FEBS Letters 589 (2015) 2987–2995

FEBS Letters

journal homepage: www.FEBSLetters.org

Review

Restraint-based three-dimensional modeling of genomes and genomic domains

François Serra [a,b], Marco Di Stefano [a,b], Yannick G. Spill [a,b], Yasmina Cuartero [a,b], Michael Goodstadt [a,b], Davide Baù [a,b], Marc A. Marti-Renom [a,b,c,*]

[a] Genome Biology Group, Centre Nacional d'Anàlisi Genòmica (CNAG), Barcelona, Spain
[b] Gene Regulation, Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), Barcelona, Spain
[c] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

ARTICLE INFO

Article history:
Received 7 April 2015
Revised 5 May 2015
Accepted 5 May 2015
Available online 14 May 2015

Edited by Wilhelm Just

Keywords:
Genome architecture
3D genome reconstruction
Chromosome Conformation Capture
Restraint-based modeling

ABSTRACT

Chromosomes are large polymer molecules composed of nucleotides. In some species, such as humans, this polymer can sum up to meters long and still be properly folded within the nuclear space of few microns in size. The exact mechanisms of how the meters long DNA is folded into the nucleus, as well as how the regulatory machinery can access it, is to a large extend still a mystery. However, and thanks to newly developed molecular, genomic and computational approaches based on the Chromosome Conformation Capture (3C) technology, we are now obtaining insight on how genomes are spatially organized. Here we review a new family of computational approaches that aim at using 3C-based data to obtain spatial restraints for modeling genomes and genomic domains.

1. Introduction

Genomes are often compared to libraries where the genetic information is stored in the form of books and text represents the linear sequence of the genome. Unfortunately, that linear (1D) representation omits the utterly complex three-dimensional (3D) organization of the genome. Indeed, the physical support of the genome (i.e., the books, the shelves, the corridors and the library building in our metaphor) may be as important as the functional elements it encodes [1]. It is now known that the dynamic structure of the complex gene networks in a genome regulates the orchestration of fundamental biological processes such as development [2], cell differentiation [3,4] or response to stimuli [5], among others. Moreover, most of such complex mechanisms are also among the most conserved features of our genomes [6,7]. Therefore, addressing the 3D structure of a genome may provide insights into fundamental questions like the *C*-value paradox [8] or the regulatory divergence between closely related species [9].

In the past decade, with the introduction and development of Chromosome Conformation Capture (3C) technologies (e.g., 3C [10], 4C [11], 5C [12], Hi-C [13], *in situ*-Hi-C [7], TCC [14], T2C [15] or Capture-C [16,17], which are here referred as 3C-based technologies), it has been possible to get insight into how the genome folds by interrogating physical interactions within the genome. Importantly, the combination of these 3C-based technologies with advanced imaging [18] has helped reducing the resolution gap in genome structure [19]. It is now known that the genome organizes into chromosome territories [20], which in turn are spaced into two compartment types [13] composed of finer units called Topologically Associating Domains or TADs [6,21,22]. Alongside these advances, the evidence that genome structure is tightly associated with its function was being reinforced by the comparison with chromatin epigenetic states [7,23]. However, two limitations are blurring the full picture of the genome organization. First, some of the emerging genomic features change depending on the scale at which we study the genome. For example, TADs are structural units that were shown to be robustly detectable over a large range of genomic resolutions. Yet, their existence is challenged when the genome is interrogated at finer scales [7]. Second, 3C-based experiments are usually carried out with tens of millions of cells, and thus are population-based measures superimposing millions of partial

* Corresponding author at: Genome Biology Group, Centre Nacional d'Anàlisi Genòmica (CNAG), Barcelona, Spain.
E-mail address: mmarti@pcb.ub.cat (M.A. Marti-Renom).

http://dx.doi.org/10.1016/j.febslet.2015.05.012
0014-5793/

snapshots of interacting chromatin fragments. All together, this indicates that the analysis of 3C-based interaction matrices requires rigorous statistical treatment that both, preserves the genuine data and removes inherent biases that can lead to misconceptions [24,25]. For instance, and due to the assumption that observed interactions between loci pairs are independent, parametric test on the co-localization of triplets of loci in genome-wide Hi-C matrices lead to the false discovery of significant interactions [26,27]. Providing a consolidated framework for 3D interpretation (i.e., chromatin modeling) of 3C-based interaction matrices could help to circumvent these issues and characterize the principles of chromatin 3D folding [28,29].

Chromatin modeling is performed using two main complementary strategies [19]. The first, known as restraint-based (RB) modeling, interprets the 3C-based data as a set of spatial restraints to build a 3D model of the chromatin fiber by satisfying the input restraints. The second, called thermodynamics-based (TB) modeling, applies polymer physics principles of the chromatin fiber to identify plausible spatial arrangements of the chromatin fiber. For example, TB approaches have been implemented to interpret the decay of interaction frequencies with the genomic distance [30], the formation of domains of active and inactive chromatin [31], epigenetic features like chromatin colors [32], chromosome territories [33–35] and co-expression data [36]. TB modeling is reviewed in separate articles in this issue.

This review will focus on RB modeling, which in contrast to TB modeling is a more recent addition to the genome structure determination field. RB modeling can be broadly classified into two main categories depending on how the restraints are implemented. The first category uses analytical approaches to directly transform an interaction matrix into a 3D object [37,38]. These approaches assume that the interaction matrix represents a single consensus structure. They are, therefore, more suitable for single-cell 3C-based studies [39]. The second category defines a set of restraints from the observed interaction data, which are then satisfied using either Monte Carlo sampling [14,40–44] or Bayesian [45–48] approaches. In contrast to the analytical methodologies, these optimization-based methods are usually used to generate thousands of possible conformations of the target chromatin region. This second category can further be divided into (a) approaches where simulations are independent of each other resulting in a population of models, which is expected to represent the variability of chromatin conformations, and (b) population-based strategies whose restraints take into account a given number of representative models simultaneously. Most of these methods have already been successfully applied to describe a large number of biological features of the chromatin fiber. For example: the presence of compartments and their tendency to aggregate by type [45], the identification of signatures of chromatin conformations specific to each cell type (which correspond to the segregation of active and inactive chromatin [49]), the identification of chromatin globules in the α-globin domain of the human genome [50,51], the ellipsoidal organization of the *Caulobacter crescentus* chromosome [52], the overall genomic organization of the yeast [53], the human [14] and the *Plasmodium* genome [54], or the spatial organization of the X inactivation center in the human genome [40].

