THE DEPARTMENT OF MICROBIOLOGY, TUMOR AND CELL BIOLOGY
Karolinska Institutet, Stockholm, Sweden

PROXIMITY IN CHROMATIN: OPPORTUNITIES FOR INNOVATIONS

Samer Yammine

Karolinska Institutet

Stockholm 2014

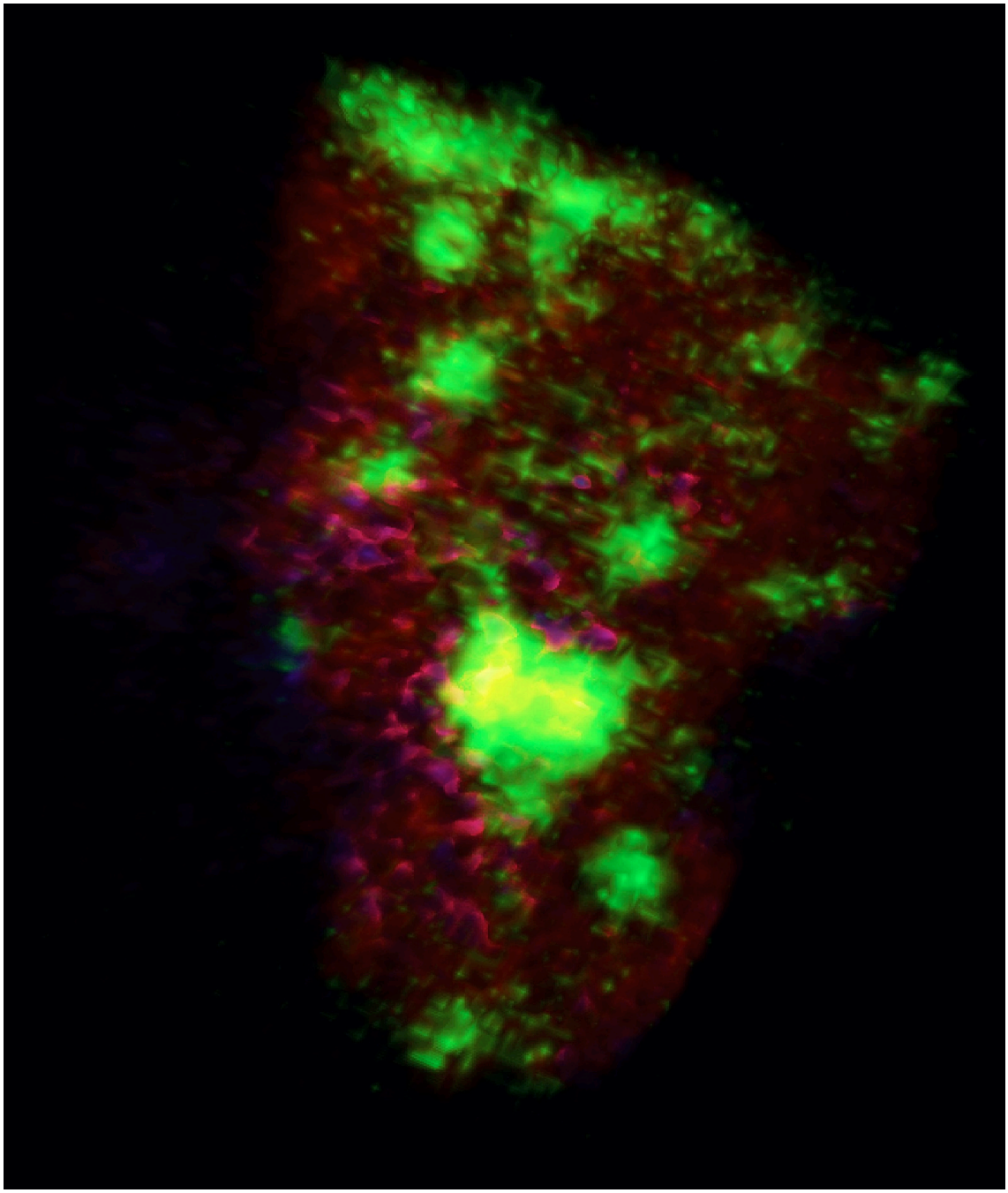


Illustration of ChrISP chromosome 11 territory.
Cover by Samer Yammine.
All previous published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet. Printed by Laserics Digital Print AB.

ISBN 978-91-7549-474-6

To the memory of Victoria & Nicolas Karam

“Stand up for what you believe in even if it means standing alone”

ABSTRACT

Mammalian chromosomes extensively communicate with each other via long-range chromatin interactions. These interactions are mostly mediated by proteins, which work as teams to control genes in the cells. These interactions could also help to unravel the mechanisms of diseases such as cancer, from new perspectives. The packaging of the chromatin fiber and how it relates to epigenetic marks that regulate its accessibility to govern lineage-specific gene expression repertoires is currently the focus of immense efforts worldwide. Moreover, how chromosomes are hierarchically folded and how they relate to each other as well as to structural hallmarks of the nucleus is a largely unchartered territory in large cell populations not to mention in individual cells.

This thesis has an emphasis on the analysis of pivotal chromatin features of single cells. Thus, interactions between a genome organizer termed CTCF and a factor involved in DNA repair, PARP1, could be demonstrated using the ISPLA technique. Such interactions likely underlie the formation of chromatin networks. Next, novel strategies/techniques were developed to visualize chromosomal structures and 3D networks by scoring for chromatin proximities within individual cells. One strategy included a novel method termed Chromatin *In Situ* Proximity (ChrISP) to visualize and identify proximities between chromatin fibers and other structural hallmarks in single cells at a resolution < 170 Å beyond that of the light microscope. Thus, large-scale changes in conformations of a single human chromosome upon the administration of reprogramming cues could be visualized. Finally, this innovation was further developed to explore differences in proximities of chromatin fibers that organize chromosome territories. The novel design, termed “rainbow ChrISP” translates physical distances in 3D, between chromatin fibres into different colors visualized with conventional microscope. This technique produced new insights into chromosome conformations and their regulation to enhance our understanding of their governing principles in single cells during development and disease.

LIST OF PUBLICATIONS

This thesis is based on the following original papers:

I. Dawn Farrar, Sushma Rai, Igor Chernukhin, Maja Jagodic, Yoko Ito, **Samer Yammine**, Rolf Ohlsson, Adele Murrell, and Elena Klenova (2010). Mutational Analysis of the Poly(ADP-Ribosyl)ation Sites of the Transcription Factor CTCF Provides an Insight into the Mechanism of Its Regulation by Poly(ADP-Ribosyl)ation. Molecular and Cellular Biology, 30(5), 1199-1216.

