The physics of three dimensional chromosome and protein organisation within the cell

Berichterstattung

Projektinformationen

THREEDCELLPHYSICS

ID Finanzhilfevereinbarung: 648050

Projektwebsite

DOI
10.3030/648050

Projekt abgeschlossen

EK-Unterschriftsdatum
16 Juni 2015

Startdatum
1 Juli 2015

Enddatum
30 Juni 2020

Finanziert unter
EXCELLENT SCIENCE - European Research Council (ERC)

Gesamtkosten
€ 1 499 236,00

EU-Beitrag
€ 1 499 236,00

Koordiniert durch
THE UNIVERSITY OF EDINBURGH
United Kingdom

Dieses Projekt findet Erwähnung in ...

Periodic Reporting for period 4 - THREEDCELLPHYSICS (The physics of three dimensional chromosome and protein organisation within the cell)

Berichtszeitraum: 2020-01-01 bis 2020-06-30

Zusammenfassung vom Kontext und den Gesamtzielen des Projekts

"The 3D structure and organisation of chromosomes within mammalian cells are very important to determine gene regulation, and therefore they are crucial for cell development, and play important roles in aging and disease as well. For instance, a liver and a brain cell of our body share the same genetic material, yet the pattern of genes which they express is vastly different, as befits their completely different functionality and cellular environment. The way genes are switched on and off relies almost invariably on 3D structure: for instance, genes which are no longer needed after a cell has differentiated are turned off by methylation of the histone proteins which are wrapped around our DNA, and this triggers the formation of a compact 3D structure which cannot be easily scanned by a polymerase (the enzymes which allows DNA transcription, the first step towards protein production). Similarly, in a senescent cells (the microscopic hallmark of aging) the pattern of open and folded chromosome regions varies substantially, and microscopy shows that protein clusters arise concomitantly with particularly compact volumes (these are called ""senescent associated foci""). Finally, diseases linked to abnormalities in 3D organisation of the chromosomes are frequent: examples are cancers and other genetic and/or developmental disorders, such as the Cornelia de Lange syndrome which is associated with loss of cohesins and CTCF, proteins involved in the formation of chromosome loops, or progeria.
Therefore, there is a high potential dividend in understanding the way genomes are organised in 3D.

Very recently, there has been a dramatic rise in experimental studies of 3D DNA, chromosome and protein organisation. It is at least as important to parallel these with computational and theoretical work. Such work is needed for many reasons: (i) to interpret and analyse the experiments, (ii) to find the fundamental mechanisms underlying this organisation, and (iii) to design new experiments to test the predictions of theories such as ours. In this context, the primary goal of our work is to provide a

major theoretical/modelling advance in our understanding of 3D chromosome/protein organisation. In our project, the modelling is done either in direct collaboration with experimentalists, or as a stand-alone study, to provide testable predictions which stimulate further work in the future. A unique feature of our work is that the scale of the simulations which we provide is much larger than previously achieved, as we rely on large computational resources which we have available in Edinburgh. A second unique aspect is that we simultaneously combine such large scale simulations both with large scale bioinformatic analyses, and with more fundamental theories involving concepts from various areas of physics, mostly statistical and soft matter physics.

More specifically, in this project we have the following main goals: (i) to provide a model to simulate the organisation of chromosomes in eukaryotic cells; (ii) to provide a model to understand DNA organisation in bacteria; (iii) to model realistic protein dynamics and structure formation, in both eukaryotes and bacteria."

Arbeit, die ab Beginn des Projekts bis zum Ende des durch den Bericht erfassten Berichtszeitraums geleistet wurde, und die wichtigsten bis dahin erzielten Ergebnisse

"Overall the THREEDCELLPHYSICS project has been successful and we have obtained results in all the areas/planned projects, with publications in all projects 1a,b and 2a,b. Following is the discussion of what I feel have been the main results, as well as their exploitation/use.

1. Development, validation and application of a 3D ""transcription factor"" model for chromosome organisation.
By using molecular dynamics simulations coupled to bioinformatic data, we have proposed a model to study the organisation of mammalian (and in general eukaryotic) chromosomes at different scale. In particular, we have studied the organisation of a number of small loci in mammalian genomes, such as the globin loci within the mouse and the SAMD4A locus within the human genome. We have then upgraded the model to study chromosome conformations in a locus important for cellular development, that of Pax6 in mouse. The latest version of the model, dubbed ""Hip-Hop"", considers chromatin as a heteromorphic polymer with variable compaction along the fibre, and this ingredient is crucial to understand the distinct ways in which the various enhancers of Pax6 work. [Hip-Hop stands for highly predictive heteromorphic polymer model.]
The transcription factor and Hip-Hop models rely on readily available bioinformatic data (DNAse hypersensitive sites and a small number of histone modifications) which are available for a large number of organisms, and cell types. Importantly, there is no fitting in our model: we can predict the 3D structure of a region of interest simply starting from these datasets; in all cases we considered the predictions conform well to experimental data of our collaborators on the papers (which employ variants of the ""chromosome conformation capture"" technique). This approach is therefore truly predictive, and we have been contacted by several groups to help them model the 3D organisation of regions that they have been interested in. For instance, Brackley is now collaborating with Dr. Daniel Rico and others in Newcastle to study the organisation of chromatin following chromosomal rearrangement in human cancer. The transcription factor model is also applicable at a much larger

