the overlapping π system of stacked nucleobases can mediate the transfer of electrical charges (both electrons and electron holes) over long distances. The study of this process has a large interest to understand its impact on biological processes as DNA repair, and can be of paramount importance in the field of DNA nanotechnology, but it is handicapped by the need of using a QM representation on a large, flexible, and highly coupled system. Different authors have tried to circumvent the problem by performing QM calculations using MM-derived ensembles. As a recent example, Lech and coworkers [24] studied the effective electronic coupling (V) in different G-motives at the INDO/S level of theory. In their study, they considered 1000 G-tetrad (G4) models built from a MM MD-derived ensemble and found that the G-tetrad orientation plays a key role in the electron hole transport within the π stacks. On the other hand, Livshits and coworkers studied the charge transfer properties in G4-DNA molecules adsorbed on a mica surface, which they experimentally observed to transport significant current over long distances (>100 nm) [25].

More importantly — and based on their theoretical calculations — they pointed out to a thermally activated hopping between multi-tetrad segments as the physical mechanism that explains the long-range conductivity. In another recent work, Bacolla *et al.* performed multiple Ionization Potential (IPs) calculations at the QM(DFT)/MM level over a set of structures sampled with classical MD simulations, to study the sequence context-dependent mutagenesis at mononucleotide repeats (A-tracts and G-tracts). Interestingly, their work suggests a key role for electron transfer in sequence-dependent mutagenesis [26]. With similar objectives, others authors have followed different approaches, for example, Kubar and Elstner proposed a multi-scale method combining a non-adiabatic propagation scheme and a linear scaling QM approach in a QM/MM coupling framework [27] to study charge hopping in a double-stranded DNA sequence.

A last area of DNA research where QM level is clearly required is the prediction of the **photophysical and spectroscopic properties**. The latter imply the evaluation of not only the ground, but also the excited-states of nucleotides, which is out of the possibilities of MM calculations. Very recently, for example, a combination of MD and QM (DFT and Time Dependent DFT (TD-DFT))/MM description of the excited states have been used to recover the UV absorption and Electronic Circular Dichroism (ECD) spectra of different DNA sequences. This opens a way to connect the ECD signals to specific structural patterns, widening the range of applicability of this spectroscopic technique. Using a QM(DFT)/MM scheme, Spata and Matsika were able to compute the UV absorption and ECD spectra of an adenine-based oligonucleotide [28], finding that the mixing between charge-transfer and excited states properties is essential to explain photophysics in DNA. Similarly, Gatusso *et al.* have recently modeled the ECD spectra of different double helix B-DNA sequences [29]. The authors proposed a general method consisting on the combination of atomistic MD to have a reasonable sampling of the configuration space with a QM/MM coupling scheme to obtain the properties of excited states of single chromophores. Finally, Zelený *et al.* performed QM(DFT)/MM surface hopping dynamic calculations to study the photophysical properties of cytosine and guanine [30], finding significant differences between the decay rates of the photo excited states corresponding to G and C. Interestingly, while the DNA environment does not hamper the photo-deactivation of cytosine, major deactivation happens for G, which is explained by the dramatic reduction of the out-of-plane motions of its NH_2 group, when inserted in duplex DNA [30].

Atomistic studies

There are many cases of interest where the electronic degrees of freedom of DNA can be ignored, and the molecule can be represented as a set of atoms whose interactions are approximated by simple classical expressions, which are parameterized to reproduce experimental observables or high-level QM calculations. This severe simplification allows dramatically accelerating the calculations [31,32,33•], and is widely used in theoretical studies of DNA.

The accuracy of atomistic classical MD simulations is determined by: firstly, the similarity between the simulated and the real system; secondly, the quality of the sampling; and lastly, the accuracy of the force-field. Major efforts in software and hardware development allowed the extension of the size of the simulated models, making them closer to real systems. Furthermore, dramatic improvement in sampling has been achieved by either increasing the length of the individual trajectories, or the number of collected replicas [33•,34]. We can expect that this tendency will continue in the near future, and most likely the force-field inaccuracy will become the Achilles heel of atomistic MD. The **refinement of force-fields** should then be considered a priority in the field.

Second generation force-fields, such as AMBER parm99 [35], were the prevalent ones for a decade, but started to show major problems around ten years ago [36] when computers allowed us to perform multi-nanosecond simulations, evidencing errors that did not appear in shorter trajectories. Re-parameterization efforts provided corrected force-fields, such as parmbsc0 [37], which has allowed to use MD to simulate DNA [31,32,38,39,40•]. Nevertheless, as the simulation regime approached the μs regime and newer systems were analyzed, errors in these improved force-fields emerged. These errors include unrealistic instability at the ends of the helix, under-twisting, or corruption of some non-canonical structures in long trajectories [33•,38,40•,41–43,44••]. This has encouraged further refinements, which are becoming now available to the community. For example, the Czech group has introduced specific corrections to improve sampling around χ and ε/ζ degrees of freedom [42,43]. In parallel, MacKerell's group has developed an updated version of the CHARMM force field (CHARMM36), which corrected several inaccuracies of previous releases of this family of force-fields [45]. More recently, our group has presented parmbsc1, a new force-field validated with more than 140 μs of trajectories, covering more than 100 different DNA structures [44••].

There are, however, some intrinsic problems that even the most recent force-fields should face, and whose correction is far from trivial. The most important one is linked to the simplicity of the non-bonded potentials. For example, Chen and Garcia have suggested that stacking is overestimated by AMBER family of force-fields [46], an idea that has been supported by other authors [47], which compared theoretical and experimental estimates of stacking free energy of nucleobases,

finding that the force-field overestimates stacking by $\sim$1.5 kcal mol^{-1}. Without arguing on the validity of the results, some caution is needed in their interpretation since: firstly, differences around 1 kcal mol^{-1} are probably within the range of accuracy of a classical force-field; secondly, experimental numbers are extremely noisy [47]; and lastly, inference of an experimental observable from atomistic simulations is always dependent on the arbitrary labeling of ''bound'' and ''unbound'' states in the simulations [48]. In any case, if such overestimation exists, its impact on the formation of DNA structures should be clarified, since the geometry of a freely stacked base pair is very different to that found in DNA. Some efforts on determining experimentally stacking free energy in DNA duplex environment were published 15 years ago, but again large uncertainties appear in the reported values. As an example, performing similar experiments, Santa Lucia and coworkers obtained a free energy of stacking for AC of around -1 kcal mol^{-1} [49], while Kool and coworkers obtained a value around -2 kcal mol^{-1} [50]. Recalibration of stacking, if needed, should start from a careful characterization of the physical origins of any potential bias, which can be related to a bad balance of hydrophobic/hydrophilic interactions, to a poor electrostatic model for nucleosides, to incorrect van der Waals terms for nucleobases, or to the intrinsic shortcomings of a pair-wise additive spherically shaped nonbonded potential. In any case, at least in our hands, a simple scaling of van der Waals parameters [46] produced incorrect representations of DNA duplexes and we cannot recommend it.

Šponer's group has explored in a series of articles the intrinsic errors associated to the use of classical force-fields [6–8,9•,51], pointing towards the neglect of a specific polarization term as one of the major source of uncertainties in current classical simulations. It is impossible to argue against the importance of polarization [52], but the field has been very reluctant to use polarized force-fields, not only because of the extra-cost on the calculation, but also because the final results were not brilliant. During these last years, remarkable advances have been made in improving polarized force-fields by McCammon and Pande [53], and by Roux and MacKerell [54•], who developed an efficient polarization algorithm based on Drude's oscillator. In our hands, this force-field provides good representation of the DNA duplex in the 10^{-1} μs regime, but at the expense of extra computation cost. MacKerell and coworkers have used their polarized DNA force-field to study, for example, fine details of DNA–ion interactions, or the electronic mechanism of base flipping [54•, 55–57]. In our opinion, after decades of theoretical and methodological discussion [52], polarized force-fields like those developed by the CHARMM-community are reaching maturity, and should be seriously considered for MD simulations of DNA.

Though imperfect, current pair-wise additive force-fields have been widely used to study many aspects of DNA. For example, the ABC consortium has collected microsecond-long trajectories of all the unique tetramer sequences in B-DNA, characterizing the sequence-dependent **physical properties** of duplex DNA under physiological conditions [40•]. One of the unexpected results emerging from this massive study is the existence of large non-harmonic movements affecting some base pair steps (bps) in certain tetramer environments. This result argue the prevalent idea that DNA deformation can be described by means of near-neighbor harmonic models [40•]. An in-depth analysis of non-harmonic deformations in DNA [41,58] characterized the atomistic mechanisms of this movement, the role of ions in DNA polymorphism, and the surprising correlation between apparently disconnected degrees of freedom in the DNA.

The idea that B-DNA is polymorphic and that different states can coexist in its equilibrium ensemble has been widely explored in the last years by different authors [30,33•,58–63], putting special emphasis on the base fraying [59–63], due to its role in DNA recognition, repair, and strand slippage. The impact of DNA polymorphism in DNA allosterism [64] has been also the subject of the work of several groups [62,65], while others centered their efforts in understanding the origin of DNA curvature [66]. Minicircles, a specially curved form of DNA traditionally challenging for the field due to the high mechanical tension that they can incorporate, have been recently revisited by different authors who explored the origins of superhelicity and the limits of the elastic response of DNA to distortion [66,67].

Different groups have made efforts to study not only small near-equilibrium relaxations of DNA but to follow **large conformational transitions**, a very challenging topic considering the difficulty to sample complex and slow processes. In an interesting article, Yang *et al.* used biased MD simulations to study the Watson–Crick to Hoogsteen transition in duplex DNA [68], finding a complex landscape with at least two major pathways, which explains well experimental data. This transition was also the focus of the work of Brooks and Al-Hashimi [69], who determined the role of cytosine protonation in the transition combining experimental measures and constant pH simulations. Sagui and coworkers studied B to Z transitions in DNA [70], suggesting a putative mechanism for this extremely complex transition. Finally, Andricioaei and coworkers explored the physical mechanism of the first stages of DNA unzipping [71].

During these last years the community has made a large effort in understanding the interplay between **ion atmosphere and DNA**. Ions are known to have a dramatic effect in modulating DNA properties. Increases in the ion concentration can change the equilibrium geometry of

DNA, and even a small quantity of certain ions can stabilize unusual conformations. Lavery, Maddocks, and Zakrzewska have been particularly active in this field, describing the sequence-dependent distribution of monovalent ions in equilibrium B-DNA, and developing analysis tools that help to describe diffuse ionic environment surrounding DNA [72•,73]. Pan *et al.* studied how ion distributions change in different forms of DNA and compared them with those obtained for RNA [74], while Šponer's group re-explored at the classical and quantum levels the unique ion atmosphere around G-quadruplexes [8]. As noted above, MacKerell's group has been also very active exploring the dependence of the DNA structure on the nature of the neutralizing cation [57,75,76], and has provided convincing evidence of ion-induced changes in the groove geometries, which are very visible when small counterions (Li^+) are used. Other authors took a step forward analyzing organic cations which might be quite abundant in certain cellular conditions. Thus, Sen *et al.* explored the effect of linear and cyclic diamines on DNA [77], while Sugimoto and coworkers [78], and Portella *et al.* [79], studied the effect of choline salts in DNA structure. The latter authors combined long MD simulations with NMR spectroscopy to characterize the nature of the choline$^+$-DNA interaction along the grooves, finding an explanation for the violation of the Watson–Crick rule occurring in the presence of choline [79].