II. Xingqi Chen*, Chengxi Shi*, **Samer Yammine***, Anita Göndör, Daniel Rönnlund , Alejandro Fernandez-Woodbridge , Noriyuki Sumida , Jerker Widengren and Rolf Ohlsson. (2014). Chromatin *In Situ* Proximity (ChrISP): Two-parameter, single-cell analysis of chromatin proximities at a high resolution. BioTechniques, *in press*

III. Xingqi Chen, **Samer Yammine**, Chengxi Shi, Anita Göndör, Mariliis Tark-Dame, Noriyuki Sumida and Rolf Ohlsson. The H3K9me2/3 mark provides spatial identity and protects against large scale reorganization of a human chromosome. *Manuscript*

IV. **Samer Yammine** and Rolf Ohlsson. Rainbow Chromatin I*n Situ* Proximity analysis of intra-chromosomal topologies in single interphase cells. *Manuscript*

* Shared first authors by alphabetical order

Related Papers not included in this thesis:

I. Anita Göndör, Alejandro Fernandez-Woodbridge, Noriyuki Sumida, Xingqi Chen, **Samer Yammine**, Moumita Biswas, Olga Loseva, Chengxi Shi, Gyorgy Stuber, Balazs Nemeti, Maria Israelsson, Thomas Helleday, Marta P Imreh and Rolf Ohlsson. TGFβ regulates physical interactions between chromosomes mediated poly(ADP-ribose). *Manuscript*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

3C	Chromosome Conformation Capture
4C	Circular Chromosome Conformation Capture
5C	Chromosome Conformation Capture Carbon Copy
BSA	Bovine Serum Albumin
ChIA-PET	Chromatin Interaction Analysis by Paired-End TAG
ChIP-seq	Chromatin Immunoprecipitation with DNA sequencing
ChrISP	Chromatin In Situ Proximity
CT	Chromosome Territory
DMR	Differentially Methylated Region
DNA	Deoxyribonucleic acid
DSB	Double Strand Break
EB	Embryoid body cells
ES	Embryonic Stem cells
FISH	Fluorescence In Situ Hybridization
FRET	Fluorescent/Förster Resonance Energy Transfer
GFP	Green Fluorescent Protein
ICC	Immunocytochemistry
ICR	Imprinting Control Region
IF	Immunofluorescence
IgG	Immunoglobulin G
IHC	Immunohistochemistry
ISH	In Situ Hybridization
ISPLA	In Situ Proximity Assay
LADs	Lamina-associated Domains
LOCKs	Large Organized Chromatin K9 modification
NL	Nuclear Lamina
PAR	Poly(ADP-ribose)
PARG	Poly(ADP-ribose) Glycohydrolase
PARP	Poly(ADP-ribose) Polymerase
PCR	Polymerase Chain Reaction
PDI	Protein DNA Interactions
PLA	Proximity Ligation Assay
RCA	Rolling Circle Amplification
RChrISP	Rainbow Chromatin In Situ Proximity
RNA	Ribonucleic acid
STED	Stimulated emission depletion
STORM	Stochastic Optical Reconstruction Microscopy
WB	Western Blotting

1 INTRODUCTION

Single Cell Studies, The New Frontier

Cellular heterogeneity arising from stochastic gene expression events in a cell population masks the individuality of single cell events [1]. Bulk measurements and data produced from large cell population often make it difficult to fully grasp the significance of cellular specificity and its relationship to the microenvironment as well its co-existence within the same cell or sub-populations of cells [1]. Averaging collected data largely fails to reflect the real representation of cell-to-cell variations. Even in apparently homogenous tissues, there is considerable cellular heterogeneity [2]. This is exemplified by niches of stem cells that can be found within differentiated cell mass of a particular tissue. Moreover, it is increasingly being realized that tumor cell populations can be very heterogeneous to contribute to cancer progression and metastasis. The focus of such complex (multi) cellular processes lie within a 6 μm round ‘’wonder’’ organelle of almost every cell.

The Nucleus, Introduced

Nucleus, the functional center of the cell, was first seen in 1710 and subsequently became the center of interest for cell theory [3]. A major step in our understanding of the nucleus required the invention of dyes that fluoresce when bound to DNA. Thenceforth, an enormous body of cell biological and biochemical experiments were performed to discover and analyze the numerous entities within the walls of this complex yet mysterious organelle [3]. The desire to reveal the different parts of the nucleus led to the major conceptual discovery, the genome.

Genomes and Proteomes

The decrypting of the genomic blueprint started with the completion of the Human Genome Project (HGP) in 2003. Scientists around the world engaged in deciphering the high quality sequences produced from the human genome were able to identify the vast majority of genes concealed within and to provide information regarding their function and structure. The relationship between genotype and phenotype was the center of the study carried by the 1000 human genome-sequencing projects in 2010. Interpreting such data in a comprehensive way to gain insights in the genetic etiology of major human diseases turned out to be the next pivotal scientific focus. Less than 2% of the genome contains coding sequences that can be transcribed and then translated to proteins [4]. More than a chicken, fewer than a grape! Around 22000 is the estimated number of genes that most databases agree upon (RefSeq, Gencode) [5]. Such discrepancy is simply due to the fact that human genes only constitute around 1% of the 3 billion A’s, C’s, G’s, T’s, the alphabets of the genomic book. Exons contain the sequence responsible for protein production, which are punctuated by the non- coding

intronic regions, allowing the cells to blend different combinations of coding exons to produce all types of proteins essential for cellular homeostasis.

The proteome is a combination of two cellular entities, proteins and the genome, which directly or indirectly regulate the expression of all proteins at any given time or circumstances [6]. Importantly, the proteome has a more dynamic composition, which differs greatly among different cellular type and function, while the genome is almost completely static. Both perspectives involve several levels of complexity; the code of genomic sequences is superimposed by the epigenetic "code" that regulates accessibility to transcription factors among other things [6]. The proteome, on the other hand, adds more intricacy by providing enormous complexity in potentials for protein-protein interactions that can be further regulated by post-translational modifications. Moreover, the proteins encoded by the genome can be diversified by alternative splicing to impact protein structures and functional protein-protein interactions.

Protein-Protein Interactions

To implement biological functions, two or more proteins frequently form complexes. Such interactions play a pivotal role in many biological processes, such as replicating DNA, import/export, catalyzing metabolic reactions, responding to external stimuli, DNA repair, maintaining cellular structures, and so forth… Proteins may bind directly or indirectly to DNA. In the latter instance, exemplified by chromatin insulators, protein-protein interactions are essential to effectuate the expressivity of the genome. Thus, insulators influence the ability of an enhancer and a promoter to communicate with each other *via* protein-protein interactions to differentially regulate expression within clusters of genes, for example.
CTCF (CCCTC-binding factor) is one of the best-known insulator proteins, which participates in creating a three dimensional (3D) DNA structure capable of inhibiting transcription patterns depending on the position of its binding sites [7]. Through the binding of its 11 zinc fingers to CTCF binding sites on chromatin or by interacting with various proteins, CTCF can however also promote transcription as well as inhibit elongation of transcription to modulate splicing patterns [8]. The biological diversity of CTCF's function likely arises from its association with its partners in crime, such as PARP-1, cohesin, polII and nucleophosmin, in context-dependent manner [7]. The ability of CTCF to simultaneously bind a specific DNA sequence along with partner proteins is likely a key to the formation of chromatin loops and bridges and linking of intra- and interchromosomal hubs [9].
One such partner is PARP1, which covalently modifies proteins by a sequential transfer of ADP-ribose from the coenzyme NAD^+ resulting in variable chain of PAR ranging from a couple to several hundred units, forming branched structures on targeted proteins [10]. These modifications are then rapidly degraded by poly- (ADP-ribose) glycohydrolase (PARG). The simultaneous presence of CTCF and PARP-1 at the imprinting control region (ICR) and their association at other genomic loci may form the basis of chromatin hubs that can respond to environmental cues. It is currently not clear if the PARylated CTCF is the cause or consequence of such chromatin hubs [11].