Through this review we will overview all the steps from the acquisition of 3C-based interaction data to the reconstruction of 3D models of chromatin (Fig. 1). This walkthrough will help us to assess the methodological choices made by today's research groups. We will finally discuss the importance of conventions in the context of a rising demand for data sharing and model visualization.

2. Data generation

3C-based experiments allow measuring the frequency at which pairs of chromatin regions physically interact in the 3D space. This frequency is proportional to the propensity for two loci to be

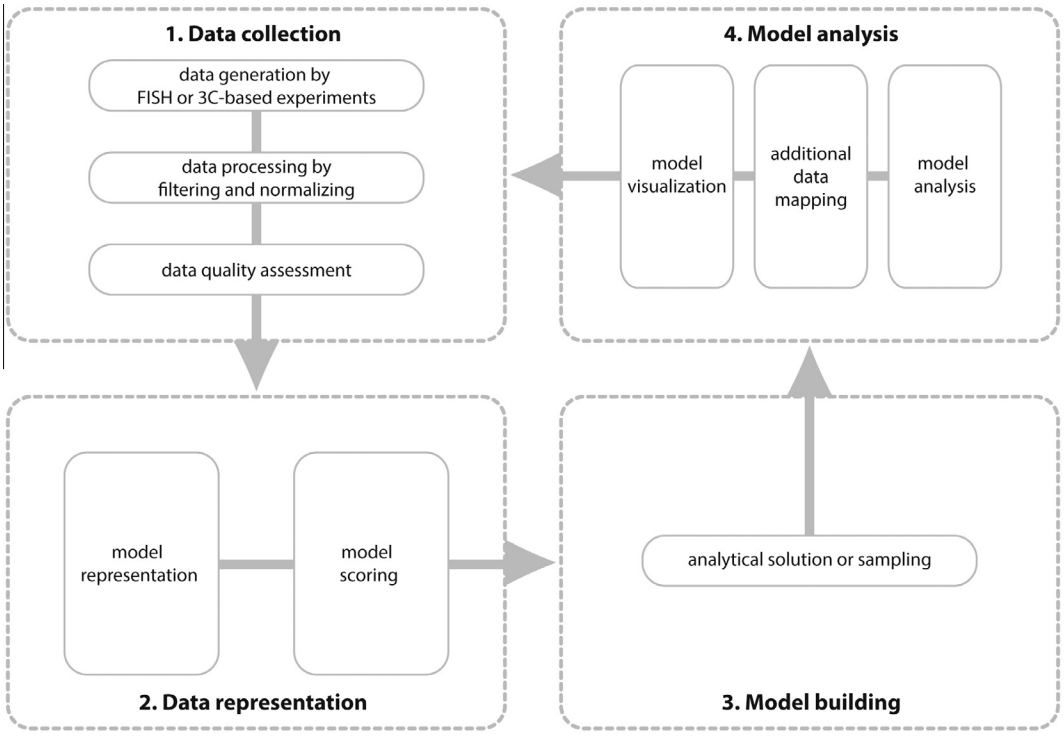

Fig. 1. The four stages of the integrative modeling cycle [63]. Those include data collection, data representation, model building and model analysis. The flow chart shows specific steps that apply to the restraint-based modeling of genomes and genomic domains.

cross-linked by formaldehyde [55]. In a typical 3C-based experiment, the chromatin of tens of millions of cells are cross-linked and then fragmented by digestion with a restriction enzyme. These cross-linked fragments are then ligated to form hybrid DNA molecules containing parts of the DNA from each side of the captured interaction. The various 3C-based methods differ, mainly, in the processes used to capture and detect the hybrid DNA molecules after ligation. In 3C experiments, ligation products are detected by PCR using locus specific primers. 4C uses inverse PCR to generate genome wide interaction profiles for single loci. 5C combines 3C with hybrid capture approaches by annealing and ligating oligonucleotides in a multiplex setting. Hi-C was the first unbiased and genome-wide adaptation of 3C. This method includes a step to introduce biotinylated nucleotides at digested sites, which facilitate the isolation of ligation products and increase subsequent sequencing efficacy. Three main variants of this protocol have been described so far. Each differs in the conditions in which the ligation step is performed: (1) in extremely diluted conditions (dilution Hi-C [13]), (2) on a solid phase support (Tethered Conformation Capture [14]), or (3) inside the nucleus (*in situ*-Hi-C [7]). Finally, Capture-C and T2C combines oligonucleotide capture technology with 3C and high-throughput sequencing [15–17]. For all the 3C-based technologies, the output of the experiment is an interaction matrix that contains the frequency of interaction between pairs of chromatin loci. Despite their contribution to the field, 3C-based technologies have two limitations. First, the experiments result in an interaction matrix representing an average picture of the conformational states that chromatin may adopt in each of the millions of cells used in the experiment. Second, current protocols can capture only about 2.5% of all possible interactions occurring in a cell [39,56]. An attempt in addressing the first point was recently accomplished with Single-Cell Hi-C maps [39]. However, the map produced for a single cell still suffers from the second limitation.

Besides 3C-based technologies, light and electron microscopy are also providing direct observations of the structural organization of genomes [57,58]. These experiments have been previously used to validate aspects of the models obtained from 3C-based experiments. Although light microscopy is subjected to the diffraction limit, which makes objects closer than 200 nm difficult to separate, there are new techniques that make it possible to reach resolutions down to 10–20 nm [59]. These sub-diffraction techniques are based on sampling observations by varying the excitation light source. Despite the many advances in high-resolution light microscopy, the folding and the compaction level of the chromatin fiber are still hardly distinguishable using microscopy. The resolution limit of electron microscopy, however, is below the scale of the chromatin fiber and would therefore make such a direct observation possible. Unfortunately, electron microscopy is very damaging, and offers no possibility of observation in living cells. Nevertheless, electron microscopy has been used to study the structure of the chromatin fiber at the nucleosome level to characterize the folding properties of the chromatin fiber [60] or to reveal the exact nucleosome positioning inside a segment of chromatin fiber [61]. Electron microscopy has also been used in the context of 3C experiments to characterize the underlying mechanisms of the 3C protocol [62].

Currently, in the context of modeling, light microscopy has been used to validate 3D models built using 3C-based data [5,50]. However, the use of integrative approaches for modeling (such as the Integrative Modeling Platform [63]) makes it possible to use data from several independent sources to build robust models with counterbalanced experimental biases. Such input data could include observations from microscopy, or general knowledge, such as the nucleus size [14,41,54]. Other types of restraints that may be integrated in the future include epigenetic knowledge such as positioning of lamina associating domains [64] or mechanical properties of chromosomes [65].