The flexibility of MD simulations allowed exploring the behavior of **DNA in non-physiological environments**. For example, Arcella *et al.* combined experimental measures with classical and quantum dynamics to characterize the behavior of a small piece of DNA when moved to the vacuum [15], exploring the conformational and topological changes related to the transfer of DNA from water to apolar environments [80]. Portella *et al.* also combined experimental and simulation techniques to characterize the surprising stabilizing properties of pyridine at acidic pH [81], describing for the first time the strong anti-cooperativity of two powerful denaturants (pyridine and the acidic media). The number of non-physiological media in which DNA has been investigated is endless and we will only mention works on lipids [82], different dendrimers [83,84], silica surface [85], graphene [86] and carbon nanotubes [87]. Finally, it is worth noting the pioneering work of Case's group in considering a very unique environment: the crystal lattice, which somehow can be considered a surrogate of the crowded cellular environment [88]. It is still too early to be sure on the suitability of force-fields in some of these environments, which are very different to those considered in the calibration of the force-field, but it is clear that a new scenario for MD simulations of DNA has emerged in the last years.

MD simulations have been extensively used to understand the structure and properties of **modified DNAs**, i.e. duplexes containing mutations, damaged by oxidative stress, containing mismatches, epigenetic modifications, or different types of covalent changes. The number of studies in this section is also endless, and we will cite those that might have had a larger impact in the understanding of the biology of modified DNAs. From this point of view, it is worth mentioning the works on the impact of the oxidization or deamination of guanine [89–92] on the properties and repairing mechanisms of DNA, as well as the studies on the DNA mismatches [93,94], or the UV-damaged DNAs [95,96].

MD has been extensively used to explore **epigenetic variants of DNA**, especially of the most prevalent of such variants: the C5-methylated cytosine (MeC). Bianchi and Zangi explored the impact of cytosine methylation on base flipping [63], and the mechanisms of recognition by proteins of DNA containing MeC [97]. Broyde and coworkers used QM/MM to explore basic structural properties of DNA containing MeC [98], and Carvalho and coworkers [99] confirmed previous claims by other authors [100,101] that methylation largely alters the elastic properties of DNA. These studies suggested that methylation-related changes might affect the nucleosome binding, a hypothesis that has been confirmed in recent studies [102,103]. Finally, it is worth citing in this section the work by Esposito *et al.* [104], who combined a variety of theoretical techniques with experimental measures to characterize the impact of DNA methylation on the photoreactivity of DNA, opening new interpretations on the evolutionary origin of DNA methylation.

MD has been traditionally used to study **unusual forms of DNA**, and among the different studies published during these last three years on this topic we should highlight here the work by Cleri and coworkers [105] on i-DNA wires, and a series of works originated from different laboratories [8,106–108] which explored different aspects of G-quadruplex. A very extended set of simulations on different unusual DNA structures (including triplexes, quadruplexes, parallel DNAs, hybrids, *etc.*) was also published in the article presenting the new parmbsc1 force-field [44••], and available to the community from the BigNASim database [109].

Finally, we should mention the avalanche of works on **DNA complexes**. A full review would be necessary to summarize the work done on this topic since multiple systems have been studied and analyzed by MD simulations, going from small drug-DNA to huge protein–DNA complexes. We limit ourselves to comment works of general interest for the understanding of DNA interactions, or studies that provided information of strong biological significance. Among the first group of works, we should note the study of protein–DNA dissociation pathways by Yonetani and Kono [110], the discussion of specific and non-specific protein binding to DNA by

Domene and coworkers [111], the study of phage maturation by Brooks and coworkers that highlighted the importance of pH-induced changes in DNA packing [112•], the work of Galindo-Murillo *et al.* on DNA intercalation [113], and finally a comprehensive work by de Ruiter and Zagrovic on the interaction of protein side chains with DNA [114]. Among the second family of studies, we should highlight the works centered on chromatin; especially those that are trying to connect epigenetic changes with chromatin structure and gene expression regulation. Worth noting the work of Erler *et al.* [115], on the role of histone tails on nucleosome structure, the analysis done by the Papoian's group on the impact of acetylation on the structure of the tails [116], and a recent study by Collepardo-Guevara *et al.*, who combined atomistic MD, coarse grained simulations, and NMR spectroscopy to provide a mechanistic explanation of the role of histone acetylation in unpacking the nucleosome fiber [117••]. To finish this section, it is necessary to cite the work of Broyde's group [118] on the role of nucleosome architecture in altering interaction of DNA with repairing enzymes.

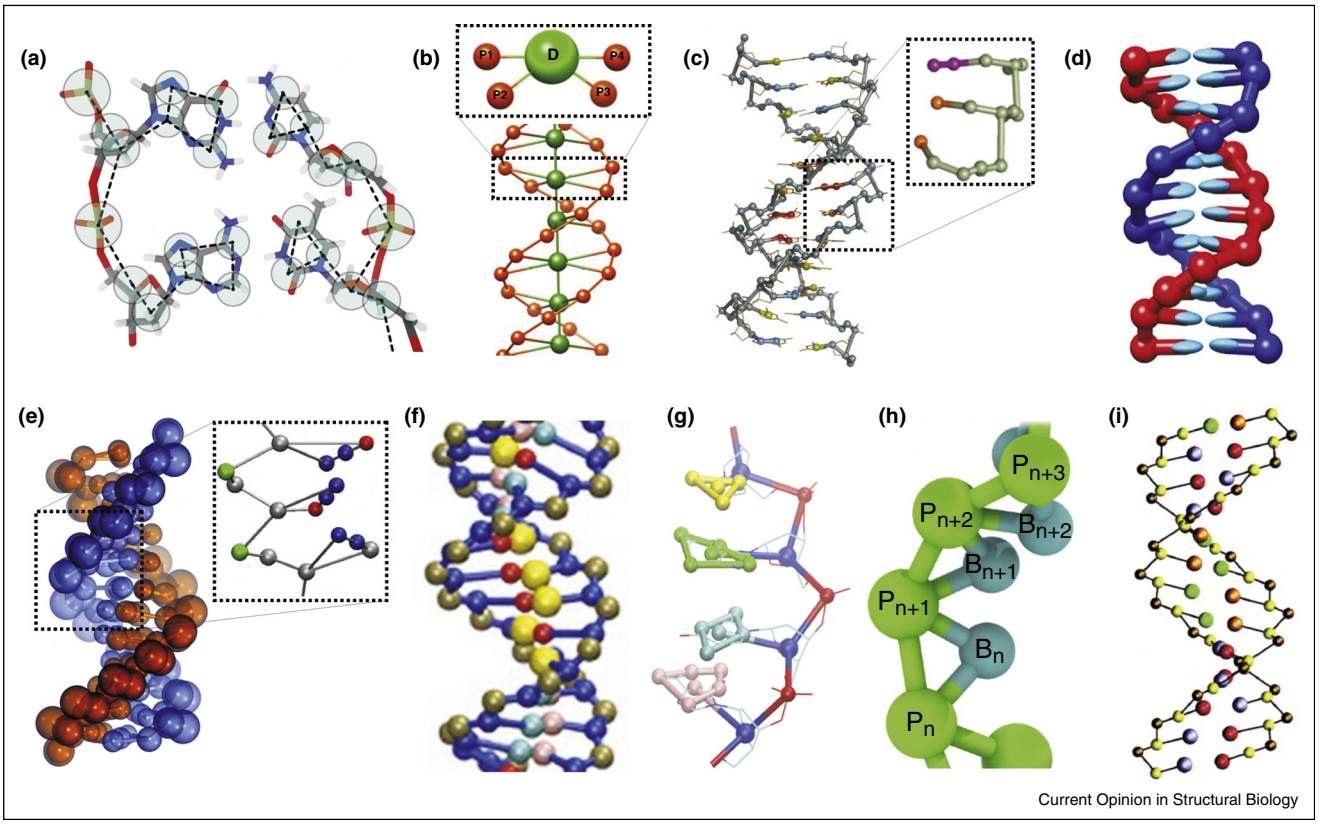
Coarse-grain studies

Coarse-graining (CG) is a common approach to handle large DNA systems that cannot be dealt with by means of atomistic models. Recent developments have expanded the accuracy of CG methods, especially for B-DNA, as extensively reviewed in a series of excellent articles by the groups of Levitt [119], Papoian [120], Marrink [121] and Noid [122] among others, and a book chapter by Leonarski and Trylska [123]. We will limit ourselves here to the latest (from 2013) advances in **particle-based CG methods**, both those oriented towards the study of DNA in biological context [124•,125••,126–133], and those designed for nanocomposites (see the recent revisions of Yingling *et al.* [134] and Ouldridge [135]). Beyond this division, the oxDNA model by Ouldridge and coworkers [136–138], the 3SPN model by de Pablo's group [131,133,134,139,140], and an extension [141•] of the DNA SIRAH model by Pantano's group [142] have been used in both fields. Despite the particle-based CG models reviewed here, it is worth to note the effort done by the community to build models at the interface of atomistic and coarse grain modeling, like the ones based on the flexibility of DNA bases considered as independent interacting rigid bodies where the ground state and the stiffness matrixes are taken from MD simulations [143,144], or the works done by Rohs and coworkers that used atomistic MC (Monte Carlo) simulations to derive a method for high-throughput DNA shape predictions [145–147].

The first decision in the development of a CG method is **the number of effective beads** used to represent each nucleotide. Most of the successful CG models use from 2 (oxDNA [135–138], and Aksimentiev's model [148]), or 3 (3SPN [131–133,139,140], BioModi [149]), to eight beads per nucleotide (see Figure 2, and Table 1 for more details). However, coarser models with five beads per base-pair step (four nucleotides), or even a single bead per nucleotide have emerged [126,127]. When more than one bead is used, all the models have chosen to place the beads in order to reproduce the position and connectivity existing between the backbone, the sugar puckering and the base (Figure 2). Due to the reduction in the number of particles, and hence in the degrees of freedom, systems containing dozens to thousands of bp have been successfully simulated (Table 1, and Figure 1), including ssDNA, dsDNA and DNA mini-circles. The second decision to be made in CG methods is the selection of the **energy functional**, which can be performed according to a *top-down* [124•,125••,129–131,133,135–138,140,141•] or to a *bottom-up* approximation [126,137,128,148]. In the ***top-down*** approach, the set of interactions is empirically parameterized, in a trial-and-error manner, to match experimentally determined thermodynamic properties (i.e. melting temperatures) or structural and mechanical features of double-stranded and single-stranded DNA. Simple equations, usually the same found in atomistic force-fields (see Subsection Atomistic Studies), are adjusted on the basis of physicochemical intuition to reproduce emergent structural or thermodynamics properties. In the ***bottom-up*** approach, effective CG interactions are extracted in a systematic and consistent way from reference atomistic simulations. Under the Statistical mechanics framework, the many-body Potential of Mean Force (PMF) for any specific coarse-grained system is completely specified by the underlying atomistic model and the chosen coarse-grain mapping. In practice, pair PMFs are used, and the parameters are determined iteratively. Finally, most of the potentials derived with a pure *bottom-up* approach are fine-tuned *a posteriori*, to reproduce experiments or phenomenological properties of a particular system. Under both approximations, the vast majority of the models use classical terms found in all-atom force fields, namely harmonic potentials for bonded terms [124•,127–130,142], and one or more expression (mostly Lennard–Jones, Coulomb, but also *ad hoc* potentials [129]) to reproduce non-bonded interactions (see Table 1 for a global view).