Chromatin Marks and histone modification

The packaging of genomic DNA around the highly conserved core histones (histone H2A, H2B, H3 and H4) forms a nucleosome that is the smallest unit of chromatin [12]. The covalent modifications of histone tails, such as acetylation, ubiquitination, sumoylation, methylation, phosphorylation and ADP ribosylation, provide a signature that can be recognized by specific factors, such as TFIIB, to influence transcription or repression (Fig.1). For instance, methylation of lysine residues on the N-terminal of a histone tail renders the formation of specific chromatin domains, such as facultative and constitutive heterochromatin or euchromatin [12]. Moreover, additional histone modifications can trigger regions of chromatin to undergo nuclear compartmentalization and adhere to the nuclear membrane, for example. Such compacted chromatin structures and repressed chromatin states are generally associated with chromatin marks, such as methylated H3 lysine 9 (H3K9me2), H3 lysine 27 (H3K27me3). Conversely, open chromatin structures and active chromatin states are commonly associated with methylation on H3 lysine 4 (H3K4me3) and H3 lysine 36 (H3K36me3). These modifications appear quite dynamic throughout the cell cycle in response to environmental stimuli. The combinatorial patterns of histone modifications thus serve as scaffolds for higher order chromatin structures to regulate contacts between histone-DNA interactions as well as between their partners.

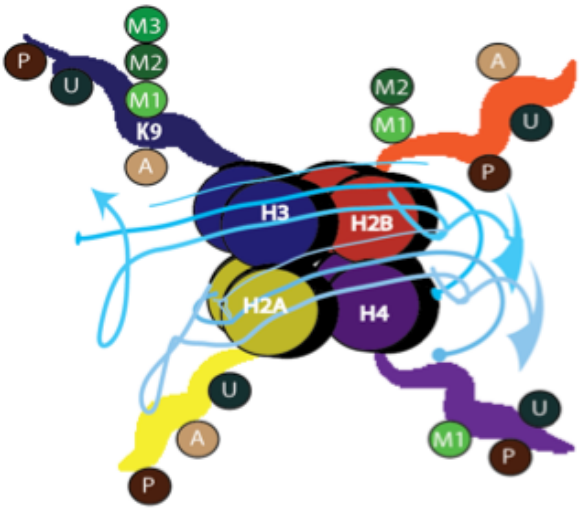


Fig.1 Nucleosome illustration

One of the prominent marks of transcriptional repression, H3K9me2, is essential for the normal phenotype as its perturbation is associated with several diseases such as prostate and gastric cancers, acute myeloid leukaemia (AML) and metabolic disorders such as diabetes [13]. Indeed, H3K9me2, which is jointly maintained by the conserved protein methyltransferases G9a/GLP, plays a central role during cellular development by regulating chromatin structure and gene expression [13]. Thus, the H3K9me2 mark, which is mainly located near genes and on chromatin blocks (up to 4.9 Mb) termed

large organized chromatin K9 modifications (LOCKs), likely plays a structural role in maintaining silenced states during lineage differentiation and disease progression [14].

Chromatin Architecture

About 2 meters long and comprising a total of 3.3 billion base pairs, DNA is meticulously packed in 23 pairs of chromosomes within a single cell. Histones provide the electrostatic attraction to tightly shape DNA; every 8 histones wrapped with 147 base pairs form a nucleosome, which together with the linker region, encompassing about 50 bp becomes the functional and structural unit of chromatin, coiled into higher-order structures [15]. In order to access this highly compacted structure during transcription, replication and repair both strands of DNA need to be available for several enzymes to act upon [16]. Two major mechanisms come to play, either enzymatic modification of histones or exposure of the underlying DNA by chromatin remodeling complexes [17]. Once these tasks have been implemented, the chromatin then reverts back to its more or less compact state.

Pretty on the inside! The 3D space of a nucleus

Chromosomes distribute within the interphase nucleus in regions termed as chromosome territories (CT) [18]. Such regions are bordered by so-called interchromatin compartments (IC) to organize specific compartments that profoundly influence transcriptional regulation [19]. Active genes or gene-rich areas dominate the interior of the nuclear space while gene-poor domains and inactive genes tend to occupy the periphery [20]. Such alignment is indeed observed when active genes decondense to loop out of its chromatin base to join the so-called transcription factories[21]. These structures form when sufficient numbers of actively transcribed genes congregate. Thus, transcriptionally active genes are most of the time located on the periphery of CT and oriented towards the center of the nucleus whereas silenced or inactive genes are usually connected with the nuclear periphery [22]. The nuclear lamina (NL), a static nuclear structure and an anchoring point for chromosomal domains, houses over 1300 identified lamin-associated domains (LADs). Such regions are characterized by low gene-expression levels and high levels of repressive chromatin marks [23]. Contrary to previous beliefs, this nuclear scaffold is dynamic in nature and provides a plasticity to higher order chromatin structures. For example, chromatin of specialized cells like rod photoreceptor cells of nocturnal animals undergoes a complete spatial reorganization in response to extracellular environment [24]. To facilitate such stochastic interactions, *in vivo* molecular crowding greatly increases the effective concentration of a component by several orders of magnitude and augments the formation of protein-protein and protein-DNA interactions. These interactions are perceived to promote preferential associations among genomic regions and limit their motion. A pattern of preferential, yet probabilistic positioning thus emerges due to the self-organizing fashion of each chromosome and its encounter with its neighbors [25].

Without doubt, an important challenge in chromatin biology focuses on the deciphering of the mechanism(s) of interactions between chromosomes in three dimensions. Once these are mapped it will be essential to understand the ramifications of this interactome on the expressivity of the genome.

Technology in focus: insights into nuclear organization

Available technologies

The most widely used *in situ* method for detecting a single protein or protein complexes exploit affinity binders. *In situ* methods deliver a localized detection of the moiety in question directly within single cells or tissues. Such approach is used for many diverse purposes under many different assays, formats or designs to expose any heterogeneity amongst cells. *In situ* analysis encounters a couple of challenges when measurements are scrutinized for sensitivity and selectivity. However, it fails to selectively distinguish between the molecule of choice and the unaccountable adjoining molecules towards providing robust measurements. In contrast, *in vivo* methods deliver the most biologically relevant data since assays are performed in living cells or intact animals, however, in many instances technical challenges and ethical issues hamper their direct application.