3. Data processing

3C-based experiments, in particular Hi-C, result in hundreds of millions of paired-end reads, whose typical length is about one hundred base pairs. After sequencing, these reads are filtered to remove experimental biases [66]. Such filtering consists of two main steps. First, it removes reads for which one or both ends were not uniquely mapped to the reference genome. It has been previously proposed that most of the single-end mapped reads would essentially belong to centromeric or telomeric regions and thus could be informative during 3D modeling [25]. Second, the filtering removes reads in which both ends are mapped into the same restriction enzyme fragment. These reads are classified into un-ligated fragments (dangling-ends, DE) or self-circularized ligation products (self-circles, SC). In both cases, the reads are discarded based on the assumption that (i) the probability of mapping homologous regions is negligible, (ii) the distribution of DE and SC is likely uniform along the genome and thus non-informative for the normalization of the data, and (iii) local features of the chromatin structure will not affect the relative proportions of DE or SC. Typically, after the filtering process, between 40% and 80% of the sequenced reads are removed. However, these discarded reads could be recycled in the context of data quality control. For example, it has been proposed that the estimation of the amount of interactions between mitochondrial DNA and genomic DNA can assess the specificity of the ligation step in 3C-based experiments [38].

Once these expected experimental biases are filtered out, a more "fine-grained" correction of the data is required to normalize the interaction frequencies. This additional filtering, called normalization, aims at removing biases that have not a clear experimental origin but arise from non-homogeneity of the genome. One of the most elegant implementations of the normalization step is based on a probabilistic non-parametric method of effective bias removal [24]. Despite its elegant implementation, it has been replaced in practice by more straightforward and fast approaches such as the iterative correction and eigenvector decomposition (ICE) [25]. Indeed, the ICE method has been extensively used for normalizing 3C-based data in the context of 3D modeling [14,37,41,47]. Yet an alternative normalization method has been recently proposed, which estimates the impact of the biases using the number of restriction enzyme cuts, the GC content and the uniqueness of the mapping of each read on a given genomic bin [45]. Despite their differences, all the normalization approaches aim at identifying and reducing the experimental biases that could have an effect on the probability of capturing an interaction between two genomic loci. It is important to note that this is a field of interest that is still evolving and new approaches are likely to appear in the near future.

4. Data quality assessment

Data quality assessment in 3C-based experiments relies on few statistical measures of the interaction matrix with the main objective of assessing whether the sequencing depth was enough to sample the interactions in all the cells in the experiment. As a rule of thumb, it has been recommended that 90% of the cells in an interaction matrix should contain non-zero values [13], or that 80% of the loci have at least 1000 contacts when the number of input reads is relatively high [7]. Once the resolution of the data has been set (i.e., the matrix has been binned to the size at which the previous rules are fulfilled), other indicators can be used to

evaluate the matrix quality. For instance, the proportion of intra- versus inter-chromosomal contacts is expected to oscillate between 40% and 60% [14,66]. Furthermore, it has been observed that, in eukaryotic cells, the decay of the interaction frequency with the genomic distance, when averaged over the sample, follows a slope of around −1.0 for interphase chromosomes [13,28,66]. However, it is important to note that this slope can vary during the cell cycle of the organism [67] and between organisms.

In the specific context of modeling, assessing whether a matrix is suitable for 3D reconstruction of the genome or a genomic region is of primary importance and is still debated. The Consensus Index [37] can be used to assess *a priori* the quality of the interaction matrix for modeling. The Consensus Index measures how far a given matrix is from satisfying all triangular inequalities. Triangular inequalities occur when the distance between two points is higher then the sum of the distances of these two points to a third one, and their satisfaction is especially important in the context of non-population based modeling strategies [37,38]. More recently, the Matrix Modeling Potential score, combining statistical properties of the interaction matrices, has been used to assess the potential of the interaction matrix for restraint-based 3D modeling [68].

5. Three-dimensional (3D) model reconstruction

Genomic 3D reconstruction from 2D interaction maps consists of two main steps: (i) transforming the observed interaction data into a set of Euclidian distances or spatial contacts and (ii) finding a 3D conformation in Cartesian space that satisfies the imposed distances or contacts. In theory, these two steps can be achieved by classic multidimensional scaling (MDS), which has been previously used in protein structure determination [69,70]. However, and in contrast to protein structure determination, genome reconstruction based on MDS approaches has intrinsic limitations. For example, 3C-based experiments contain significant amounts of noise as well as structural variability. This variability is due to the fact that they are performed over tens of millions of cells, and because the chromatin fiber is very flexible and can adopt multiple conformations. All together, the resulting experimental 3C-based matrices contain many triangular inequalities, which can only be satisfied through alternative optimization-based 3D reconstructions. In this section, we discuss and classify different methods to solve the 3D reconstruction problem (Table 1); we describe the different 3D reconstruction methods by the way they represent chromatin, how the independent solutions are scored and how the solutions are found.

5.1. Representation

The first important choice for 3D modeling of chromatin is the way the chromatin fiber would be represented as a physical object (polymer) of consecutive particles (monomers). Although in principle different shapes can be used for a single particle [71], most of the modeling strategies use either point particles [37,38,45,47,49] or spheres of a given radius [14,40,41,43,44,48]. Usually, the genomic size of each of these segments is decided during the 3C-based data processing as previously discussed, and the final model is a homo-polymer where each particle has the same size (i.e., binned interaction matrices). However, depending on the experimental data or their analysis, the model can be also represented by a hetero-polymer with particles of varying size. For example, a 5C experiment results in interaction matrices in which the bins map onto the restriction fragments left after digestion, which have different sizes. As a consequence, the modeling usually involves a polymer where particle radii are proportional to the amount of DNA in the fragment [72]. Even when the input data are equally binned matrices, hetero-polymers can still been used for modeling. For instance, the entire human genome has been represented with 428 particles of varying size, where each particle was obtained by clustering a binned interaction matrix based on its cross-correlation [14]. Having particles of varying radii therefore allows adapting the models to the resolution of the data.

5.2. Scoring

The next step in 3D modeling of genome and genomic domains is to devise a scoring function that will be optimized to yield the Cartesian coordinates of all particles in the model. The scoring function is thus used during sampling to identify, within a set of possible conformations, those that best satisfy the input data. It is important to note that scoring is not needed when no optimization is performed [37,38]. In this review we define the generic scoring function as:

$$S = U_{3C} + U_{Biol} + U_{Phys}$$

where U_{3C} accounts for the restraints inferred from the input 3C-based data, U_{Biol} scores restraints from other experimental observations independent from 3C-based data, and U_{Phys} scores restraints based on general polymer physics properties. The final goal of the sampling is to find a conformation of all particles that minimizes the value of S, that is, where the imposed restraints are minimally violated. Next, we describe in more detail each of the terms of this generic scoring function.