scale, for instance to whole chromosomes, as we have done in a work where we focussed on dissecting the polymer physics mechanisms behind the formation of topological domains in human chromosomes (published in Nucleic Acids Research). This work constituted the first example of a fitting-free model being used to simulate domain formation in a whole chromosome.

2. Proposal of a new model for epigenetic dynamics coupled to chromosome folding.
In the work under heading 1., we used epigenetic data (e.g. histone modifications, DNase hypersensitivity) which was available for the cell line of interest. In other words, we considered a fixed epigenetic landscape. Some important questions, however, are how this epigenetic landscape is set up in the first place, when a cell differentiates, and how it can be re-established after each round of cell division. To address this new question, we have proposed a model where the 1D epigenetic information (along the chromosome) is dynamic, and is coupled to the 3D chromosome folding dynamics. Our model was the first to address the direct coupling between 1D epigenetic dynamics and 3D chromosome dynamics, and we have shown in that paper that this coupling provides a generic route towards a discontinuous transition between an epigenetically disordered and open chromosome to a folded and more compact chromosome where one of the epigenetic state dominates. Because the transition is discontinuous, there is hysteresis, equivalently memory, hence our model provides a novel fundamental mechanism to explain how cells may be able to ""remember"" their epigenetic state from one generation to the next. This biophysical model has been extended to include the presence of genomic bookmarks, which we have shown provide an avenue to the formation of multiple and coexisting epigenetic domains, as opposed to a single active or inactive domain which is the only possible stable steady states of the simpler model (lacking bookmarks) used previously. The more refined model was also applied to the epigenetic dynamics in the fruitfly genome, leading to a good agreement with the experiments.

3. Proposal of a model for supercoiling-dependent transcription.
In both bacterial and eukaryotic cells, DNA is ""supercoiled"": i.e. the DNA double helix is either underwound or overwound, and this is important to determine whether a gene is expressed (transcribed) or not, because the local DNA structure affects the likeliness with which a polymerase will bind to the promoter of a given gene. We have proposed a new 1D model which couples the dynamics of supercoiling to the stochastic dynamics of genetic transcription. This model leads to a set of surprising predictions: the coupling between supercoiling and polymerase binding naturally leads to bursty transcription, and to the upregulation of divergent genes: both facts are generically observed in transcription both in bacteria and in eukaryotes, such as yeast or the fly. We have also predicted a new phenomenon: supercoiling should trigger the setup of transcription waves in arrays of parallel genes, which can nowadays in principle be engineered in vitro or in vivo through gene editing.

4. Development of new alternative models for ""loop extrusion"".
This work provides an explanation for the bias found in high-throughput chromosome conformation capture experiments favouring looping between convergent (over divergent, or parallel) binding sites for the CTCF protein. This convergent looping bias has puzzled researchers since its discovery in 2014 (Lieberman-Aidan's lab), and the only model which could account for it was the so-called ""loop extrusion"" model, which postulated a motor activity for cohesin. Our new model, published in Physical Review Letters in 2017 as an Editor's Suggestio, argues that also a model where cohesin acts as a

molecular slip-links which slides along chromatin diffusively, with no motor activity (and compatibly with existing single molecule experiments) can still explain the convergent looping bias. This work has attracted interest in the community, and for instance we have been invited to write an Extra View article in Nucleus to highlight its findings.

5. Investigation of the relation between chromosome organisation and disease or senescence.
These works provide the first simulation study of the functional implications of genomic structure and organisation to HIV, senescence and progeria. The HIV retrovirus integrates into DNA, and it was known experimentally that it does so more efficiently when DNA is associated with histone octamers to form chromatin. Our work has discovered a simple biophysical mechanism which underlies this preference, and which depends upon the fact that DNA wrapped around histones is substantially curved: this lowers the energy barrier to integration as the cost for DNA bending is already spent. At a larger scale, our model shows that the combination between bending energy cost, DNA accessibility and nuclear organisation can explain the experimentally observed preference of HIV to integrate within active chromatin which is peripheral in the nucleus.
In another work, we showed instead that chromosome relocations in senescence and progeria create distinct changes in the chromatin contact patterns which can be observed experimentally. Existing data on senescence which were analysed prior to our work showed that contacts became more long-ranged in senescence, and our model provided the first theoretical explanation for this fact. It is notable that our model predicted that in progeria, instead, chromatin contacts should be more short-ranged, and this prediction was successfully tested on experimentally available data which were not previously used to make this conclusion."