Location, Location, Location

Immunohistochemistry (IHC) is a simple, yet routinely used method in medical diagnostics and research field. This technique was developed in the early 19th century to detect proteins, antigens or any other moiety by the principle of an antibody-antigen (Ab-Ag) reaction [26]. When the method is used to score for tissue sections while maintaining the original architecture of the surrounding tissue, it is referred as immunohistochemical staining while when it is used to detect cellular antigens in cell cultures, it is then entitled immunocytochemistry (ICC) [26].
In principle, ICC can only mark one antigen at a time. However, with the advancements made in the field, fluorescent multiplexed detection can now be achieved.

ICC can be implemented with two different approaches, either a direct or indirect detection. Direct detection is when the primary antibody used for antigen detection is directly labeled with an enzyme or fluorescent molecule while indirect detection is when two steps are added in the detection of the antigen. Primary antibody is used to detect the antigen followed by a secondary labeled antibody to detect the primary antibody to increase sensitivity [26]. When more than one affinity binder is used, the sensitivity is increased while the background and non-specificity decreases. It is not always trivial in many cases to use multiple binders due to the high risk of cross-reactivity issues rising from the non-specific binding of antibodies to each other or towards different antigens.

Proteins can also be located and visualized by integrating a detectable identifier. This is exemplified by introducing constructs containing a DNA vector encoding a detectable protein to cells. Once the fusion protein is expressed joined to the protein of interest sub-cellular localization can be studied and documented. Green fluorescent protein

(GFP) is one of the most commonly used tools to study either living or fixed transformed cells or tissues [27]. Similar to GFP, novel methodologies have been designed to introduce different types of tags to the protein of interest that can be afterwards visualized with numerous fluorescent dyes, such as SNAP and CLIP tags systems, to name a few. However, these techniques report information at the single protein level and come up short when the aim is to assay whether proteins participate in networks inside cells and/or their posttranslational modifications. To tackle these limitations other techniques are required, such as Western blotting (WB) for detection of posttranslational modification of proteins and Förster resonance energy transfer (FRET) to determine protein-protein interactions [28]. In WB analysis, the proteins are first separated on gel according to size and visualized by specific antibody detection either with an enzyme or fluorescent reporter. Concerning FRET, two fluorescent moieties tagging two potentially interacting proteins should come into proximity (10 nm), subsequently have the correct orientation, so that once excited with a light source, the "acceptor" can absorb the energy released from the "donor" and emit a longer wavelength indicating the interaction or the closeness of the two proteins. Both WB and FRET have advantages; WB for instance, allows the separation of proteins by size, and conformation, thus allowing the detection of several targets, whereas, in FRET the protein of interest is exclusively tagged and the expressed fusion protein can be tracked in living cells with a simple microscope, eliminating the artifacts introduced due to cell fixation and permeabilization. On the contrary, WB analysis requires high amount of cell lysate, tedious optimizations and it is prone to false or subjective results considering the cells' mixed population as well as the antibody specificity. FRET in turns is simply not suitable for clinical samples or cancer tissues due to the technical requirements. In addition, fusion proteins might interfere with the endogenous protein levels and properties leading to a change in their subcellular localization and patterns of interactions [28].

Propinquity – An overview

As described earlier, affinity binders are commonly used to detect proteins of interest. Several binders have been exploited, each under specific, well-tailored experimental conditions. Aptamers [29] (DNA / RNA), affibodies [30] and antibodies were the golden choice of many technical innovations by virtue of their high affinity to the target of interest as well as the possibility to label them with countless modifications (enzyme, DNA, RNA, fluorophores). Antibodies, selected on the basis of their ability to bind their antigen from an infinite repertoire of unique binders are produced in immunized animals. Out of the several existing subclasses, IgG is mainly used, either in monoclonal or polyclonal forms.

As discussed previously, sensitivity is one of the major drawbacks which scientists try to improve constantly either *via* signal amplification or by background elimination. Immuno-PCR is one of such progress, where an oligonucleotide probe is covalently attached to the antibody and upon target detection a circular DNA probe is added [31]. Similar to target based amplification known as PCR, signal-based amplification strategy is used in this case. Rolling Circle Amplification (RCA) is initiated from the primed oligonucleotide on the antibody by the Ø29 enzyme, sequentially amplifying

the circlular DNA oligo in an exponential mode. The antibody-tethered end product of this amplification is then subjected to hybridization with fluorescently labeled short oligonucleotides detected as micrometer-sized spheroids under a fluorescent microscope [31]. One major drawback arises after high sensitivity, that of signal non-specificity due to binder cross-reactivity. This is a common flaw encountered with almost all methods relying solely on the affinity of one binder, which in several instances binds to other proteins with analogous epitopes, misses its target or cross-reacts with other moieties.

To overcome this additional pitfall, an innovative and ingenious concept was conceived. Proximity Ligation Assay (PLA) [32], utilizing two or more binders, simultaneously targets the protein of interest for greater enhanced selectivity (Fig.2). Briefly, two or more binders with covalently linked oligonucleotides are used, once those binders recognize their target, the excess is then removed. Two additional oligonucleotides are then introduced, one to join both oligonucleotides on the binders in a linear manner and the other to complete the formation of a circular product. A complete circle is then end-joined by a DNA ligase followed by an amplification step initiated by the DNA polymerase Ø29 from the priming end of one of the binder oligonucleotides. Concatenated copies of the DNA circle (~ 1000 folds) replicated via RCA can then be visualized upon hybridization of fluorescently labeled sequence specific oligonucleotides. Thus, this technique resolves both selectivity and sensitivity issues [33] and its applications are simply limited by the imagination of the user. I will touch upon several innovations conceived from this concept in the following sections.