Restraints imposed by the 3C-based data (U_{3C}) are inferred using an inverse relationship with the frequency of interaction between two particles in the model (that is, between two bins or restriction fragments in the genome). In generic terms, it can be expressed as:

$$D_{ij} \propto \left(\frac{1}{F_{ij}}\right)^{\alpha}$$

where D_{ij} is the spatial distance between particle i and particle j, F_{ij} is the normalized interaction frequency measured in the 3C-based experiment, and α an exponent parameter that, depending on the method, is either fixed, estimated or empirically optimized. The α parameter takes values close to 1 when it is fixed prior to modeling, because it matches the characteristic exponent of the crumpled or fractal globule model [13,30,73]. This model is valid for distances in the range of 0.7–10 Mb. Nonetheless, α may vary for other polymer models like, for instance, the equilibrium model where the nominal value is 0.67 [35]. Indeed, several studies point out that α may vary depending on many aspects of the experiment, including cell synchronization [66], clades [47] or chromatin state [44]. When α is optimized, the conversion of interaction frequencies into distances relies on several assumptions that can arguably bias the modeling process as well. The most relevant approximations are the folding of the DNA into a 30 nm fiber, whose existence in vivo is still controversial [74], and the level of chromatin compaction, which can vary from 6 to 11 nucleosomes per 11 nm of chromatin [61,75–78]. An elegant way to circumvent the conversion of interaction frequencies into distances is to compare directly the observed interaction frequencies to the simulated contacts of the modeled chromatin structure [14,40]. The drawback of this method is that it is, *a priori*, scale less, but the final size/occupancy can be still recovered *a posteriori* using nucleus size or chromatin density.

Restraints imposed by other biological data or knowledge (U_{Biol}) rely on direct observations of the cells used for the experiment. For example, the nucleus/cell size and shape [14,41,43,52] or the relative positioning of the nucleolus [41] could be used to confine the model to a given physical space. Other biological restraints can rely

Table 1
Summary of different modeling strategies. F_{ij} is the observed interaction frequency between two particles i and j, D_{ij} is the target distance usually inferred from F_{ij} and r_{ij} is the distance computed on the models. N is the total number of particles.

Method *available online	Representation	Scoring				Sampling	Models
		U_{3C}		U_{Biol}	U_{Phys}		
		$F_{ij} \rightarrow D_{ij}$ conversion	Functional form				
ChromSDE* [37]	Points	$D_{ij} = \begin{cases} \left(\frac{1}{F_{ij}}\right)^{\alpha} \text{ if} F_{ij} > 0 \\ \infty \text{ if } F_{ij} = 0 \end{cases}$ α is optimized	$\sum_{(i,j\|D_{ij}<\infty)} \frac{(r_{ij}^2 - D_{ij}^2)}{D_{ij}} - \lambda \sum_{(i,j)} r_{ij}^2$ where λ is set to 0.01	N/A	N/A	Deterministic semidefinite programming to find the coordinates	Consensus
ShRec3D* [38]	Points	$D_{ij} = \begin{cases} \left(\frac{1}{F'_{ij}}\right)^{\alpha} & \text{if } F'_{ij} > 0 \\ \frac{N^2}{\sum_{(i,j)} F'_{ij}} & \text{if } F'_{ij} = 0 \end{cases}$ F'_{ij} is the original F_{ij} corrected to satisfy all triangular inequalities with the shortest path reconstruction	N/A	N/A	N/A	Deterministic transformations of D_{ij} into coordinates	Consensus
TADbit* [43]	Spheres	$D_{ij} \propto \begin{cases} \alpha F_{ij} + \beta & \text{if } F_{ij} < \gamma' \text{ or } F_{ij} > \gamma \\ \frac{S_i + S_j}{2} & \text{if } \|i-j\| = 1 \end{cases}$ α and β are estimated from the max and the min F_{ij}, from the optimized max distance and from the resolution. $\gamma' < \gamma$ are optimized too. s_i is the radius of particle i	$\sum_{(i,j)} k_{ij}(r_{ij} - D_{ij})^2$ where $k_{ij} = 5$ if $\|i-j\| = 1$ or proportional to F_{ij} otherwise	Yes	U_{excl} and U_{bond} have harmonic forms	Monte Carlo (MC) sampling with Simulated annealing and Metropolis scheme	Resampling
BACH* [45]	Points	$D_{ij} \propto \frac{B_i B_j}{F_{ij}^{\alpha}}$. The biases B_i and B_j and α are optimized	$b_{ij} D_{ij}^{1/\alpha} + c_{ij} \log(D_{ij})$ where b_{ij} and c_{ij} are optimized parameters	No	No	Sequential importance and Gibbs sampling with hybrid MC and adaptive rejection	Population
Giorgetti et al. [40]	Spheres	Particles interact with pair-wise well potentials of depths B_{ij} and contact radius a, which is larger than a hard-core radius and smaller than a maximum contact radius. The parameters are optimized over all the population of models		No	N/A	MC sampling with metropolis scheme	Population
Duan et al. [41]	Spheres	$\overline{F_{\|i-j\|}} = \frac{\sum_{k=0}^{N-\|i-j\|} F_{(k,k+\|i-j\|)}}{N-\|i-j\|}$ is the average of F_{ij} at genomic distance $\|i-j\|$ expressed in kb. $D_{ij} = \overline{F_{\|i-j\|}} \times 7.7 \times \|i-j\|$ assuming that α 1 kb maps onto 7.7 nm	$\sum_{(i,j)} (r_{ij} - D_{ij})^2$	Yes	U_{excl} and U_{bond} have harmonic forms	Interior-point gradient-based method	Resampling
MCMC5C* [49]	Points	$D_{ij} \propto \frac{1}{F_{ij}^{\alpha}}$ where is optimized	$\sum_{(i,j)} (F_{ij} - r_{ij}^{-1/\alpha})^2$	N/A	N/A	MC sampling with Markov chain based algorithm	Resampling
PASTIS* [47]	Points	$D_{ij} \propto \frac{1}{F_{ij}^{\alpha}}$ where α is optimized	$b_{ij} D_{ij}^{1/\alpha} + c_{ij} \log(D_{ij})$ where b_{ij} and c_{ij} are optimized parameters	No	No	Interior point and isotonic regression algorithms	Resampling
Meluzzi and Arya [48]	Spheres	$\sum_{(i,j))} k_{ij} r_{ij}^2$ where k_{ij} are adjusted such that the contact probabilities computed on the models match the F_{ij}		No	U_{excl} is a pure repulsive LJ potential. U_{bond} and U_{bend} have harmonic forms	Brownian dynamics	Resampling
AutoChrom3D* [44]	Points	$D_{ij} \propto \begin{cases} \alpha F_{ij} + \beta & \text{if } F_{min} < F_{ij} < F_{\gamma} \\ \alpha' F_{ij} + \beta' & \text{if } F_{\gamma} < F_{ij} < F_{max} \end{cases}$ where F_{min} (F_{max}) are the min(max) of F_{ij}. The parameters (α, β), (α', β') and F_{γ} are found using the nuclear size, the resolution and the decay of F_{ij} with $\|i-j\|$	$\sum_{(i,j)} \frac{(r_{ij} - D_{ij})^2}{D_{ij}^2}$	Yes	N/A	Non-linear constrained	Consensus
Kalhor et al. [14]	Spheres	$D_{ij} = R_{contact}$ to enforce the pair contact, if the normalized contact frequency F_{ij} is higher than 0.25. Otherwise the contact is not enforced	$\sum_{models} \sum_{(i,j)} k_{ij} (r_{ij} - D_{ij})^2$ where k_{ij} is different for pairs of particles, on different chromosomes, on the same chromosome, or connected	Yes	U_{excl} and U_{bond} have harmonic forms	Conjugate gradients sampling with Simulated annealing scheme	Population