Fortschritte, die über den aktuellen Stand der Technik hinausgehen und voraussichtliche potenzielle Auswirkungen (einschließlich der bis dato erzielten sozioökonomischen Auswirkungen und weiter gefassten gesellschaftlichen Auswirkungen des Projekts)

"1. The transcription factor and Hip-Hop models for chromatin and chromosome organisation developed in THREEDCELLPHYSICS go beyond the state of the art because they are the first fitting-free model used to model chromosome organisation within mammalian nuclei, including large scale simulations of a whole human chromosome.
This work has already had an impact within the biology community as well: following we list some activities which have been stimulated by our work.
(i) Following the work on globin we have been asked by a number of biology lab to model chromosome regions of interest in their research. These labs include A. Papantonis's lab in Goettingen, Nick Gilbert's, Tamir Chandra's and Bob Hill's labs in Edinburgh, Ajaz-ul-Wani's lab in Kashmere. The work with A. Papantonis has resulted in a paper, where our model has been compared to a new experimental technique to assess chromosome conformation using native forces, or crosslinking. The work with Gilbert has resulted in the publication of a paper in Mol. Cell describing the HiP-HoP model, and that with Chandra in the publication of a Cell Reports paper. The work with ul-Wani is ongoing and focuses on another organism (Drosophila), highlighting the generality of our

approach, which can be used genome-wide and for a variety of eukaryotes.
(ii) Following the work on the transcription factor model we have also been asked to write an extra view in Nucleus: this is essentially an auto-commentary which increases the visibility of the original paper.
(iii) Due to the interest raised by our simulation work performed overall within the THREEDCELLPHYSICS project, I have been asked to coedit with Nick Gilbert (Edinburgh) a Special Issue in Chromosome Research on Genome Organisation, and we have also written a review in Nature Methods covering mainly our transcription factor and Hip-Hop model.

Regarding wider societal impact, it is expected that the conformation of chromatin in genomic loci associated with specific diseases may provide important clues to the understanding of the microscopic and molecular basis of such pathologies, and we are now collaborating with Nick Gilbert in the Western General Hospital in Edinburgh to further pursue these possibilities in the future. This collaboration involves two PhD students (Michael Chiang and Giada Forte), funded by the Carnegie Trust and the Precision Medicine Centre for Doctoral Training (both funded independently of the current ERC grant). Brackley is also independently collaborating with the Newcastle Department of Biology to use the Hip-Hop model to predict 3D chromosomal structure in human cancers.

2. The work on epigenetic modelling goes beyond the state of the art as there was previously no model which directly addressed the coupling between 1D epigenetic dynamics and 3D chromatin/chromosome folding.
This work, published in Phys. Rev. X in December 2016, has already attracted attention among the community. It has been highlighted in Physics, via a Focus article written by Philip Ball and entitled ""How cells remember who they are"", see http://physics.aps.org/articles/v9/147.

3. The work on supercoiling-mediated transcription goes beyond the state of the art as it provides the first ever model which considers the coupling between transcription and supercoiling dynamics. The prediction of transcription waves (triggered by supercoiling) is particularly novel and we hope it will stimulate further experimental work to find these in the lab.

4. The work on diffusive loop extrusion goes beyond the state of the art as it provides a new explanation for the experimentally observed bias favouring convergent loop formation. Our model does not require any motor activity for cohesin, at variance with previous work. As there is to date no evidence of a motor activity for cohesin on chromatin, this model has attracted significant attention in the biology and biophysics community; for example our first paper on the topic was selected as an Editor's Choice in Physical Review Letters, and the journal Nucleus has asked us to write a paper (Extra View) on these results.

5. The work on nuclear body biogenesis (published in Biophysical Journal) goes beyond the state of the art as it introduces the new biophysical concept of ""protein switching"", which provides a possible explanation for the self-assembly of nuclear bodies which are simultaneously very stable and highly dynamic. It also provides a methodological advance, as it couples stochastic biochemical reactions (""switching"") to a standard Brownian dynamics code: this approach may be useful in other contexts in biophysics.

6. The work on HIV and senescence goes beyond the state of the art as it shows, for the first time, that biophysical considerations are important in these two topics, which were not previously studied by means of physical approaches. The discovery that genome bending and accessibility is key to understand the biases in HIV integration has been recognised by the biological community, for instance I have been invited to a 1-day workshop in Leuven on HIV integration in July 2018 following the work we had done, which was then available as a preprint on the biorXiv."

Example of some of the research done in THREEDCELLPHYSICS at different lengthscales.

Letzte Aktualisierung: 17 März 2021

Permalink: https://cordis.europa.eu/project/id/648050/reporting/de