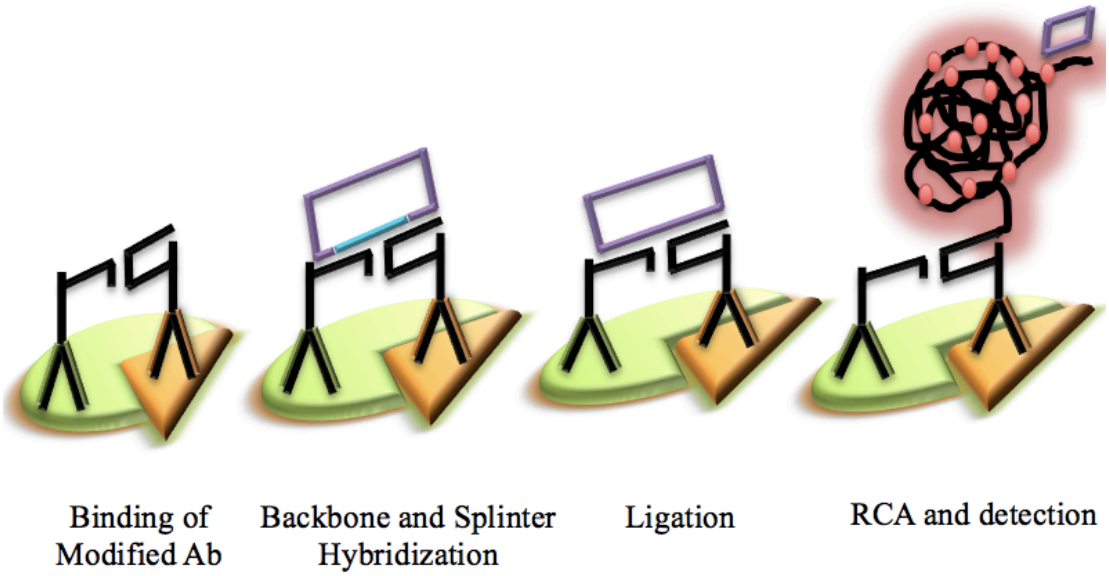
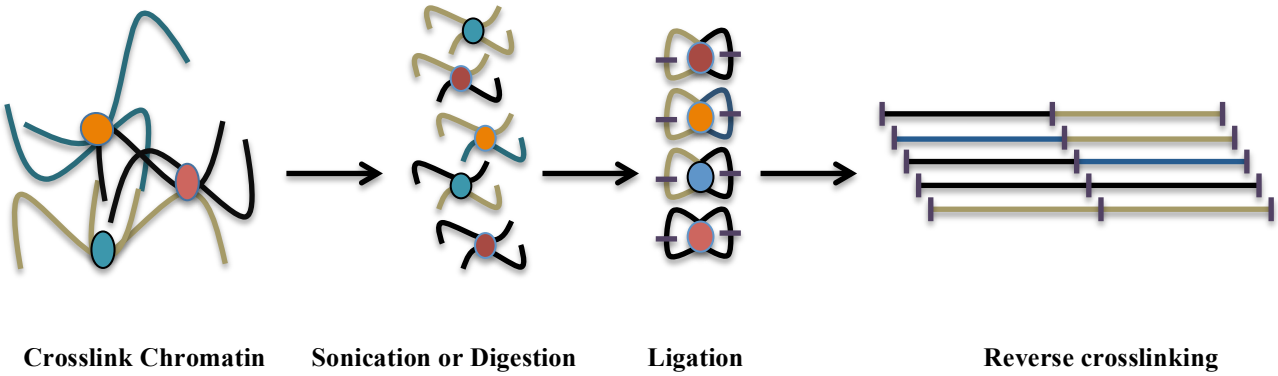


FIG.2 *in situ* PLA (ISPLA) a step-by-step approach.

Three-Dimensional (3D) genomics network

The development of the so-called "C" techniques has enabled the visualization of the spatial organization of chromosomal regions at a resolution level not possible just a few years ago [34]. Chromosome Conformation Capture (3C), used to analyse long-range interactions within chromosomes and its derivatives such as 4C/5C/Hi-C methodologies are established on the *in situ* ligation of proximal cross-linked chromatin structures to measure chromatin fiber proximities in high throughput manners [34]. Table.1 provides brief overview regarding the major similarities and differences between these techniques. While most of these techniques consume massive amount of cells, they all fail to generate more than a hand full number of interactions per cell while the estimated number of interactions in a single cell is in millions and fails to detect cell-to-cell variation and dynamics. In addition, these methods do not provide any information about the possible proteins involved in those three-dimensional structures.



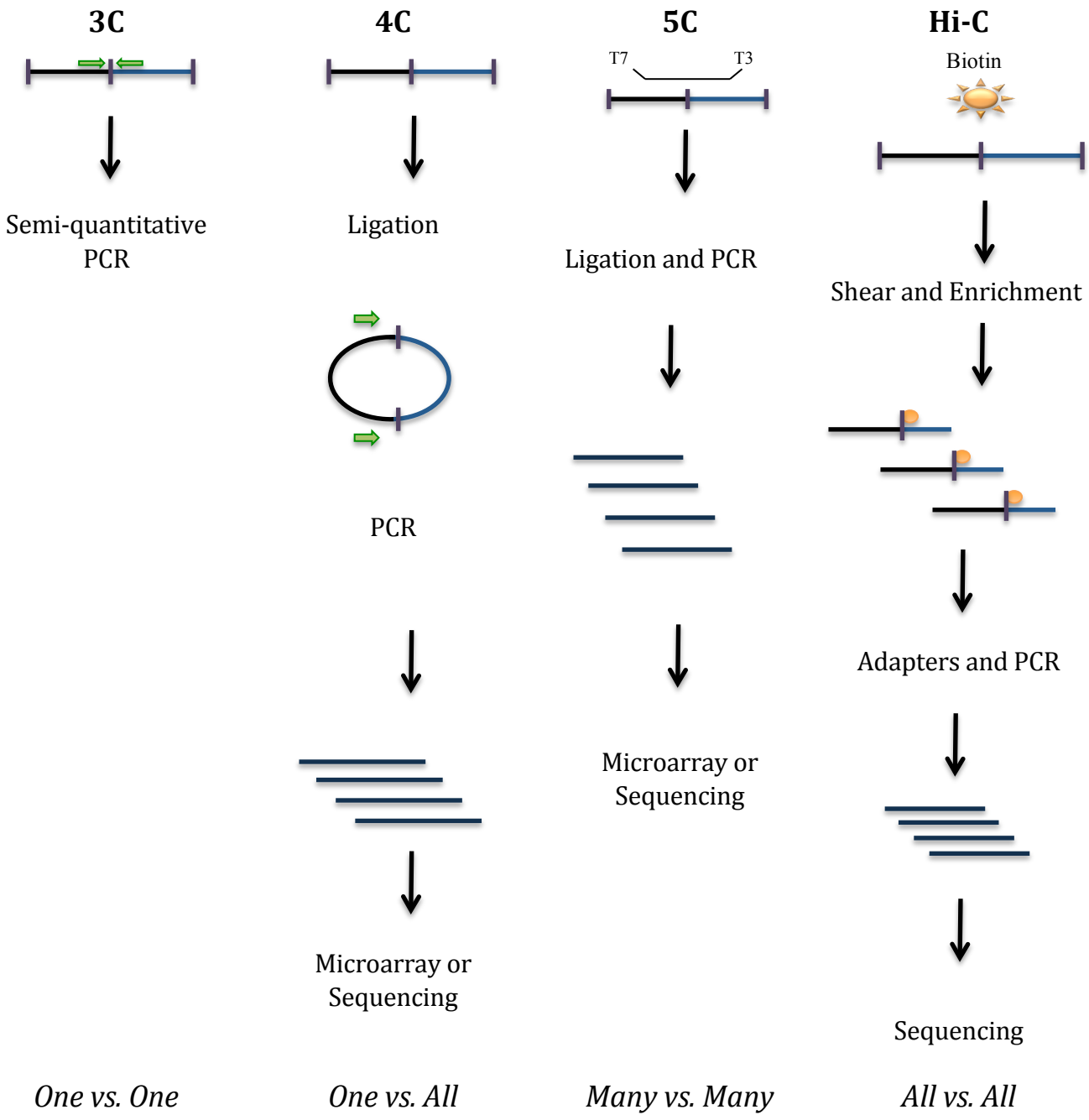


Table 1. Overview of the "C" techniques.