In more detail, and beyond the popular oxDNA, 3SPN, and SIRAH *top-down* models, Pasquali and Derreumaux extended their RNA model named HiRE-RNA to dsDNA [129], modifying the equilibrium values of the bond and angle interactions and adjusting the hydrogen bond terms. The model is useful in DNA folding and self-assembly of small oligomers. Only few months ago, the group of Marrink published a new MARTINI-DNA force-field [124•], which uses six beads per pyrimidines and seven beads per purines, and an elastic network to preserve the secondary structure (limiting the applicability of the model out of canonical helices). On the other

Figure 2



Mapping strategies for coarse-grained DNA from different models, as illustrated by the corresponding authors. **(a)** Model developed by Marrink and coworkers. Pyrimidines are represented with six beads and purines with seven beads. Adapted with permission from [124•]. **(b)** The model of Nordenskiöld and coworkers used five beads per bps to represent dsDNA. Adapted with permission from [127]. **(c)** Model from Derreumaux and coworkers. Six and seven beads are used to represent pyrimidines and purines respectively. Adapted with permission from [129]. **(d)** oxDNA model from Ouldridge's group. Two beads, with four interaction sites are used to represent each base. Adapted with permission from [137]. **(e)** SIRAH model from Pantano and coworkers. 6 beads per base are used to reduce the complexity. Adapted with permission from [142]. **(f)** Model developed by Nguyen and coworkers. Three beads are used per base. Adapted with permission from [149]. **(g)** DNA model from Scheraga's group. Bases are represented by six to eight beads. Adapted with permission from [128]. **(h)** Model from Aksimentiev's group. In this single-strand coarse-grained representation of DNA each base is reproduced by two beads. Adapted with permission from [148]. **(i)** 3SPN model of de Pablo and coworkers. Each base is represented by three beads (adapted with permission from [184]).

hand, two pure *bottom-up* models were recently published. Vercauteren's group [126] derived the interactions potentials between CG sites using a combined Iterative Boltzmann Inversion (IBI) and Newton Inversion (NI) schemes to fit all-atom MD simulations. This promising model, with back-mapping capabilities, was able to correctly reproduce ring closure probabilities and mini-circle topologies. Nordenskiöld and coworkers [127] used an Inverse Monte Carlo (IMC) algorithm to derive the set of CG interactions from all-atom MD simulations. The model reproduced correctly the salt-dependent persistence length of DNA. Finally, other two interesting hybrid (*bottom-up*/*top-down*) models were recently published: Scheraga's group, who extended the NARES-2P [150] model to reproduce DNA folding starting from two short single-stranded oligomers [128]; and Aksimentiev and coworkers [148] who published a model for ssDNA based on IBI from MD simulations, refined by fine tuning the parameters to reproduce experimentally measured radii of gyration.

No matter the approach used to derive the force-field, nor the final application of the model, one of the key difficulties specific to the DNA coarse graining is the correct handling of **long-range electrostatics**, something that is crucial to correctly represent one of the most highly charged naturally occurring polyelectrolytes [151]. Only a few models incorporate electrostatic explicitly [127,128,130,141•] assigning partial charges to the DNA beads, but in all the cases the environment around DNA has been taken into account at some degree, by using implicit or explicit approaches. Most models used implicit Langevin dynamics (with increased viscosity of the medium) coupled to the Debye–Hückel approach (when electrostatics is explicitly taken into account) to treat the ionic strength (Table 1). While in most cases the solvent is treated as a

Table 1

Key features of the main CG DNA models developed or widely used in the last 3 years

Name/group of the model	Number of beads per base	Max. num. base/base-pairs simulated	Type of potential	DNA environment	Other components	Main application
oxDNA Ouldridge	2 beads with 4 interaction sites. (2 total).	6200 bp (dsDNA).	Top-down. Fitted to melting temp. + ssDNA and dsDNA structure.	Implicit solvent + 500 nM added salt (Langevin dynamics).	No.	ssDNA. dsDNA. Nanotechnology. Biology. Crowded systems.
SIRAH Pantano	2 beads backbone. 1 bead sugar. 3 beads base. (6 total).	104 bp (dsDNA).	Top-down. Fitted to melting temp. + dsDNA structure. Harmonic bonds and angles. Coulomb + LJ.	Explicit (WT4) and implicit solvent + Debye–Hückel. Explicit ions.	Water. Ions. Proteins.	ssDNA. dsDNA. Biology.
3SPN.0/ 1/2/2C de Pablo	1 bead backbone. 1 bead sugar. 1 bead base. (3 total).	144 bases (ssDNA). 1490 bp (dsDNA).	Top-down. Fitted to thermal denaturation exp. Harmonic bonds and angles. Coulomb + non-bonded.	Implicit solvent (Langevin) + Debye–Hückel.	No.	ssDNA. dsDNA. Biology. Nanotechnology. Confined DNA.
BioModi Nguyen	1 bead backbone. 1 bead sugar. 1 bead base. (3 total).	350 bases (ssDNA). 32 bp (dsDNA).	Top-down. Fitted to MD and exp. structures. Discrete MD potentials.	Implicit solvent + Debye–Hückel.	Proteins. Polymers.	ssDNA. dsDNA. Biology and Nanotechnology. Crowded systems.
MARTINI Marrink	1 bead backbone. 2 beads sugar. 3 beads Y. 4 beads R. (6/7 total).	40 bases (ssDNA). 100 bp (dsDNA).	Top-down/bottom-up. Fitted to densities of liquids, partition coefficients, and MD. Harmonic bonds and angles + elastic network.	Explicit solvent and ions.	Lipids. Water. Polarized water. Carbohydrates. Polymers. Ions. Proteins.	ssDNA. dsDNA. Biology.
UNRES like-DNA Scheraga	1 bead backbone. 1 bead sugar. 4 beads C. 5 beads G and T. 6 beads A. (6/7/8 total).	60 bases (ssDNA). 60 bp (dsDNA).	Bottom-up/Top-down. Non-linear PLS algorithm to fit all-atom PMF calculations + reproduce B-DNA structure. Harmonic bonds + angles. LJ-like + electrostatics.	Implicit solvent + Debye–Hückel.	Proteins. DNA (coarser level: Nares-2P).	ssDNA. dsDNA. Biology
HiRe-DNA Derreumaux	3 beads backbone. 2 beads sugar. 1 bead Y. 2 beads R. (6/7 total).	16 bp (dsDNA).	Top-down. Fitted to exp. structures (NDB). Harmonic bonds and angles. Modified LJ + Hbond terms.	Implicit solvent.	RNA (HiRe-RNA).	dsDNA. Biology.
Nordenskiöld	Central bead represents 4 sugars + 4 bases. 4 beads phosphates. (5 total).	200 bp (dsDNA).	Bottom-up. Fitted to MD by IMC method. Harmonic bonds and angles. Coulomb + LJ.	Explicit ions. Implicit solvent (Langevin dynamics).	Ions.	dsDNA. Biology.
Aksimentiev	1 bead backbone. 1 bead base. (2 total).	200 bases (ssDNA).	Bottom-up/Top-down. Fitted to MD by IBI method + radius of gyration.	Implicit solvent (Langevin dynamics). Implicit ions.	No.	ssDNA. Biology.
Vercauteren	1 bead. (1 total).	500 bp (dsDNA) linear and circular.	Bottom-up. Fitted to MD by IBI and NI methods.	Implicit solvent. Explicit ions.	Ions.	dsDNA. Biology. Mini-circles.

Table 1 (*Continued*)

Name/group of the model	Number of beads per base	Max. num. base/base-pairs simulated	Type of potential	DNA environment	Other components	Main application
Stachiewicz and Molski	2 beads backbone. 1 bead sugar. 3 beads base. (6 total). SIRAH scheme.		Top-down. SIRAH potential with modified parameters.	Explicit solvent and ions from MARTINI ff.	Water. Ions.	dsDNA. Nanotechnology.

continuum, some explicit models for water and ions have been developed to work with specific coarse-grained models. For example, the WT4 model, which condenses 11 water molecules in four beads, was meant to work in conjunction with the SIRAH force-field [142]. In the same way, the regular and polarized water models by Marrink's group, where each bead represents four water molecules, have been developed to work with the MARTINI force-field [124•].

The latest developments in the field are addressed to mix protein and DNA CG models to allow the calculation of **protein–DNA complexes**, or even medium-size chromatin fibers. In this regard, the new model from the Marrink's group [124•] completes the MARTINI force-field, allowing the unified CG simulation of DNA, proteins, water, carbohydrates, ions and lipids. The model developed by Nguyen and coworkers was thought to combine with ePRIME [152], and be delivered as BioModi [149], a unified CG force-field for DNA, proteins and general polymers. The DNA model from Scheraga's group [128] was derived with the same philosophy used to derive the UNRES [153] force-field for proteins, and is expected to be compatible. Finally, the incorporation of proteins to the SIRAH force-field [154], allows the comprehensive simulation of DNA, proteins, water and ions in a multiscale (all-atom/coarse-grain) manner [125••].

In the nanotechnological field, the mentioned models have been used to understand the 3-dimensional arrangements of self-assemblies of large DNA nanostructures like nanotetrahedrons, three-armed star motif, and other macromolecular conformations belonging to the domains of DNA origami [136]. Advances have also been made in the field of DNA nanodevices like for example nanotweezers [136,138], where hybridization dynamics has been described in terms of strand-exchange, internal displacement, and zippering. Burnt-bridges DNA motors have also been simulated by means of coarse-grain models [137]. Denaturation and renaturation has been studied in crowded or confined spaces, also highlighting the existence of a zippering and slithering mechanism underlying DNA hybridization [140,141•,149]. Hybridization is also a crucial mechanism in DNA replication and translation. In this regard, several coarse-grained models were applied to understand certain processes of biological interest, like duplex formation starting from short ssDNA oligomers [127,131,132,142]. Other models have shown to be useful in reproducing the structure of DNA/RNA duplexes [129], or structural properties such as the sequence-dependent/salt-dependent persistence length [127,131,132], or the DNA curvature. The model from Vercauteren and coworkers was able to reproduce the topology of several DNA mini-circles with different link numbers, simulated using explicit K^+Cl^- ions [126]. Commendable examples have been reported from the use of unified CG force-fields, with systems including DNA, proteins, explicit water and ions [124•], and hybrid MM/CG approaches [125••].