* These methods are publicly available.

on previous knowledge, like the tendency of a specific region (e.g., centromeres) to aggregate in the nuclear space or the tethering of specific genomic regions to nuclear or cellular landmarks [41,54].

Finally, restraints can also be based on general polymer physics (U_{Phys}), to ensure for chain connectivity (neighbor restraints), to avoid or limit overlapping of the particles (excluded volume) or to account for bending rigidity. Accordingly, physical restraints are grouped into three major categories: (i) chromatin connectivity restraints or resistance to stretching (U_{bond}), (ii) excluded volume restraints forbidding two particles to occupy the same space (U_{excl}), and (iii) bending force applied to consecutive beads (U_{bend}). The first and second parameters aim at maintaining the connectivity of the chain as well as respecting particle occupancies. The bending rigidity term (U_{bend}) is relevant at genomic separations up to 1–2 chromatin persistence lengths and should be applied only if it is coherent with the 3C-based experiment resolution. It has been proposed that the persistence length of chromatin fiber ranges from 170 to 220 nm [79] and spans about 30–35 kb [80]. Therefore, applying the bending energy contribution makes sense for models at resolution not larger than about 50 kb. For further discussion on this point we refer the reader to other articles of this special FEBS letter.

5.3. Sampling

The sampling process consists in finding one or many (ensemble) optimal solutions that minimize the previously described scoring function. For that purpose, several optimization strategies can be employed. The first strategy assumes that a single *consensus* (average) model can satisfy all imposed restraints. This strategy neglects the structural variability inherent to the 3C-based experiment output. Under this category, we find modeling strategies that analytically identify a single consensus solution [37,38,44] as well as modeling strategies that require optimization or sampling [47]. Despite the conceptual simplicity and algorithmic efficiency of analytical approaches, which are computationally very fast, the likely usability and accuracy of the resulting models is dependent on the quality of the input data, and in particular, on the amount of structural variability present in the input matrix. The second type of sampling strategy implies that the input matrix has not been obtained from a single structure but rather an ensemble of conformations. To identify them, those strategies sample the conformational space thousands of times (*resampling*) starting from a different random configuration. They thus obtain an ensemble of solutions from which each model similarly satisfies the all input restraints [41,43,48,49]. Finally, the third strategy tries to account for the structural variability of the experimental data by optimizing in parallel a given number of models (*population*), each accounting for a portion of the observed interactions in the input matrix [14,40,45]. This last strategy is, in principle, closer to the experimental sampling of the 3C-based experiment. However it is also computationally very demanding and relies on the optimization of a number of parameters proportional to the number of fitted models.

6. Assessing the quality of the models

3C-based experimental data used to model chromatin are subjected to structural variability and experimental/genomic biases. Additionally, no consensus/systematic quality checks have been established for the many 3C-based flavors. Therefore, being able to assess the quality of the resulting 3D models is essential. Next, we briefly describe a series of tests that can help assessing the accuracy of the models obtained from 3C-based experiments. Those range from computational bootstrapping of the data during modeling to direct observation of the modeled chromatin fiber by orthogonal experimental approaches.

First, 3D models can be checked for consistency by varying the amount of data used during modeling. This consists in bootstrapping over the entire set of 3C-based data and building series of 3D models. The resulting models from each independent search are then checked for their consistency. Such consistency measures have been previously used in other fields of integrative modeling [81]. This can also be achieved by repeating the 3D modeling at different resolutions [47]. Alternatively, one can also cross-validate the modeling for consistency using different 3C-based datasets obtained with different restriction enzymes or from biological/technical replicas of the experiment [5,45]. Second, additional experimental data not taken into account for modeling can be used to indirectly validate the 3D conformation of the resulting models. For example, epigenetic marks have been previously used to validate the structure of the α-globin genomic domain in human [50] as well as the response of TADs to progesterone treatment [5]. It is important to stress that the biological knowledge included in the pool of restraints used for modeling (U_{Biol}, see previous section) cannot be used for validation purposes. Finally, the most convincing assessment of the accuracy of the 3D models is the direct observation of the chromatin fiber through confocal microscopy [19,22]. However, this verification is also subjected to the structural variability of the chromatin fiber and limited by its resolution.

7. Obtaining insights from the models

As briefly discussed in the introduction, statistical inference of chromatin structural features suffers from either low resolution, as in the case of direct observation by microscopy, or inconsistent models, as in the case of 3C-based experiments. Furthermore, the analysis of the 3C-based interaction matrices is usually limited to a number of features, like identifying loops/co-localization [7,27], TADs [6,22] or compartments [13]. By reconstructing a 3D object that best satisfies the input interaction matrix other chromatin features may be accessible [43]. For example, one can evaluate chromatin density/compaction (the number of nucleotides per nanometer of chromatin fiber); consistency of the generated models (how variable the obtained conformations after modeling are); accessibility of the chromatin fiber (how accessible is the model to an object of a given size), or differential analysis of two populations of models to contrast changes in conformation. Finally, 3D reconstructed models have been also previously used for de-multiplexing the input matrix into a limited number of plausible conformational states. Unfortunately, most of the proposed chromatin measures from the 3D models cannot, at the present time, be experimentally validated. Therefore, it is important to assess their consistency using additional biological data or bootstrapping the input data (see previous section).