To fully understand the dynamics behind nuclear architecture, the visualization of chromatin (DNA + Protein) in a super resolution manner is required. The core technique used for spatial exploration of particular DNA targets ranging from complete CTs to distinctive gene loci, is fluorescence *in situ* hybridization (FISH), which involves hybridization of fluorescently labelled complementary sequences of the researched region [35]. FISH is complemented by immunostaining to identify protein-DNA-RNA co-localization within the same cell. Unfortunately, this fusion is not always fruitful due to the fact that numerous epitopes are destroyed under the harsh FISH handling conditions. Enduring heat denaturation and permabilization procedures are deemed to be the main factors behind the survival of few epitopes and a novel approach should be designed to cope with million other protein epitopes. The dozens of FISH protocols that are available are highlighted with their pros and cons in Table. 2.

Single cell studies of chromatin proximities in DNA FISH applications are severely hampered by its low 3D resolution [35]. Although the lateral resolution is limited to ~250 nm by diffraction in the confocal microscope, this resolution can be considerably increased in the Stimulated emission depletion (STED) microscope [36]. The axial resolution of STED microscopes is, however, comparable to those of confocal microscopes making it virtually impossible to obtain high-resolution 3D topologies of chromatin structures using even STED/ Stochastic Optical Reconstruction Microscopy (STORM) microscopic applications [37].

Despite the tremendous efforts invested on improving the FISH protocols to date, application-specific fine-tuning of sensitivity, multiplicity and fundamental ways to correlate molecular data (interactions in space and time) from the “C” techniques to the behaviour of single cells remains necessary [38].

Design, Application and novelty

ACM-FISH	Multicolor/ alpha (centromere), classical (1q12), and midi (1p36.3) for simultaneous detection of numerical and structural chromosomal abnormalities in sperm cells.
ArmFISH	42-color M-FISH variant that allows the detection of chromosomal abnormalities at the resolution of chromosome arms (p- and q-arms of all 24 human chromosomes, except the p-arm of the Y and acrocentric chromosomes).
catFISH	Cellular compartment analysis of temporal (cat) activity by FISH is an ingenious experimental approach devised to investigate the dynamic interactions of neuronal populations associated with different behaviors or cognitive challenges
CO-FISH	Chromosome orientation uses single stranded DNA probes to produce strand-specific hybridization to determine the orientation of tandem repeats within centromeric regions of chromosomes.
CARD-FISH	Catalyzed reporter deposition-FISH, refers to the signal amplification obtained by peroxidase activity through the deposition of a large number of fluorescently labeled tyramine molecules in which the horseradish peroxidase (HRP)-labeled probe has bound.
CB-FISH	CB-FISH involves hybridization on binucleated cells in which cytokinesis has been blocked by treatment with cytochalasin-B (CB).
COBRA-FISH	Combined binary ratio, produce more pseudocolors, allowing the resolution of more than 24 colors within a specimen.
COD-FISH	Concomitant oncoprotein detection-FISH, a technique to visualize not only the loci signals for a particular oncogene, but also the protein product derived from this gene.
Comet-FISH	Comet-FISH is a combination of the comet assay and FISH analysis.
D-FISH	Double fusion FISH, use of two (or two sets of) differentially labeled large probes for the detection of recurring chromosomal translocations in hematological malignancies.
COMBO-FISH	Combinatorial oligonucleotide, is a method with no requirement for sample denaturation prior to hybridization.
Cryo-FISH	Cryo-FISH makes use of ultrathin cryosections (150 nm thick) of well-fixed, sucrose-embedded cells applied to the study of spatial interrelationships of chromosome territories in the cell nucleus.
DBD-FISH	DBD - DNA breakage detection FISH by means of an alkali DNA unwinding solution and protein removal.
e-FISH	e-FISH is a BLAST-based FISH simulation program able to accurately predict the outcome of hybridization experiments.
Flow-FISH	PNA-labeled telomere probes are used to visualize and measure the length of telomere repeats with flow cytometry.
Halo-FISH	Performed on cells that are first permeabilized and then extracted with high salt to remove soluble proteins.
Fiber-FISH	Allows high resolution mapping of genes and chromosomal regions on fibers of chromatin or DNA, permitting physical ordering of DNA probes down to a resolution of 1000 bp.
Fusion-Signal FISH	Devised for the identification of the 9;22 Philadelphia translocation in peripheral blood and bone marrow cells in CML patients.
Harlequin-FISH	Cell cycle-controlled chromosome analysis in human lymphocytes that allows a precise quantification of induced chromosome damage for human biodosimetry purposes.
LNA-FISH	Use of Locked nucleic acids oligonucleotides in FISH experiments for improved resolution and sensitivity.
Immuno-FISH	Permits the visualization of antigens within the sample, so that both DNA and proteins can be analyzed on the same sample.
M-FISH	Multiplex-FISH, 24-color karyotyping, based on combinatorial labeling.
ML-FISH	Multilocus FISH, simultaneous use in multicolor FISH of multiple probes to screen for multiple microdeletion syndromes in patients
PCC-FISH	Premature chromosome condensation, use of chromosome-specific painting probes to determine chromosome damage after irradiation.
QD-FISH	Quantum dot-conjugates probes for FISH analysis on human metaphase chromosomes.
PNA-FISH	Peptide nucleic acids, to measure individual telomere lengths on metaphase chromosomes.
Q-FISH	Quantitative-FISH methodology permits the measurement of probe signal intensity.
Rainbow-FISH	Allows the simultaneous detection and quantification of up to seven different microbial groups.
Raman-FISH	Combines FISH technology with Raman microspectroscopy for ecophysiological investigations.
Reverse-FISH	FISH probe comprises DNA from the material of interest for characterizing marker chromosomes and chromosome amplifications in cancer.
RNA-FISH	Allows detection of RNA within cells to analyze the transcriptional activity of endogenous genes.

ReD-FISH	Replicative detargeting, allows the replication timing of specific sequences to be determined.
RING-FISH	Use of high concentrations of polynucleotide probes for an increase in sensitivity and visualization of any part of the genetic material of a bacterial cell, regardless of copy number.
RxFISH	Color banding technique that is also described as chromosome bar-coding.
Split-Signal FISH	Fast and sensitive dual-color FISH assay for the detection of frequently occurring chromosome translocations affecting specific genes in hematopoietic malignancies.
3-D FISH	Analyze spatial positioning and relative organization of chromosomes and subchromosomal regions within cell nuclei.
T-FISH	Tyramide-FISH, ultrasensitive detection.
Zoo-FISH	Crossspecies chromosome painting FISH consists of hybridizing libraries of DNA sequences, also known as chromosome paints, from one species to the chromosomes of another species, to identify regions of synteny.
Tissue-FISH	FISH on frozen, fixed, or embedded in paraffin wax tissues samples.

Table. 2 Overview of the FISH techniques.