Mesoscopic studies

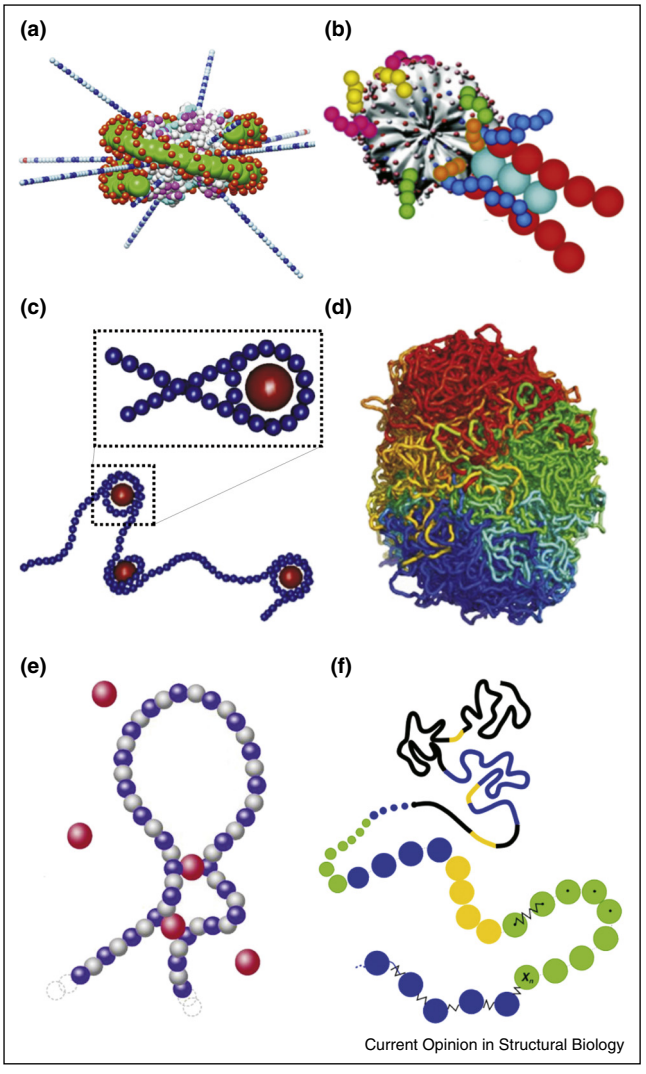
On a larger scale, DNA of around 1 m in length (in human cell) has to compress into a nucleus of ~6 μm in diameter. The first level of compaction (by a factor ~1.7) is achieved by wrapping 147 bp of duplex DNA around a histone octamer forming the nucleosome. Nucleosomes are connected by 20–80 bp long linkers forming a nucleosome string called chromatin. Early in vitro experiments [155] suggested the compaction of the nucleosome string into a 30 nm fiber, but the situation in vivo is probably more complex (Figure 1) and depends on many variables such as the linker length, ionic environment, the presence/absence of linker histones and the effect of chromatin remodelers [156,157]. General consensus is that chromatin probably adopts dynamically a mix of structures depending on the cell activity [158,159]. This fluidity helps the cell to modulate chromatin accessibility and accordingly the DNA expression level [158], but complicates the theoretical description of the in vivo chromatin structure. Despite recent advances, even coarse-grained methods are unable to manage the size and complexity of chromatin. In this section we will give an overview of the latest theoretical models specifically developed to study chromatin organization. In addition, we would like to point out recent reviews on models examining chromatin folding [160–164].

We can divide computational approaches of chromatin structure into three basic models: ***bottom-up***, ***top-down*** and ***intermediate***. The ***bottom-up*** approaches make use of the atomistic properties of nucleosome and linker DNA. Properties like electrostatics and accurate three-dimensional shape of the constituents are transferred into a coarse-grained model of chromatin.

Multiple nucleosomes with their linker DNA are connected to create kbp long chains which are used to simulate compaction, accessibility and other chromatin features under different internal parameters such as variations in DNA linker length [156,165], ionic environment [157,166], presence of linker histones [167] and different intra-chain and inter-chain physical interactions [156,157,165–169]. In the ***top-down*** models the chromatin structure is derived by implementing experimental restraints coming from chromosome conformation capture techniques into a simple model of the chromatin fiber. Techniques such as Hi-C [170] provide low resolution (at best 1 kb with in situ Hi-C [171]) information on prevalent contacts of the chromatin fiber which, transformed into distance or spatial contact restraints, can be used to visualize the target chromatin region. There are in general two modeling strategies to convert the experimental output into a three-dimensional object (recently reviewed in [164]). One category of models directly transforms the contacts analytically into a single 3D structure while another set of models uses optimization-based methods to generate multiple possible configurations [164]. Introduction of higher level of detail can be achieved by the ***intermediate*** chain-of-beads approach, which involves two features: small-scale chromatin properties and overall genome organization based on experimental results, for example from Hi-C [170], FISH [172] or cryo-EM [159] experiments. At this scale chromatin is usually modeled as a polymer chain (one monomer unit can comprise less than one kb [173•] up to several kb's [174,175]) with energy terms representing intra-chain and inter-chain interactions and incorporating specific constraints (for example nucleus size, general chromatin shape or Hi-C contacts) [175–177]. Constant improvement in Hi-C techniques [171] and the recent irruption of ultra-resolution fluorescence microscopy in the field [178] suggests that there is room for the 'intermediate' approach to improve the level of resolution, with the long-range objective to reach at least nucleosome-level resolution.

Different coarse-grained models of DNA and nucleosome have recently been developed in the scope of the ***bottom-up*** approach. For example, Nordenskiöld's group presented a novel model of the nucleosome (Figure 3a) with flexible histone tails and detailed representation of nucleosomal DNA to probe the influence of mono-valent, di-valent and tri-valent counterions in intra-nucleosomal and inter-nucleosomal interactions [157]. The model was extended in a multiscale study to a super-coarse-grained representation of the nucleosome which made it possible to study aggregation of an ensemble of up to 5000 nucleosomes [179•]. Schlick and coworkers developed a coarse-grained chromatin model (Figure 3b) which incorporates (in addition to an excluded volume term) a Debye–Hückel representation of the nucleosome core particle as well as flexible histone tails, linker histones and

Figure 3



Illustrations of different *bottom-up*, mixed, and *intermediate* mesoscale models. **(a)** In the advanced nucleosome model of Nordenskiöld each amino acid is represented as a bead while the nucleosomal DNA is modeled as illustrated in Figure 2b; adapted with permission from [157]. **(b)** Schlick's group represents the nucleosome core as a charged irregular surface, whit flexible histone tails (five beads per tail), shown in yellow, purple, green, blue and orange. The linker histone consists of three beads (shown in turquoise), the linker DNA is a chain of beads comprising 10 bp per beads (shown in red); adapted with permission from [156]; **(c)** Chromatin model of Ohyama's group. DNA is illustrated as a worm-like chain (shown in blue; six bp per bead) and wrapped around rigid spherical nucleosome core particles (shown in red) with a specific entry-exit angle; adapted with permission from [183•]. **(d)** The model of Dekker's group reduces the details of chromatin to a polymer chain of 128,000 spheres with 10 nm in diameter (600 bp per bead) to reproduce Hi-C results; adapted with permission from [173•]. **(e)** The 'strings and binders switch' model of Nicodemi and coworkers uses a self-avoiding worm-like chain of 512 spherical beads (20 kb per bead) while beads can act as a binding site (blue) for diffusive binders (red); adapted with permission from [174]. **(f)** Block copolymer model of Jost *et al*. Chromatin is modeled as a self avoiding bead-spring polymer of up to 131 monomers (10 kbp per monomer) while specific attractive short-range interactions between a monomers of the same color account for different epigenomic states in the model (adapted with permission from [175]).

linker DNA represented by means of a worm-like chain model [180]. The Schlick model allows to study quite long (~50 nucleosomes) fibers [181] and has been used to study, for example, the influence of linker histone H1 [182], the DNA linker length [156,181], or the effect of certain epigenetic modifications on chromatin arrangement [117••]. Finally, worth to mention is the work by Müller *et al.* [165] who simulated a fiber where nucleosomes are represented as cylinders connected by a chain of linker DNA spheres.

Trying to move from the ***bottom-up*** to the ***intermediate*** level Ohyama's group developed a model using a 'beads-on-a-string'-like representation of the chromatin fiber (Figure 3c) where linker DNA corresponds to a chain of beads (each bead has 2 nm in diameter; 6 bp/bead) and nucleosomes are modeled as spheres with nucleosomal DNA wrapped around it [183•]. Persistence lengths of individual linker DNA pieces were used to account for elastic bending of the linker DNA. Chromatin chains were grown with Monte Carlo methods and filtered according to the size of the nucleus. Despite the lack of electrostatic and accurate steric interactions in the model the group was able to build a reasonable model of the yeast chromatin [183•].

Within the pure ***intermediate*** methods the objective is to move to much longer models than those accessible to the ***bottom-up*** approach, relying on the experimental data to correct the intrinsic limitations of the simple physical model used [159,170,172]. Quite surprisingly, very simple polymer models such as C-SAC [177] are able to reproduce well some experimental details on chromatin compaction such as the observed scaling behavior of contact probability vs. genomic distance in interphase human chromatin just by introducing restrictions on the nuclear volume. Similarly, Gehlen *et al.* [176] were able to reproduce overall structural architecture of yeast chromatin by putting spatial constraints involving the positions of centromers, telomeres and nucleolus.

Trying to gain higher resolution and to introduce more details on chromatin composition Dekker's group developed a polymer model (Figure 3d) to match the observed Hi-C data of human mitotic chromosomes [173•]. Chromatin was represented as a polymer chain of 128,000 beads (600 bp per bead). Attractive and repulsive Lennard–Jones potentials capture the interactions of the polymer. The potential was softened at short distances so that monomers can pass through each other to mimic topoisomerase II action. Additional restraints to promote looping and spatial elastic constraints to attract the polymer towards the central axis of the confined volume were set to match best the experimental Hi-C data. Nicodemi's group developed an alternative model based on a self-avoiding polymer chain of 20 kb beads [174] which includes the possibility to incorporate the effect of diffusive binders. Those binder particles can float in solution or bind to the chain modulating local properties of the fiber (Figure 3e). By tuning the parameters the chromatin chain can respond to different environmental conditions and can guide the chain into certain structural configurations, which was successfully applied to the representation of human mitotic chromatin [174]. Jost *et al.* [175] introduced a block co-polymer model (each monomer represents 10 kbp of DNA) which incorporates epigenomic features into a continuous polymer model (Figure 3f). Specific attractive interactions between monomers within a certain segment simulate the nature of each epigenomic domain. The model incorporates Hi-C restraints and has been successfully used in the study of Drosophila chromatin structure [175].

As computer power increases and as experimental data increase in resolution, we can expect the irruption of multiscale approaches [117••,183•] to guide the future development of *in-silico* chromatin models.