8. Visualizing the models

3C-based techniques unveil the spatial nature of chromatin for which a limited number of available genome browsers, such as EpiGenome [82] or Juicebox [7], can assist its analysis. These browsers display matrices aligned with classic genomic tracks, which allows inferring mechanisms but requires the experienced researcher to create a mental image of the folding. However, reconstructing 3D models using 3C-based experiments result in sets of Cartesian coordinates for which none of the popular genomic browser is ready now days. These spatial datasets can be viewed using tools that have been previously developed for protein structure visualization such as Chimera [83]. More recently, the Genome3D software was developed to visualize 3D models of

genomes and genomic domains with a special attention to the multi-scale property of the data [84]. However, there is still the need for a browser that allows the user to visualize 3D models alongside genome-wide datasets. Examples of browsers from other fields clearly show the utility of integrating these disparate data types, e.g., MulteeSum for gene expression [85] or Aquaria for proteins [86]. Key to advancing the development of these types of tools is the creation of standard formats for interchange of genomic spatial datasets. These could be within existing frameworks such as the PDB-X (http://mmcif.wwpdb.org/docs/faqs/pdbx-mmcif-faq-general.html) or by establishing new criteria to respond to the particular conditions brought by broad-scale, multiple-resolution, data resolution [87]. Above all, visualization of genomic structures and mechanisms brings not only technical challenges, but also requires an evolving grammar to describe and explore our emerging understanding of genome folding.

9. Future perspective

During the recent years, a plethora of new methods have emerged that make use of Chromosome Conformation Capture experimental data to model 3D structures. Here we have briefly introduced the main concepts behind this new family of 3D reconstruction methods, which will likely rapidly evolve in the coming years. In our opinion, the 3D genomics community should take advantage of this to address the likely challenges ahead. First, experimental datasets obtained by 3C-based approaches are of increasing resolution (now a days ~1 Kb for the human genome maps [7]) and thus validating, storing and disseminating such large datasets is a challenge. Second, obtaining and integrating diverse datasets of experimental data from orthogonal approaches (*i.e.*, those beyond imaging and 3C-based experiments) is of most importance to generate more accurate 3D models of genomes and genomic domains. Third, it is also important to merge physics-based modeling with current restraint-based modeling. Forth, we will also need to properly store and disseminate the resulting 3D models from the diverse sets of data. And fifth, visualizing the stored 3D models, as well as mapping the wealth of available genomic data into them, is essential for obtaining biological insight.

Acknowledgments

We thank Michael Nilges for comments on standardization of data formats. We acknowledge support from the Spanish Ministerio de Economia y Competitividad [BFU2013-47736-P], the European Research Council [609989] and the Human Frontiers Science Program [RGP0044].

References

[1] Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C. and Snyder, M. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74.
[2] Berlivet, S., Paquette, D., Dumouchel, A., Langlais, D., Dostie, J. and Kmita, M. (2013) Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. PLoS Genet. 9, e1004018.
[3] Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., Diao, Y., Liang, J., Zhao, H., Lobanenkov, V.V., Ecker, J.R., Thomson, J.A. and Ren, B. (2015) Chromatin architecture reorganization during stem cell differentiation. Nature 518, 331–336.
[4] Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., Thurman, R.E., Cheng, Y., Gulsoy, G., Dennis, J.H., Snyder, M.P., Stamatoyannopoulos, J.A., Taylor, J., Hardison, R.C., Kahveci, T., Ren, B. and Gilbert, D.M. (2014) Topologically associating domains are stable units of replication-timing regulation. Nature 515, 402–405.
[5] Le Dily, F., Baù, D., Pohl, A., Vicent, G.P., Serra, F., Soronellas, D., Castellano, G., Wright, R.H., Ballare, C., Filion, G., Marti-Renom, M.A. and Beato, M. (2014) Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes Dev. 28, 2151–2162.
[6] Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380.
[7] Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. and Aiden, E.L. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680.
[8] Gall, J.G. (1981) Chromosome structure and the C-value paradox. J. Cell Biol. 91, 3s–14s.
[9] Chambers, E.V., Bickmore, W.A. and Semple, C.A. (2013) Divergence of mammalian higher order chromatin structure is associated with developmental loci. PLoS Comput. Biol. 9, e1003017.
[10] Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing chromosome conformation. Science 295, 1306–1311.
[11] Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat. Genet. 38, 1348–1354.
[12] Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C., Green, R.D. and Dekker, J. (2006) Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res. 16, 1299–1309.
[13] Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S. and Dekker, J. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293.
[14] Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F. and Chen, L. (2011) Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat. Biotechnol. 30, 90–98.
[15] Kolovos, P., van de Werken, H.J., Kepper, N., Zuin, J., Brouwer, R.W., Kockx, C.E., Wendt, K.S., van, I.W.F., Grosveld, F. and Knoch, T.A. (2014) Targeted Chromatin Capture (T2C): a novel high resolution high throughput method to detect genomic interactions and regulatory elements. Epigenet. Chromatin 7, 10.
[16] Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R. and Higgs, D.R. (2014) Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat. Genet. 46, 205–212.
[17] Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., Tavares-Cadete, F., Sugar, R., Javierre, B.M., Nagano, T., Katsman, Y., Sakthidevi, M., Wingett, S.W., Dimitrova, E., Dimond, A., Edelman, L.B., Elderkin, S., Tabbada, K., Darbo, E., Andrews, S., Herman, B., Higgs, A., LeProust, E., Osborne, C.S., Mitchell, J.A., Luscombe, N.M. and Fraser, P. (2015) The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. Genome Res. 25, 582–597.
[18] Bickmore, W.A. (2013) The spatial organization of the human genome. Annu. Rev. Genomics Hum. Genet..
[19] Marti-Renom, M.A. and Mirny, L.A. (2011) Bridging the resolution gap in structural modeling of 3D genome organization. PLoS Comput. Biol. 7, e1002125.
[20] Cremer, T., Kreth, G., Koester, H., Fink, R.H., Heintzmann, R., Cremer, M., Solovei, I., Zink, D. and Cremer, C. (2000) Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. Crit. Rev. Eukaryot. Gene Expr. 10, 179–212.
[21] Nora, E.P., Dekker, J. and Heard, E. (2013) Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? BioEssays 35, 818–828.
[22] Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Bluthgen, N., Dekker, J. and Heard, E. (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381–385.
[23] Duan, Z. and Blau, C.A. (2012) The genome in space and time: does form always follow function? How does the spatial and temporal organization of a eukaryotic genome reflect and influence its functions? BioEssays 34, 800–810.
[24] Yaffe, E. and Tanay, A. (2011) Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat. Genet. 43, 1059–1065.
[25] Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J. and Mirny, L.A. (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods 9, 999–1003.
[26] Hu, M., Deng, K., Qin, Z. and Liu, J.S. (2013) Understanding spatial organizations of chromosomes via statistical analysis of Hi-C data. Quant. Biol. 1, 156–174.
[27] Witten, D.M. and Noble, W.S. (2012) On the assessment of statistical significance of three-dimensional colocalization of sets of genomic elements. Nucleic Acids Res. 40, 3849–3855.
[28] Dekker, J., Marti-Renom, M.A. and Mirny, L.A. (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat. Rev. Genet. 14, 390–403.
[29] Rosa, A. and Zimmer, C. (2014) Computational models of large-scale genome architecture. Int. Rev. Cell Mol. Biol. 307, 275–349.