Technology in focus: Functional 3D regulatory DNA networks – A new era

To pave a new era of functional 3D regulatory DNA networks and their stochastic formations will require that such analyses be performed in single cells [39]. This approach must include analyses between sub-chromosomal domains within and between chromosomes, of patterns of chromatin crosstalk during development and cancer and to enable the distribution of chromatin-bound protein-protein complexes within the nuclear architecture with a special reference to chromosomal networks. Data analyzed by ChIP (protein-DNA) and 3C (DNA-DNA) methods perpetually yield a common 3D interactome overlooking the unique and inclusive spatial arrangement of chromosomal interactions (short- or long-range) in a single cell. Thus, these data sets can only be implicitly used when massive heterogeneity is considered in chromosome folding and positioning [39]. Tools to analyze these interactions from a single locus towards a single chromosome have started to emerge and a whole new era of inter-cellular-individualism has ushered in.

Initial concepts to visualize individual protein-DNA interactions (PDIs) was based on the *in situ* PLA assay that use padlock probes to detect unique DNA-Protein complex sequences in the genome [40]. Previously, PDIs where only scrutinized in bulk and extracted information was simply the averaging of the millions of stochastic interactions spotted in heterogeneous population of cells or tissues [40]. PLA generated signals from endogenous protein-DNA complexes overcome main limitations of FISH and IF to distinguish between two adjacent fluorescent variants and evades the low spatial resolution of microscopy. Similar approach combining *In Situ* Hybridization (ISH) and PLA has been used to study histone modifications at specific gene loci in histological sections [41].

Novel technologies built upon the PLA principle are thus already paving the way towards new horizons in visualizing, uncovering, analyzing of protein-DNA-protein hybrid interactions.

Soliloquy – thinking out loud

Developing novel techniques for both molecular biology and diagnostics is rather a tedious and extensive process. Numerous criteria must be attained in each single step throughout the practice. Such rational must be pertained to accomplish reproducible yet applicable concept. For developing novel methodologies, a scientist should start by thoroughly standardizing commonly used steps in routinely performed experiments. For example, fixation and sonication, both need finely optimized parameters; so that once reproduced in different facilities, they should both generate as similar as possible and trustworthy patterns and outcomes. This should be followed by a general definition for the subsequent terms, such as sensitivity, specificity, limit of detection, dynamic range and foremost of all, throughput. Without the above guidelines, technical development would suffer tremendously. The work portrayed in the succeeding section discloses these issues and the challenges tackled while building upon PLA.

2 AIMS OF THE THESIS

The overall ambition of the studies presented in this thesis was to invent new technologies that will facilitate the understanding of epigenetic processes. Specifically I have addressed the following aims in my thesis:

- To delineate protein complexes engaged in chromatin insulation in single cells

- To innovate a new technique that is able to identify higher order chromatin structures in single cells at an unrivalled resolution

- To use this technique to identify the spatial identity of interphase chromosomes and its relationship with chromatin marks and transcriptional competence

- To translate conformations within a single chromosome into a color code to uncover novel aspects of chromatin organization

3 RESULTS AND DISCUSSION

This thesis demonstrates the innovation of the Chromatin *In Situ* Proximity (ChrISP) technique to study DNA-Protein-DNA interactions in fixed cells at an unprecedented resolution. The technique, which is based on the *In Situ* Proximity Assay (ISPLA) principle, has several applications. While the first paper highlights the use of the original ISPLA method to detect endogenous protein-protein interactions under different environmental cues, the following papers visualize chromatin fibre proximities within a single chromosome in single cell using the ChrISP technique. The conclusions generated by these observations include the identification of the spatial identity of a single chromosome and how this relates to structural hallmarks within a single chromosome. Moreover, a ChrISP derivative termed rainbow ChrISP or rChrISP visualizes different levels of chromatin proximities in complex patterns to generate novel insight into interphase chromosome structures.

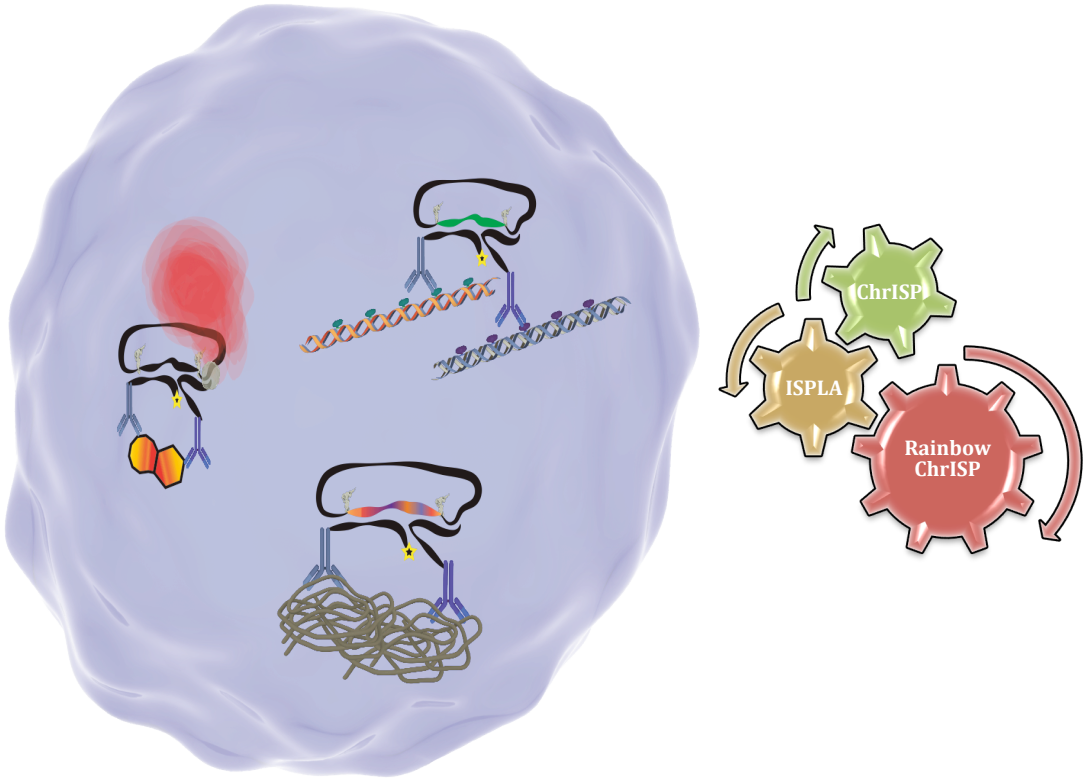


Fig. 3 Overview of technique development in present investigation

Paper I: Mutational Analysis of the Poly(ADP-ribosyl)ation Sites of the Transcription Factor CTCF Provides an Insight into the Mechanism of its Regulation by Poly(ADP-ribosyl)ation

The eleven zinc finger factor termed CTCF has been described as a master weaver of the genome due to various reasons [42]. The most important function of CTCF, is its binding to the *H19* imprinting control region (ICR). This region is responsible for the parent of origin-dependent mono-allelic expression of *Igf2* and *H19* and is inherited in two flavours: The paternal allele is methylated whereas the maternal allele is unmethylated [43]. CTCF binds to only the unmethylated *H19* ICR allele [44] to insulate the *Igf2* promoters from the downstream enhancer on the maternal allele [44]. The CTCF – dependent chromatin insulator function [45] depends on poly(ADP-ribosyl)ation (PARylation) [9]. Among all the 17 members of PAR polymerase (PARP) family, (PARP-1) is the most likely candidate catalyzing the PARylation of CTCF [46]. This conclusion is further supported by the fact that CTCF activates PARP-1, a process that has been proposed to lead to DNA hypomethylation [47]. In this study, we performed several experiments to investigate and identify PARylation sites in CTCF.