Conclusions

Increase in computer power and improvement in algorithms allow us to dream of the possibility of gaining a holistic view of DNA from theoretical calculations. We can envision a near future where, by moving in a continuum of methodologies, we will be able to explore from fine details of the electronic distributions at a given DNA step, to large chromatin rearrangements occurring throughout the cell cycle. Fast hardware, powerful algorithms, and clever integration of experimental data will be needed to make those expectations real.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgements

MO thanks Spanish Ministry of Science (BIO2012-32868 and BFU2014-61670-EXP), the Catalan SGR, the Instituto Nacional de Bioinformática, and the European Research Council (ERC SimDNA), the H2020 program (MuG and BioExcel projects) for support. MO is an ICREA academia researcher. PDD is a PEDECIBA and SNI (ANII, Uruguay) researcher. HG is a JdC researcher.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Soler JM, Artacho E, Gale JD, García A, Junquera J, Ordejón P, Sánchez-Portal D: **The SIESTA method for ab initio order-N materials simulation**. *J Phys Condens Matter* 2002, **14**:2745-2779.

2. Genovese L, Neelov A, Goedecker S, Deutsch T, Ghasemi SA, Willand A, Caliste D, Zilberberg O, Rayson M, Bergman A *et al.*: **Daubechies wavelets as a basis set for density functional pseudopotential calculations**. *J Chem Phys* 2008, **129**:14109.

3. Mohr S, Ratcliff LE, Boulanger P, Genovese L, Caliste D, Deutsch T, Goedecker S: **Daubechies wavelets for linear scaling density functional theory**. *J Chem Phys* 2014, **140**:204110.

4. Hobza P, Šponer J: **Structure, energetics, and dynamics of the nucleic acid base pairs: nonempirical ab initio calculations**. *Chem Rev* 1999, **99**:3247-3276.

5. Šponer J, Leszczynski J, Hobza P: **Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases**. *Biopolymers* 2001, **61**:3-31.

6. Mládek A, Banáš P, Jurečka P, Otyepka M, Zgarbová M, Šponer J: **Energies and 2′-hydroxyl group orientations of RNA backbone conformations. Benchmark CCSD(T)/CBS database, electronic analysis, and assessment of DFT methods and MD simulations**. *J Chem Theory Comput* 2014, **10**:463-480.

7. Mládek A, Krepl M, Svozil D, Čech P, Otyepka M, Banáš P, Zgarbová M, Jurečkad P, Šponer J: **Benchmark quantum-chemical calculations on a complete set of rotameric families of the DNA sugar-phosphate backbone and their comparison with modern density functional theory**. *Phys Chem Chem Phys* 2013, **15**:7295-7310.

8. Gkionis K, Kruse H, Platts JA, Mládek A, Koča J, Šponer J: **Ion binding to quadruplex DNA stems, comparison of MM and QM descriptions reveals sizable polarization effects not included in contemporary simulations**. *J Chem Theory Comput* 2014, **10**:1326-1340.

9. • Šponer J, Mládek A, Špačková N, Cang X, Cheatham TE, Grimme S: **Relative stability of different DNA guanine quadruplex stem topologies derived using large-scale quantum-chemical computations**. *J Am Chem Soc* 2013, **135**:9785-9796.

The authors performed accurate QM (DFT-D3) calculations to study the relative stability of different folds of the cation-stabilized guanine quadruplexes (G-DNA), obtaining good agreement with the experimental findings. Moreover, they provided evidence on the limitations of classical force fields in the study of nucleic acids.

10. Kruse H, Havrila M, Šponer J: **QM computations on complete nucleic acids building blocks: analysis of the Sarcin–Ricin RNA motif using DFT-D3, HF-3c, PM6-D3H, and MM approaches**. *J Chem Theory Comput* 2014, **10**:2615-2629.

11. Lech C, Heddi B, Phan A: **Guanine base stacking in G-quadruplex nucleic acids**. *Nucleic Acids Res* 2013, **41**:2034-2046.

12. •• Parker TM, Hohenstein EG, Parrish RM, Hud NV, Sherrill DC: **Quantum-mechanical analysis of the energetic contributions to π stacking in nucleic acids versus rise, twist, and slide**. *J Am Chem Soc* 2013, **135**:1306-1316.

The Symmetry-Adapted Perturbation Theory (SAPT) was used in model QM calculations to estimate the energetic components of base stacking and their relation with the helical parameters. According to their results, the London dispersion forces are the most important attractive component in base stacking, and the backbone needs to be considered to obtain good stacking geometries, principally in RNA.

13. Sochacka E, Szczepanowski RH, Cypryk M, Sobczak M, Janicka M, Kraszewska K, Bartos P, Chwialkowska A, Nawrot B: **2-Thiouracil deprived of thiocarbonyl function preferentially base pairs with guanine rather than adenine in RNA and DNA duplexes**. *Nucleic Acids Res* 2015, **43**:2499-2512.

14. Brovarets' OO, Hovorun DM: **Why the tautomerization of the G.C. Watson–Crick base pair via the DPT does not cause point mutations during DNA replication? QM and QTAIM comprehensive analysis**. *J Biomol Struct Dyn* 2014, **32**:1474-1499.

15. Arcella A, Dreyer J, Ippoliti E, Ivani I, Portella G, Gabelica V, Carloni P, Orozco M: **Structure and dynamics of oligonucleotides in the gas phase**. *Angew Chem Int Ed Engl* 2015, **54**:467-471.

16. Molina R, Stella S, Redondo P, Gómez H, Marcaida MJJ, Orozco M, Prieto J, Montoya G: **Visualizing phosphodiester-bond hydrolysis by an endonuclease**. *Nat Struct Mol Biol* 2015, **22**:65-72.

17. Nam K, Cui Q, Gao J, York DM: **Specific reaction parametrization of the AM1/d hamiltonian for phosphoryl transfer reactions: H, O, and P atoms**. *J Chem Theory Comput* 2007, **3**:486-504.

18. Sgrignani J, Magistrato A: **QM/MM MD simulations on the enzymatic pathway of the human flap endonuclease (hFEN1) elucidate common cleavage pathways to RNase H enzymes**. *ACS Catal* 2015, **5**:3864-3875.

19. Aranda J, Zinovjev K, Roca M, Tuñón I: **Dynamics and reactivity in thermus aquaticus N6-adenine methyltransferase**. *J Am Chem Soc* 2014, **136**:16227-16239.

20. Mlýnský V, Banáš P, Šponer J, van der Kamp MW, Mulholland AJ, Otyepka M: **Comparison of ab initio, DFT, and semiempirical QM/MM approaches for description of catalytic mechanism of hairpin ribozyme**. *J Chem Theory Comput* 2014, **10**:1608-1622.

21. Garrec J, Patel C, Rothlisberger U, Dumont E: **Insights into intrastrand cross-link lesions of DNA from QM/MM molecular dynamics simulations**. *J Am Chem Soc* 2012, **134**:2111-2119.

22. Chandan P, Garrec J, Dupont C, Dumont E: **What singles out the G[8-5]C intrastrand DNA cross-link? Mechanistic and structural insights from quantum mechanics/molecular mechanics simulations**. *Biochemistry* 2013, **52**:425-431.

23. Lior-Hoffmann L, Wang L, Wang S, Geacintov NE, Broyde S, Zhang Y: **Preferred WMSA catalytic mechanism of the nucleotidyl transfer reaction in human DNA polymerase elucidates error-free bypass of a bulky DNA lesion**. *Nucleic Acids Res* 2012, **40**:9193-9205.

24. Lech CJ, Phan AT, Michel-Beyerle M-EE, Voityuk AA: **Electron-hole transfer in G-quadruplexes with different tetrad stacking geometries: a combined QM and MD study**. *J Phys Chem B* 2013, **117**:9851-9856.

25. Livshits GI, Stern A, Rotem D, Borovok N, Eidelshtein G, Migliore A, Penzo E, Wind SJ, Di Felice R, Skourtis SS, Cuevas JC, Gurevich L, Kotlyar AB, Porath D: **Long-range charge transport in single G-quadruplex DNA molecules**. *Nat Nanotechnol* 2014, **9**:1040-1046.

26. Bacolla A, Zhu X, Chen H, Howells K: **Local DNA dynamics shape mutational patterns of mononucleotide repeats in human genomes**. *Nucleic Acids Res* 2015, **43**:5065-5080.

27. Kubar T, Elstner M: **A hybrid approach to simulation of electron transfer in complex molecular systems**. *J R Soc Interface* 2013, **10**:20130415.

28. Spata VA, Matsika S: **Role of excitonic coupling and charge-transfer states in the absorption and CD spectra of adenine-based oligonucleotides investigated through QM/MM simulations**. *J Phys Chem A* 2014, **118**:12021-12030.

29. Gattuso H, Assfeld X, Monari A: **Modeling DNA electronic circular dichroism by QM/MM methods and Frenkel Hamiltonian**. *Theor Chem Acc* 2015, **134**:36.

30. Zelený T, Ruckenbauer M, Aquino AJA, Muller T, Lankaš F, Dršata T, Hase WL, Nachtigallova D, Lischka H: **Strikingly different effects of hydrogen bonding on the photodynamics of individual nucleobases in DNA: comparison of guanine and cytosine**. *J Am Chem Soc* 2012, **134**:13662-13669.

31. Pérez A, Luque FJ, Orozco M: **Frontiers in molecular dynamics simulations of DNA**. *Acc Chem Res* 2012, **45**:196-205.

32. Cheatham TE, Case DA: **Twenty-five years of nucleic acid simulations**. *Biopolymers* 2013, **99**:969-977.

33. • Galindo-Murillo R, Roe DR, Cheatham TE: **On the absence of intrahelical DNA dynamics on the μs to ms timescale**. *Nat Commun* 2014, **5**:5152.

A comprehensive MD study using MD-specific hardware that provide useful insights into the dynamics of DNA in the multi-microsecond time scale.

34. Galindo-Murillo R, Roe DR, Cheatham TE: **Convergence and reproducibility in molecular dynamics simulations of the DNA duplex d(GCACGAACGAACGAACGC)**. *Biochim Biophys Acta* 2015, **1850**:1041-1058.

35. Cheatham TE, Cieplak P, Kollman PA: **A modified version of the Cornell *et al.* force field with improved sugar pucker phases and helical repeat**. *J Biomol Struct Dyn* 1999, **16**:845-862.

36. Várnai P, Zakrzewska K: **DNA and its counterions: a molecular dynamics study**. *Nucleic Acids Res* 2004, **32**:4269-4280.

37. Pérez A, Marchán I, Svozil D, Sponer J, Cheatham TE, Laughton CA, Orozco M: **Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers**. *Biophys J* 2007, **92**:3817-3829.

38. Pérez A, Luque FJ, Orozco M: **Dynamics of B-DNA on the microsecond time scale**. *J Am Chem Soc* 2007, **129**:14739-14745.

39. Lavery R, Zakrzewska K, Beveridge D, Bishop TC, Case DA, Cheatham T, Dixit S, Jayaram B, Lankas F, Laughton C *et al.*: **A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA**. *Nucleic Acids Res* 2010, **38**:299-313.

40. • Pasi M, Maddocks JH, Beveridge D, Bishop TC, Case DA, Cheatham T, Dans PD, Jayaram B, Lankas F, Laughton C *et al.*: **μABC: a systematic microsecond molecular dynamics study of tetranucleotide sequence effects in B-DNA**. *Nucleic Acids Res* 2014, **42**:12272-12283.