[30] Mirny, L.A. (2011) The fractal globule as a model of chromatin architecture in the cell. Chromosome Res. 19, 37–51.
[31] Barbieri, M., Chotalia, M., Fraser, J., Lavitas, L.M., Dostie, J., Pombo, A. and Nicodemi, M. (2013) A model of the large-scale organization of chromatin. Biochem. Soc. Trans. 41, 508–512.
[32] Jost, D., Carrivain, P., Cavalli, G. and Vaillant, C. (2014) Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. Nucleic Acids Res. 42, 9553–9561.
[33] Emanuel, M., Radja, N.H., Henriksson, A. and Schiessel, H. (2009) The physics behind the larger scale organization of DNA in eukaryotes. Phys. Biol. 6, 025008.
[34] Hahnfeldt, P., Hearst, J.E., Brenner, D.J., Sachs, R.K. and Hlatky, L.R. (1993) Polymer models for interphase chromosomes. Proc. Natl. Acad. Sci. USA 90, 7854–7858.
[35] Münkel, C. and Langowski, J. (1998) Chromosome structure predicted by a polymer model. Phys. Rev. E 57, 5888.
[36] Di Stefano, M., Rosa, A., Belcastro, V., di Bernardo, D. and Micheletti, C. (2013) Colocalization of coregulated genes: a steered molecular dynamics study of human chromosome 19. PLoS Comput. Biol. 9, e1003019.
[37] Zhang, Z., Li, G., Toh, K.C. and Sung, W.K. (2013) 3D chromosome modeling with semi-definite programming and Hi-C data. J. Comput. Biol. 20, 831–846.
[38] Lesne, A., Riposo, J., Roger, P., Cournac, A. and Mozziconacci, J. (2014) 3D genome reconstruction from chromosomal contacts. Nat. Methods 11, 1141–1143.
[39] Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A. and Fraser, P. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 502, 59–64.
[40] Giorgetti, L., Galupa, R., Nora, E.P., Piolot, T., Lam, F., Dekker, J., Tiana, G. and Heard, E. (2014) Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. Cell 157, 950–963.
[41] Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y.J., Lee, C., Shendure, J., Fields, S., Blau, C.A. and Noble, W.S. (2010) A three-dimensional model of the yeast genome. Nature 465, 363.
[42] Ay, F., Bunnik, E.M., Varoquaux, N., Vert, J.P., Noble, W.S. and Le Roch, K.G. (2014) Multiple dimensions of epigenetic gene regulation in the malaria parasite *Plasmodium falciparum*: gene regulation via histone modifications, nucleosome positioning and nuclear architecture in *P. falciparum*. BioEssays.
[43] Baù, D. and Marti-Renom, M.A. (2012) Genome structure determination via 3C-based data integration by the integrative modeling platform. Methods 58, 300–306.
[44] Peng, C., Fu, L.Y., Dong, P.F., Deng, Z.L., Li, J.X., Wang, X.T. and Zhang, H.Y. (2013) The sequencing bias relaxed characteristics of Hi-C derived data and implications for chromatin 3D modeling. Nucleic Acids Res. 41, e183.
[45] Hu, M., Deng, K., Qin, Z., Dixon, J., Selvaraj, S., Fang, J., Ren, B. and Liu, J.S. (2013) Bayesian inference of spatial organizations of chromosomes. PLoS Comput. Biol. 9, e1002893.
[46] Fraser, J., Rousseau, M., Shenker, S., Ferraiuolo, M.A., Hayashizaki, Y., Blanchette, M. and Dostie, J. (2009) Chromatin conformation signatures of cellular differentiation. Genome Biol. 10, R37.
[47] Varoquaux, N., Ay, F., Noble, W.S. and Vert, J.P. (2014) A statistical approach for inferring the 3D structure of the genome. Bioinformatics 30, i26–i33.
[48] Meluzzi, D. and Arya, G. (2013) Recovering ensembles of chromatin conformations from contact probabilities. Nucleic Acids Res. 41, 63–75.
[49] Rousseau, M., Fraser, J., Ferraiuolo, M.A., Dostie, J. and Blanchette, M. (2011) Three-dimensional modeling of chromatin structure from interaction frequency data using Markov chain Monte Carlo sampling. BMC Bioinformatics 12, 414.
[50] Baù, D., Sanyal, A., Lajoie, B.R., Capriotti, E., Byron, M., Lawrence, J.B., Dekker, J. and Marti-Renom, M.A. (2011) The three-dimensional folding of the alpha-globin gene domain reveals formation of chromatin globules. Nat. Struct. Mol. Biol. 18, 107–114.
[51] Sanyal, A., Baù, D., Marti-Renom, M.A. and Dekker, J. (2011) Chromatin globules: a common motif of higher order chromosome structure? Curr. Opin. Cell Biol. 23, 325–331.
[52] Umbarger, M.A., Toro, E., Wright, M.A., Porreca, G.J., Baù, D., Hong, S.H., Fero, M.J., Zhu, L.J., Marti-Renom, M.A., McAdams, H.H., Shapiro, L., Dekker, J. and Church, G.M. (2011) The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. Mol. Cell 44, 252–264.
[53] Tjong, H., Gong, K., Chen, L. and Alber, F. (2012) Physical tethering and volume exclusion determine higher-order genome organization in budding yeast. Genome Res. 22, 1295–1305.
[54] Ay, F., Bunnik, E.M., Varoquaux, N., Bol, S.M., Prudhomme, J., Vert, J.P., Noble, W.S. and Le Roch, K.G. (2014) Three-dimensional modeling of the *P. falciparum* genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. Genome Res. 24, 974–988.
[55] Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y. and Dekker, J. (2012) Hi-C: a comprehensive technique to capture the conformation of genomes. Methods 58, 268–276.
[56] Dekker, J. and Mirny, L. (2013) Biological techniques: chromosomes captured one by one. Nature 502, 45–46.
[57] Rapkin, L.M., Anchel, D.R., Li, R. and Bazett-Jones, D.P. (2012) A view of the chromatin landscape. Micron 43, 150–158.
[58] Flors, C. and Earnshaw, W.C. (2011) Super-resolution fluorescence microscopy as a tool to study the nanoscale organization of chromosomes. Curr. Opin. Chem. Biol. 15, 838–844.