A land-mark study by the ABC consortium, where the microsecond dynamics of all unique tretamers of DNA is explored.

41. Dans PD, Pérez A, Faustino I, Lavery R, Orozco M: **Exploring polymorphisms in B-DNA helical conformations**. *Nucleic Acids Res* 2012, **40**:10668-10678.

42. Zgarbová M, Luque FJ, Sponer J, Cheatham TE, Otyepka M, Jurečka P: **Toward Improved description of DNA backbone: revisiting epsilon and zeta torsion force field parameters**. *J Chem Theory Comput* 2013, **9**:2339-2354.

43. Krepl M, Zgarbová M, Stadlbauer P, Otyepka M, Banáš P, Koča J, Cheatham TE, Jurečka P, Sponer J: **Reference simulations of noncanonical nucleic acids with different χ variants of the AMBER force field: quadruplex DNA, quadruplex RNA and Z-DNA**. *J Chem Theory Comput* 2012, **8**:2506-2520.

44. •• Ivani I, Dans PD, Noy A, Pérez A, Faustino I, Hospital A, Walther J, Andrio P, Goñi R, Balaceanu A *et al.*: **Parmbsc1: a refined force-field for DNA simulations**. *Nat Methods* 2015 http://dx.doi.org/10.1038/nmeth.3658.

The presentation of the new parmbsc1 for-field, with more than 140 ms of dynamics on around 100 different DNA systems.

45. Huang J, MacKerell AD: **CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data**. *J Comput Chem* 2013, **34**:2135-2145.

46. Chen AA, García AE: **High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations**. *Proc Natl Acad Sci U S A* 2013, **110**:16820-16825.

47. Brown RF, Andrews CT, Elcock AH: **Stacking free energies of all DNA and RNA nucleoside pairs and dinucleoside-monophosphates computed using recently revised AMBER parameters and compared with experiment**. *J Chem Theory Comput* 2015, **11**:2315-2328.

48. Colominas C, Luque FJ, Orozco M: **Dimerization of formamide in gas phase and solution. An ab initio MC-MST study**. *J Phys Chem A* 1999, **103**:6200-6208.

49. Bommarito S, Peyret N, SantaLucia J: **Thermodynamic parameters for DNA sequences with dangling ends**. *Nucleic Acids Res* 2000, **28**:1929-1934.

50. Guckian KM, Schweitzer BA, Ren RX-F, Sheils CJ, Tahmassebi DC, Kool ET: **Factors contributing to aromatic stacking in water: evaluation in the context of DNA**. *J Am Chem Soc* 2000, **122**:2213-2222.

51. Banáš P, Mládek A, Otyepka M, Zgarbová M, Jurečka P, Svozil D, Lankaš F, Šponer J: **Can we accurately describe the structure of adenine tracts in B-DNA? Reference quantum-chemical computations reveal overstabilization of stacking by molecular mechanics**. *J Chem Theory Comput* 2012, **8**:2448-2460.

52. Luque FJ, Dehez F, Chipot C, Orozco M: **Polarization effects in molecular interactions**. *Wiley Interdiscip Rev Comput Mol Sci* 2011, **1**:844-854.

53. Huang J, Lopes PEM, Roux B, MacKerell AD: **Recent advances in polarizable force fields for macromolecules: microsecond simulations of proteins using the classical drude oscillator model**. *J Phys Chem Lett* 2014, **5**:3144-3150.

54. • Savelyev A, MacKerell AD: **All-atom polarizable force field for DNA based on the classical Drude oscillator model**. *J Comput Chem* 2014, **35**:1219-1239.

Presentation of CHARMM polarizable force-field adapted to perform MD simulations of DNA.

55. Savelyev A, MacKerell AD: **Balancing the interactions of ions, water, and DNA in the Drude polarizable force field**. *J Phys Chem B* 2014, **118**:6742-6757.

56. Lemkul JA, Savelyev A, MacKerell AD: **Induced polarization influences the fundamental forces in DNA base flipping**. *J Phys Chem Lett* 2014, **5**:2077-2083.

57. Savelyev A, MacKerell AD: **Competition among Li(+), Na(+), K(+), and Rb(+) monovalent ions for DNA in molecular dynamics simulations using the additive CHARMM36 and Drude polarizable force fields**. *J Phys Chem B* 2015, **119**:4428-4440.

58. Dans PD, Faustino I, Battistini F, Zakrzewska K, Lavery R, Orozco M: **Unraveling the sequence-dependent polymorphic behavior of d(CpG) steps in B-DNA**. *Nucleic Acids Res* 2014, **42**:11304-11320.

59. Zgarbová M, Otyepka M, Šponer J, Lankaš F, Jurečka P: **Base pair fraying in molecular dynamics simulations of DNA and RNA**. *J Chem Theory Comput* 2014, **10**:3177-3189.

60. Banavali NK: **Partial base flipping is sufficient for strand slippage near DNA duplex termini**. *J Am Chem Soc* 2013, **135**:8274-8282.

61. Banavali NK: **Analyzing the relationship between single base flipping and strand slippage near DNA duplex termini**. *J Phys Chem B* 2013, **117**:14320-14328.

62. Xu X, Ge H, Gu C, Gao YQ, Wang SS, Thio BJR, Hynes JT, Xie XS, Cao J: **Modeling spatial correlation of DNA deformation: DNA allostery in protein binding**. *J Phys Chem B* 2013, **117**:13378-13387.

63. Bianchi C, Zangi R: **Base-flipping propensities of unmethylated, hemimethylated, and fully methylated CpG sites**. *J Phys Chem B* 2013, **117**:2348-2358.

64. Kim S, Brostromer E, Xing D, Jin J, Chong S, Ge H, Wang S, Gu C, Yang L, Gao YQ *et al.*: **Probing allostery through DNA**. *Science* 2013, **339**:816-819.

65. Dršata T, Zgarbová M, Špačková N, Jurečka P, Šponer J, Lankaš F: **Mechanical model of DNA allostery**. *J Phys Chem Lett* 2014, **5**:3831-3835.

66. Taranova M, Hirsh AD, Perkins NC, Andricioaei I: **Role of microscopic flexibility in tightly curved DNA**. *J Phys Chem B* 2014, **118**:11028-11036.

67. Mitchell JS, Harris SA: **Thermodynamics of writhe in DNA minicircles from molecular dynamics simulations**. *Phys Rev Lett* 2013, **110**:148105.

68. Yang C, Kim E, Pak Y: **Free energy landscape and transition pathways from Watson–Crick to Hoogsteen base pairing in free duplex DNA**. *Nucleic Acids Res* 2015, **43**:7769-7778.

69. Nikolova EN, Goh GB, Brooks CL, Al-Hashimi HM: **Characterizing the protonation state of cytosine in transient G·C Hoogsteen base pairs in duplex DNA**. *J Am Chem Soc* 2013, **135**:6766-6769.

70. Moradi M, Babin V, Roland C, Sagui C: **Reaction path ensemble of the B-Z-DNA transition: a comprehensive atomistic study**. *Nucleic Acids Res* 2013, **41**:33-43.

71. Mentes A, Florescu AM, Brunk E, Wereszczynski J, Joyeux M, Andricioaei I: **Free-energy landscape and characteristic forces for the initiation of DNA unzipping**. *Biophys J* 2015, **108**:1727-1738.

72. • Lavery R, Maddocks JH, Pasi M, Zakrzewska K: **Analyzing ion distributions around DNA**. *Nucleic Acids Res* 2014, **42**:8138-8149.

A comprehensive study of the ion atmosphere around DNA from extended MD simulations.

73. Pasi M, Maddocks JH, Lavery R: **Analyzing ion distributions around DNA: sequence-dependence of potassium ion distributions from microsecond molecular dynamics**. *Nucleic Acids Res* 2015, **43**:2412-2423.

74. Pan F, Roland C, Sagui C: **Ion distributions around left- and right-handed DNA and RNA duplexes: a comparative study**. *Nucleic Acids Res* 2014, **42**:13981-13996.

75. Savelyev A, MacKerell AD: **Differential impact of the monovalent ions Li(+), Na(+), K(+), and Rb(+) on DNA conformational properties**. *J Phys Chem Lett* 2015, **6**:212-216.

76. Savelyev A, MacKerell AD: **Differential deformability of the DNA minor groove and altered BI/BII backbone conformational equilibrium by the monovalent ions Li^+, Na^+, K^+, and Rb^+ via water-mediated hydrogen bonding**. *J Chem Theory Comput* 2015, **11** 150826153042001.

77. Sen A, Sahu D, Ganguly B: **In silico studies toward understanding the interactions of DNA base pairs with protonated linear/cyclic diamines**. *J Phys Chem B* 2013, **117**:9840-9850.

78. Nakano M, Tateishi-Karimata H, Tanaka S, Sugimoto N: **Choline ion interactions with DNA atoms explain unique stabilization of A–T base pairs in DNA duplexes: a microscopic view**. *J Phys Chem B* 2014, **118**:379-389.

79. Portella G, Germann MW, Hud NV, Orozco M: **MD and NMR analyses of choline and TMA binding to duplex DNA: on the origins of aberrant sequence-dependent stability by alkyl cations in aqueous and water-free solvents**. *J Am Chem Soc* 2014, **136**:3075-3086.

80. Arcella A, Portella G, Collepardo-Guevara R, Chakraborty D, Wales DJ, Orozco M: **Structure and properties of DNA in apolar solvents**. *J Phys Chem B* 2014, **118**:8540-8548.

81. Portella G, Terrazas M, Villegas N, González C, Orozco M: **Can a denaturant stabilize DNA? Pyridine reverses DNA denaturation in acidic pH**. *Angew Chem Int Ed Engl* 2015, **54**:10488-10491.

82. Bagai S, Sun C, Tang T: **Lipid-modified polyethylenimine-mediated DNA attraction evaluated by molecular dynamics simulations**. *J Phys Chem B* 2014, **118**:7070-7076.

83. Mandal T, Kumar MVS, Maiti PK: **DNA assisted self-assembly of PAMAM dendrimers**. *J Phys Chem B* 2014, **118**:11805-11815.

84. Ficici E, Andricioaei I: **On the possibility of facilitated diffusion of dendrimers along DNA**. *J Phys Chem B* 2015, **119**:6894-6904.

85. Shi B, Shin YK, Hassanali AA, Singer SJ: **DNA binding to the silica surface**. *J Phys Chem B* 2015, **119**:11030-11040.

86. Bobadilla AD, Seminario JM: **Assembly of a noncovalent DNA junction on graphene sheets and electron transport characteristics**. *J Phys Chem C* 2013, **117**:26441-26453.

87. Bascom G, Andricioaei I: **Single-walled carbon nanotubes modulate the B- to A-DNA transition**. *J Phys Chem C Nanomater Interfaces* 2014, **118**:29441-29447.

88. Liu C, Janowski PA, Case DA: **All-atom crystal simulations of DNA and RNA duplexes**. *Biochim Biophys Acta* 2015, **1850**:1059-1071.

89. Kara M, Zacharias M: **Influence of 8-oxoguanosine on the fine structure of DNA studied with biasing-potential replica exchange simulations**. *Biophys J* 2013, **104**:1089-1097.

90. Geršata T, Kara M, Zacharias M, Lankaš F: **Effect of 8-oxoguanine on DNA structure and deformability**. *J Phys Chem B* 2013, **117**:11617-11622.

91. Krepl M, Otyepka M, Banáš P, Šponer J: **Effect of guanine to inosine substitution on stability of canonical DNA and RNA duplexes: molecular dynamics thermodynamics integration study**. *J Phys Chem B* 2013, **117**:1872-1879.

92. Kuznetsov NA, Bergonzo C, Campbell AJ, Li H, Mechetin GV, de los Santos C, Grollman AP, Fedorova OS, Zharkov DO, Simmerling C: **Active destabilization of base pairs by a DNA glycosylase wedge initiates damage recognition**. *Nucleic Acids Res* 2015, **43**:272-281.

93. Rossetti G, Dans PD, Gomez-Pinto I, Ivani I, Gonzalez C, Orozco M: **The structural impact of DNA mismatches**. *Nucleic Acids Res* 2015, **43**:4309-4321.

94. DeRocco VC, Sass LE, Qiu R, Weninger KR, Erie DA: **Dynamics of MutS-mismatched DNA complexes are predictive of their repair phenotypes**. *Biochemistry* 2014, **53**:2043-2052.

95. Guo X, Liu Z, Song Q, Wang L, Zhong D: **Dynamics and mechanism of UV-damaged DNA repair in indole-thymine dimer adduct: molecular origin of low repair quantum efficiency**. *J Phys Chem B* 2015, **119**:3446-3455.

96. Dumont E, Geršata T, Guerra CF, Lankaš F: **Insights into the structure of intrastrand cross-link DNA lesion-containing oligonucleotides: G[8-5m]T and G[8-5]C from molecular dynamics simulations**. *Biochemistry* 2015, **54**:1259-1267.

97. Bianchi C, Zangi R: **Molecular dynamics study of the recognition of dimethylated CpG sites by MBD1 protein**. *J Chem Inf Model* 2015, **55**:636-644.

98. Yang J, Lior-Hoffmann L, Wang S, Zhang Y, Broyde S: **DNA cytosine methylation: structural and thermodynamic characterization of the epigenetic marking mechanism**. *Biochemistry* 2013, **52**:2828-2838.

99. Carvalho ATP, Gouveia L, Kanna CR, Wärmländer SKTS, Platts JA, Kamerlin SCL: **Understanding the structural and dynamic consequences of DNA epigenetic modifications: computational insights into cytosine methylation and hydroxymethylation**. *Epigenetics* 2014, **9**:1604-1612.

100. Severin PMD, Zou X, Gaub HE, Schulten K: **Cytosine methylation alters DNA mechanical properties**. *Nucleic Acids Res* 2011, **39**:8740-8751.

101. Pérez A, Castellazzi CL, Battistini F, Collinet K, Flores O, Deniz O, Ruiz ML, Torrents D, Eritja R, Soler-López M *et al.*: **Impact of methylation on the physical properties of DNA**. *Biophys J* 2012, **102**:2140-2148.

102. Portella G, Battistini F, Orozco M: **Understanding the connection between epigenetic DNA methylation and nucleosome positioning from computer simulations**. *PLoS Comput Biol* 2013, **9**:e1003354.

103. Jimenez-Useche I, Yuan C: **The effect of DNA CpG methylation on the dynamic conformation of a nucleosome**. *Biophys J* 2012, **103**:2502-2512.

104. Esposito L, Banyasz A, Douki T, Perron M, Markovitsi D, Improta R: **Effect of C5-methylation of cytosine on the photoreactivity of DNA: a joint experimental and computational study of TCG trinucleotides**. *J Am Chem Soc* 2014, **136**:10838-10841.

105. Singh RP, Blossey R, Cleri F: **Structure and mechanical characterization of DNA i-motif nanowires by molecular dynamics simulation**. *Biophys J* 2013, **105**:2820-2831.

106. Luo D, Mu Y: **All-atomic simulations on human telomeric G-quadruplex DNA binding with thioflavin T**. *J Phys Chem B* 2015, **119**:4955-4967.

107. Rebič M, Mocci F, Laaksonen A, Uličný J: **Multiscale simulations of human telomeric G-quadruplex DNA**. *J Phys Chem B* 2015, **119**:105-113.

108. Islam B, Sgobba M, Laughton C, Orozco M, Sponer J, Neidle S, Haider S: **Conformational dynamics of the human propeller telomeric DNA quadruplex on a microsecond time scale**. *Nucleic Acids Res* 2013, **41**:2723-2735.

109. Hospital A, Andrio P, Cugnasco C, Codo L, Becerra Y, Dans PD, Battistini F, Torres J, Goñi R, Orozco M *et al.*: **BIGNASim: a NoSQL database structure and analysis portal for nucleic acids simulation data**. *Nucleic Acids Res* 2015 http://dx.doi.org/10.1093/nar/gkv1301.

110. Yonetani Y, Kono H: **Dissociation free-energy profiles of specific and nonspecific DNA–protein complexes**. *J Phys Chem B* 2013, **117**:7535-7545.

111. Furini S, Barbini P, Domene C: **DNA-recognition process described by MD simulations of the lactose repressor protein on a specific and a non-specific DNA sequence**. *Nucleic Acids Res* 2013, **41**:3963-3972.

112. • May ER, Arora K, Brooks CL: **pH-induced stability switching of the bacteriophage HK97 maturation pathway**. *J Am Chem Soc* 2014, **136**:3097-3107.

A nice study showing at the atomistic level how pH oscillations can control packing of DNA.

113. Galindo-Murillo R, Garcia-Ramos JC, Ruiz-Azuara L, Cheatham TE, Cortes-Guzman F: **Intercalation processes of copper complexes in DNA**. *Nucleic Acids Res* 2015, **43**:5364-5376.

114. De Ruiter A, Zagrovic B: **Absolute binding-free energies between standard RNA/DNA nucleobases and amino-acid sidechain analogs in different environments**. *Nucleic Acids Res* 2015, **43**:708-718.

115. Erler J, Zhang R, Petridis L, Cheng X, Smith JC, Langowski J: **The role of histone tails in the nucleosome: a computational study**. *Biophys J* 2014, **107**:2911-2922.

116. Winogradoff D, Echeverria I, Potoyan DA, Papoian GA: **The acetylation landscape of the H4 histone tail: disentangling the interplay between the specific and cumulative effects**. *J Am Chem Soc* 2015, **137**:6245-6253.

117. •• Collepardo-Guevara R, Portella G, Vendruscolo M, Frenkel D, Schlick T, Orozco M: **Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: a multiscale computational study**. *J Am Chem Soc* 2015, **137**:10205-10215.

Experiments, atomistic MD and coarse-grained models are combined to decipher the mechanistic basis of some histone epigenetic variants. The impact of lysine acetylation on histone tails is investigated with molecular dynamics and transferred to the coarse-grained model showing that this epigenetic change reduces inter-nucleosomal interactions and promotes opening of the chromatin fiber.

118. Cai Y, Kropachev K, Terzidis MA, Masi A, Chatgilialoglu C, Shafirovich V, Geacintov NE, Broyde S: **Differences in the access of lesions to the nucleotide excision repair machinery in nucleosomes**. *Biochemistry* 2015, **54**:4181-4185.

119. Sim AYL, Minary P, Levitt M: **Modeling nucleic acids**. *Curr Opin Struct Biol* 2012, **22**:273-278.

120. Potoyan D, Savelyev aA, Papoian Ga: **Recent successes in coarse-grained modeling of DNA**. *Wiley Interdiscip Rev Comput Mol Sci* 2013, **3**:69-83.

121. Ingólfsson HI, Lopez Ca, Uusitalo JJ, de Jong DH, Gopal SM, Periole X, Marrink SJ: **The power of coarse graining in biomolecular simulations**. *Wiley Interdiscip Rev Comput Mol Sci* 2013, **4**:225-248.

122. Noid WG: **Perspective: coarse-grained models for biomolecular systems**. *J Chem Phys* 2013, **139**:090901.

123. Nicolaï A, Delarue P, Senet P: **Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes**2014.

124. • Uusitalo JJ, Ingólfsson HI, Akhshi P, Tieleman DP, Marrink SJ: **Martini coarse-grained force field: extension to DNA**. *J Chem Theory Comput* 2015 http://dx.doi.org/10.1021/acs.jctc.5b00286.

An extension of the MARTINI force-field for DNA. This new incorporation to the unified MARTINI FF allowed for the simulation of Protein–DNA complexes in explicit solvent and ions.

125. •• Machado MR, Pantano S: **Exploring the LacI-DNA dynamics by multiscale simulations using the SIRAH force field**. *J Chem Theory Comput* 2015 http://dx.doi.org/10.1021/acs.jctc.5b00575.

One of the most outstanding simulation, to date, using an hybrid MM/CG approach for the solutes (protein–DNA), and for the solvent (the waters and the ions are modelled at two levels MM/CG simultáneously, using the WT4 model from the SIRAH force-field.).

126. Naômé A, Laaksonen A, Vercauteren DP: **A coarse-grained simulation study of the structures, energetics, and dynamics of linear and circular DNA with its ions**. *J Chem Theory Comput* 2015 http://dx.doi.org/10.1021/acs.jctc.5b00113.

127. Korolev N, Luo D, Lyubartsev AP, Nordenskiöld L: **A coarse-grained DNA model parameterized from atomistic simulations by inverse Monte Carlo**. *Polymers (Basel)* 2014, **6**:1655-1675.

128. Maciejczyk M, Spasic A, Liwo A, Scheraga HA: **DNA duplex formation with a coarse-grained model**. *J Chem Theory Comput* 2014, **10**:5020-5035.

129. Cragnolini T, Derreumaux P, Pasquali S: **Coarse-grained simulations of RNA and DNA duplexes**. *J Phys Chem B* 2013, **117**:8047-8060.

130. Hinckley DM, Freeman GS, Whitmer JK, De Pablo JJ: **An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: structure, thermodynamics, and dynamics of hybridization**. *J Chem Phys* 2013, **139**:144903.

131. Freeman GS, Hinckley DM, Lequieu JP, Whitmer JK, de Pablo JJ: **Coarse-grained modeling of DNA curvature**. *J Chem Phys* 2014, **141**:165103.

132. Hinckley DM, Lequieu JP, De Pablo JJ: **Coarse-grained modeling of DNA oligomer hybridization: length, sequence, and salt effects**. *J Chem Phys* 2014, **141**:035102.

133. Frederickx R, In'T Veld T, Carlon E: **Anomalous dynamics of DNA hairpin folding**. *Phys Rev Lett* 2014, **112**:1-5.

134. Li NK, Kim HS, Nash J, Lim M, Yingling YG: **Progress in molecular modelling of DNA materials**. *Mol Simul* 2014 http://dx.doi.org/10.1080/08927022.2014.913792.

135. Ouldridge TE: **DNA nanotechnology: understanding and optimisation through simulation**. *Mol Phys* 2014, **113**:1-15.