[59] Huang, B., Babcock, H. and Zhuang, X. (2010) Breaking the diffraction barrier: super-resolution imaging of cells. Cell 143, 1047–1058.
[60] Scheffer, M.P., Eltsov, M. and Frangakis, A.S. (2011) Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei. Proc. Natl. Acad. Sci. USA 108, 16992–16997.
[61] Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.M., Zhu, P. and Li, G. (2014) Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science 344, 376–380.
[62] Gavrilov, A.A., Gushchanskaya, E.S., Strelkova, O., Zhironkina, O., Kireev, I.I., Iarovaia, O.V. and Razin, S.V. (2013) Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. Nucleic Acids Res. 41, 3563–3575.
[63] Russel, D., Lasker, K., Webb, B., Velazquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B. and Sali, A. (2012) Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. PLoS Biol. 10, e1001244.
[64] Greil, F., Moorman, C. and van Steensel, B. (2006) DamID: mapping of in vivo protein–genome interactions using tethered DNA adenine methyltransferase. Methods Enzymol. 410, 342–359.
[65] Marko, J.F. (2008) Micromechanical studies of mitotic chromosomes. Chromosome Res. 16, 469–497.
[66] Lajoie, B.R., Dekker, J. and Kaplan, N. (2015) The Hitchhiker's guide to Hi-C analysis: practical guidelines. Methods 72, 65–75.
[67] Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A. and Dekker, J. (2013) Organization of the mitotic chromosome. Science 342, 948–953.
[68] Trussart, M., Serra, F., Baù, D., Junier, I., Serrano, L. and Marti-Renom, M.A. (2015) Assessing the limits of restraint-based 3D modeling of genomes and genomic domains. Nucleic Acids Res..
[69] Havel, T.F. and Snow, M.E. (1991) A new method for building protein conformations from sequence alignments with homologues of known structure. J. Mol. Biol. 217, 1–7.
[70] Havel, T.F. and Wuthrich, K. (1985) An evaluation of the combined use of nuclear magnetic resonance and distance geometry for the determination of protein conformations in solution. J. Mol. Biol. 182, 281–294.
[71] Langowski, J. and Heermann, D.W. (2007) Computational modeling of the chromatin fiber. Semin. Cell Dev. Biol. 18, 659–667.
[72] Baù, D. and Marti-Renom, M.A. (2011) Structure determination of genomic domains by satisfaction of spatial restraints. Chromosome Res. 19, 25–35.
[73] Grosberg, A., Nechaev, S.K. and Shakhnovich, E.I. (1988) The role of topological limitations in the kinetics of homopolymer collapse and self-assembly of biopolymers. Biofizika 33, 247–253.
[74] Chen, P., Zhu, P. and Li, G. (2014) New insights into the helical structure of 30-nm chromatin fibers. Protein Cell 5, 489–491.
[75] Bassett, A., Cooper, S., Wu, C. and Travers, A. (2009) The folding and unfolding of eukaryotic chromatin. Curr. Opin. Genet. Dev. 19, 159–165.
[76] Gerchman, S.E. and Ramakrishnan, V. (1987) Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. Proc. Natl. Acad. Sci. USA 84, 7802–7806.
[77] Ghirlando, R. and Felsenfeld, G. (2008) Hydrodynamic studies on defined heterochromatin fragments support a 30-nm fiber having six nucleosomes per turn. J. Mol. Biol. 376, 1417–1425.
[78] Robinson, P.J., Fairall, L., Huynh, V.A. and Rhodes, D. (2006) EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proc. Natl. Acad. Sci. USA 103, 6506–6511.
[79] Bystricky, K., Heun, P., Gehlen, L., Langowski, J. and Gasser, S.M. (2004) Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. Proc. Natl. Acad. Sci. USA 101, 16495–16500.
[80] Felsenfeld, G. and Groudine, M. (2003) Controlling the double helix. Nature 421, 448–453.
[81] Alber, F., Kim, M.F. and Sali, A. (2005) Structural characterization of assemblies from overall shape and subcomplex compositions. Structure 13, 435–445.
[82] Zhou, X., Lowdon, R.F., Li, D., Lawson, H.A., Madden, P.A., Costello, J.F. and Wang, T. (2013) Exploring long-range genome interactions using the WashU Epigenome Browser. Nat. Methods 10, 375–376.
[83] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera – a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
[84] Asbury, T.M., Mitman, M., Tang, J. and Zheng, W.J. (2010) Genome3D: a viewer-model framework for integrating and visualizing multi-scale epigenomic information within a three-dimensional genome. BMC Bioinformatics 11, 444.
[85] Meyer, M., Munzner, T., DePace, A. and Pfister, H. (2010) MulteeSum: a tool for comparative spatial and temporal gene expression data. IEEE Trans. Visual Comput. Graphics 16, 908–917.
[86] O'Donoghue, S.I., Sabir, K.S., Kalemanov, M., Stolte, C., Wellmann, B., Ho, V., Roos, M., Perdigao, N., Buske, F.A., Heinrich, J., Rost, B. and Schafferhans, A. (2015) Aquaria: simplifying discovery and insight from protein structures. Nat. Methods 12, 98–99.

[87] Andrej Sali, Helen M. Berman, Torsten Schwede, Jill Trewhella, Gerard Kleywegt, Stephen K. Burley, John Markley, Haruki Nakamura, Paul Adams, Alexandre M.J.J. Bonvin, Wah Chiu, Matteo Dal Peraro, Frank Di Maio, Thomas E. Ferrin, Kay Grünewald, Aleksandras Gutmanas, Richard Henderson, Gerhard Hummer, Kenji Iwasaki, Graham Johnson, Catherine L. Lawson, Jens Meiler, Marc A. Marti-Renom, G.T. Montelione, Michael Nilges, Ruth Nussinov, Ardan Patwardhan, Juri Rappsilber, Randy J. Read, Helen Saibil, Gunnar F. Schröder, Charles Schwieters, Claus A.M. Seidel, Dmitri Svergun, Maya Topf, Eldon L. Ulrich, S. Velankar, J.D. Westbrook, Outcome of the first wwPDB hybrid/integrative methods task force workshop, Structure (2015) (accepted for publication).