136. Doye JPK, Ouldridge TE, Louis AA, Romano F, Šulc P, Matek C, Snodin BEK, Rovigatti L, Schreck JS, Harrison RM *et al.*: **Coarse-graining DNA for simulations of DNA nanotechnology**. *Phys Chem Chem Phys* 2013, **15**:20395-20414.

137. Šulc P, Ouldridge TE, Romano F, Doye JPK, Louis AA: **Simulating a burnt-bridges DNA motor with a coarse-grained DNA model**. *Nat Comput* 2013 http://dx.doi.org/10.1007/s11047-013-9391-8.

138. Ouldridge TE, Šulc P, Romano F, Doye JPK, Louis A: **DNA hybridization kinetics: zippering, internal displacement and sequence dependence**. *Nucleic Acids Res* 2013, **41**:8886-8895.

139. De Pablo JJ: **Coarse-grained simulations of macromolecules: from DNA to nanocomposites**. *Annu Rev Phys Chem* 2011, **62**:555-574.

140. Li H, Wang Z, Li N, He X, Liang H: **Denaturation and renaturation behaviors of short DNA in a confined space**. *J Chem Phys* 2014, **141**:044911.

141. • Stachiewicz A, Molski A: **A coarse-grained MARTINI-like force field for DNA unzipping in nanopores**. *J Comput Chem* 2015 http://dx.doi.org/10.1002/jcc.23874.

This work is interesting due to its nature and the phylosophy behind. The authors combined (after some re-paremeterizations), the DNA model from the SIRAH force-field, with the water and ion model from the MARTINI FF. This combination is applied to a system with nanotechnological interest, despite the fact that SIRAH and MARTINI were both thought to simulate biological systems.

142. Setubal JC, Almeida NF (Eds): *Advances in Bioinformatics and Computational Biology*. Springer International Publishing; 2013.

143. Petkevičiūtė D, Pasi M, Gonzalez O, Maddocks JH: **cgDNA: a software package for the prediction of sequence-dependent coarse-grain free energies of B-form DNA**. *Nucleic Acids Res* 2014, **42**:e153.

144. Dršata T, Lankaš F: **Multiscale modelling of DNA mechanics**. *J Phys Condens Matter* 2015, **27**:323102.

145. Zhou T, Yang L, Lu Y, Dror I, Dantas Machado AC, Ghane T, Di Felice R, Rohs R: **DNAshape: a method for the high-throughput prediction of DNA structural features on a genomic scale**. *Nucleic Acids Res* 2013, **41**:W56-W62.

146. Zhoua T, Shenb N, Yanga L, Abed N, Hortonc J, Mannd RS, Bussemakerf HJ, Gordânc R, Rohs R: **Quantitative modeling of transcription factor binding specificities using DNA shape**. *Proc Natl Acad Sci U S A* 2015, **112**:4654-4659.

147. Abe N, Dror I, Yang L, Slattery M, Zhou T, Bussemaker HJ, Rohs R, Mann RS: **Deconvolving the recognition of DNA sequence from shape**. *Cell* 2015, **161**:307-318.

148. Maffeo C, Ngo TTM, Ha T, Aksimentiev A: **A coarse-grained model of unstretched single-stranded DNA derived from atomistic simulation and single-molecule experiment**. *J Chem Theory Comput* 2014 http://dx.doi.org/10.1021/ct500193u.

149. Markegard CB, Fu IW, Reddy KA, Nguyen HD: **Coarse-grained simulation study of sequence effects on DNA hybridization in a concentrated environment**. *J Phys Chem B* 2015, **119**:1823-1834.

150. He Y, Maciejczyk M, Ołdziej S, Scheraga HA, Liwo A: **Mean-field interactions between nucleic-acid-base dipoles can drive the formation of a double helix**. *Phys Rev Lett* 2013, **110**:098101.

151. Maffeo C, Yoo J, Comer J, Wells DB, Luan B, Aksimentiev A: **Close encounters with DNA**. *J Phys Condens Matter* 2014, **26**:413101.

152. Fu IW, Markegard CB, Chu BK, Nguyen HD: **The role of electrostatics and temperature on morphological transitions of hydrogel nanostructures self-assembled by peptide amphiphiles via molecular dynamics simulations**. *Adv Healthc Mater* 2013, **2**:1388-1400.

153. Liwo A, Khalili M, Scheraga HA: **Ab initio simulations of protein-folding pathways by molecular dynamics with the united-residue model of polypeptide chains**. *Proc Natl Acad Sci U S A* 2005, **102**:2362-2367.

154. Darré L, Machado MR, Brandner AF, González HC, Ferreira S, Pantano S: **SIRAH: a structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics**. *J Chem Theory Comput* 2015, **11**:723-739.

155. Horowitz-Scherer RA, Woodcock CL: **Organization of interphase chromatin**. *Chromosoma* 2005, **115**:1-14.

156. Collepardo-Guevara R, Schlick T: **Chromatin fiber polymorphism triggered by variations of DNA linker lengths**. *Proc Natl Acad Sci U S A* 2014, **111**:8061-8066.

157. Fan Y, Korolev N, Lyubartsev AP, Nordenskiöld L: **An advanced coarse-grained nucleosome core particle model for computer simulations of nucleosome-nucleosome interactions under varying ionic conditions**. *PLOS ONE* 2013, **8**:e54228.

158. Nishino Y, Eltsov M, Joti Y, Ito K, Takata H, Takahashi Y, Hihara S, Frangakis AS, Imamoto N, Ishikawa T *et al.*: **Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure**. *EMBO J* 2012, **31**:1644-1653.

159. Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu R-M, Zhu P, Li G: **Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units**. *Science* 2014, **344**:376-380.

160. Korolev N, Fan Y, Lyubartsev AP, Nordenskiöld L: **Modelling chromatin structure and dynamics: status and prospects**. *Curr Opin Struct Biol* 2012, **22**:151-159.

161. Ozer G, Luque A, Schlick T: **The chromatin fiber: multiscale problems and approaches**. *Curr Opin Struct Biol* 2015, **31**:124-139.

162. Rosa A, Zimmer C: **Computational models of large-scale genome architecture**. *Int Rev Cell Mol Biol* 2014, **307**:275-349.

163. Nicodemi M, Pombo A: **Models of chromosome structure**. *Curr Opin Cell Biol* 2014, **28**:90-95.

164. Serra F, Di Stefano M, Spill YG, Cuartero Y, Goodstadt M, Baù D, Marti-Renom MA: **Restraint-based three-dimensional modeling of genomes and genomic domains**. *FEBS Lett* 2015 http://dx.doi.org/10.1016/j.febslet.2015.05.012.

165. Müller O, Kepper N, Schöpflin R, Ettig R, Rippe K, Wedemann G: **Changing chromatin fiber conformation by nucleosome repositioning**. *Biophys J* 2014, **107**:2141-2150.

166. Collepardo-Guevara R, Schlick T: **Crucial role of dynamic linker histone binding and divalent ions for DNA accessibility and gene regulation revealed by mesoscale modeling of oligonucleosomes**. *Nucleic Acids Res* 2012, **40**:8803-8817.

167. Collepardo-Guevara R, Schlick T: **The effect of linker histone's nucleosome binding affinity on chromatin unfolding mechanisms**. *Biophys J* 2011, **101**:1670-1680.

168. Kulaeva OI, Zheng G, Polikanov YS, Colasanti AV, Clauvelin N, Mukhopadhyay S, Sengupta AM, Studitsky VM, Olson WK: **Internucleosomal interactions mediated by histone tails allow distant communication in chromatin**. *J Biol Chem* 2012, **287**:20248-20257.

169. Grigoryev SA, Arya G, Correll S, Woodcock CL, Schlick T: **Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions**. *Proc Natl Acad Sci U S A* 2009, **106**:13317-13322.

170. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO *et al.*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **326**:289-293.

171. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping**. *Cell* 2014, **159**:1665-1680.

172. Rouquette J, Cremer C, Cremer T, Fakan S: **Functional nuclear architecture studied by microscopy: present and future**. *Int Rev Cell Mol Biol* 2010, **282**:1-90.

173. • Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J: **Organization of the mitotic chromosome**. *Science* 2013, **342**:948-953.
This article combines results of Hi-C and 5C experiments of mitotic chromosomes with a polymer model. A two stage simulation process for the polymer combining linear compaction and axial compression promotes chromatin loops along a linear conformation which fits best to the experimental data.

174. Barbieri M, Chotalia M, Fraser J, Lavitas L-M, Dostie J, Pombo A, Nicodemi M: **Complexity of chromatin folding is captured by the strings and binders switch model**. *Proc Natl Acad Sci U S A* 2012, **109**:16173-16178.

175. Jost D, Carrivain P, Cavalli G, Vaillant C: **Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains**. *Nucleic Acids Res* 2014, **42**:9553-9561.

176. Gehlen LR, Gruenert G, Jones MB, Rodley CD, Langowski J, O'Sullivan JM: **Chromosome positioning and the clustering of functionally related loci in yeast is driven by chromosomal interactions**. *Nucleus* 2012, **3**:370-383.

177. Gürsoy G, Xu Y, Kenter AL, Liang J: **Spatial confinement is a major determinant of the folding landscape of human chromosomes**. *Nucleic Acids Res* 2014, **42**:8223-8230.

178. Ricci MA, Manzo C, García-Parajo MF, Lakadamyali M, Cosma MP: **Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo**. *Cell* 2015, **160**:1145-1158.

179. • Lyubartsev AP, Korolev N, Fan Y, Nordenskiöld L: **Multiscale modelling of nucleosome core particle aggregation**. *J Phys Condens Matter* 2015, **27**:064111.
In this multi-scale study a super-coarse-grained model of the nucleosome core particle (NCP) is introduced where effective potentials between the super-coarse-grained NCP's are derived from the already available finer NCP model. With the new model aggregation of up to 5000 NCP's under different environmental conditions are studied.

180. Arya G, Zhang Q, Schlick T: **Flexible histone tails in a new mesoscopic oligonucleosome model**. *Biophys J* 2006, **91**:133-150.

181. Perišić O, Collepardo-Guevara R, Schlick T: **Modeling studies of chromatin fiber structure as a function of DNA linker length**. *J Mol Biol* 2010, **403**:777-802.

182. Luque A, Collepardo-Guevara R, Grigoryev S, Schlick T: **Dynamic condensation of linker histone C-terminal domain regulates chromatin structure**. *Nucleic Acids Res* 2014, **42**:7553-7560.

183. • Kimura H, Shimooka Y, Nishikawa J, Miura O, Sugiyama S, Yamada S, Ohyama T: **The genome folding mechanism in yeast**. *J Biochem* 2013, **154**:137-147.
This work shows that with a comparatively simple mesoscale model of nucleosome and linker DNA chromosomal compaction in yeast can accurately be described. The resulting nucleosomal arrays -flexibility and persistence length of linker DNA fragments are derived experimentally- illustrate that the architecture of chromosomes in yeast is principally determined by physical properties of DNA and the nucleus size.

184. Knotts TA, Rathore N, Schwartz DC, de Pablo JJ: **A coarse grain model for DNA**. *J Chem Phys* 2007, **126**:084901.