CTCF and PARP-1 colocalize in the nucleoplasm

To strengthen the link between CTCF and PARP1, it was essential to visualize co-localization of CTCF and PARP1. Using the *in situ* proximity ligation assay (ISPLA), it was indeed verified that CTCF and PARP1 are extensively co-localized in the nucleoplasm.

Generation of a mutant deficient in PARylation, which was employed to investigate the importance of CTCF PARylation in transcriptional regulation and the control of cell proliferation.

There appears to be little preferences for PARlatable amino residues although glutamic acid and lysine residues appear overrepresented [46]. It was determined that the cluster of glutamic acid residues at the N-terminal end was the prime PARylable targets within CTCF. To generate a mutant CTCF deficient in PARylation, glutamic acid clusters were mutated to an alanine. The resulting mutant CTCF was identical to the wild type protein in several aspects, such as size, localization and ability to bind DNA except that it could not be PARylated. When overexpressed, the mutant CTCF abrogated the insulator function in a manner similar to the wild type CTCF in the presence of PJ-34, a strong PARP1 inhibitor. These results implicate that PARylation of CTCF is vital for the chromatin insulator function.

CTCF forms a functional complex with PARP-1 at CTCF binding sites, independent of the PARylation state of CTCF.

An earlier key observation demonstrated that CTCF remained bound to the *H19* ICR even when extensively PARylated [9]. This is in contrast to most other factors, such as p53, that are evicted from chromatin upon PARylation in all likelihood due to the strong negative charge produced by PARylation [48] [49]. This unique feature of CTCF likely depends on the separation of the DNA binding zinc finger domain from that of

PARylated N-terminal domain. Such studies indicated that the mutated residues in the N-terminal end of CTCF did not affect its ability to bind the *H19* ICR, the mutant CTCF was transfected into MCF-7 cells. Subsequent chromatin immunoprecipitation assay revealed that the mutant CTCF was indeed able to interact with its binding sites in the living cells.

Summary and perspective

Previous observations claimed that the productive interaction between CTCF and PARP1 is involved in the chromatin insulator function of mammalian cells. The poly(ADP-ribosyl)ation (PARylation)of CTCF resulting from this interaction may facilitate/stabilize long-range chromatin fibre interactions underlying the insulator function. To be able to conclusively state that this post-translational modification of CTCF is directly involved in chromatin insulation, a mutant CTCF that could not be PARylated was generated. Using mutant CTCF, it was demonstrated that the insulator function of the *H19* ICR was indeed impeded despite that the mutant CTCF and PARP1 interacted at the ICR. Using ISPLA, we further demonstrated that CTCF and PARP1 co-localizes throughout the nucleus to make the case that such CTCF-PARP1 complexes have a genome-wide function in demarcating expression domains. This current paper has fulfilled aim number 1: **To delineate protein complexes engaged in chromatin insulation in single cells.**

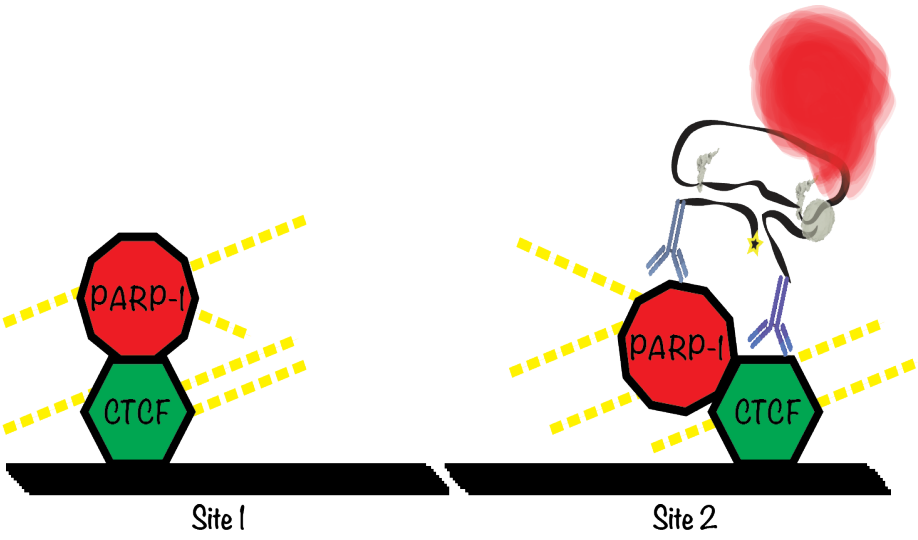


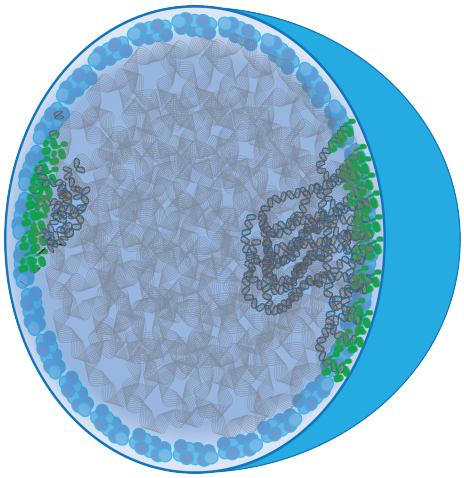
Fig. 4 CTCF and PARP-1 co-localizes at *H19* ICR region.

Paper II: Chromatin *In Situ* Proximity (ChrISP): Two-parameter, single-cell analysis of chromatin proximities at a high resolution

The question underlying the generation of this manuscript was: Is it possible to combine the "location" delivered by 3D DNA FISH and the "proximity" produced by the "C" technologies into a novel method that quantitatively visualize proximities at high resolution in single cells?

A method to visualize chromatin fibers proximity in single cells

The basic strategy was to analyse the proximities between chromatin fibres by labeling selected DNA probes with digoxygenin or biotin followed by ISPLA using the rolling circle amplification (RCA) approach for detection. However, we found that the rolling circle amplification was sterically hindered by structural hallmarks of the nucleus to prevent unbiased detection. Moreover, the rolling circle amplification step generated large blobs often approaching 1 micron in diameter that precluded in-depth analysis of structural features. To overcome this problem the rolling circle amplification step was omitted and replaced by a fluorescently labeled splinter. Figure 5 shows ChrISP signal clustering on nuclear periphery.



Analogous to the ISPLA principle, fluorescent modified splinter would thus bridge the proximity between two altered antibodies decorating nearby modified chromatin fibers. The splinter is stabilized by a backbone oligo and ligation. This result in continuum of fluorescent signals that efficiently spotted proximities between chromatin fibres visualized at an unmatched resolution. Using the same approach *in vitro*, on DNA fibers FISH, a resolution <170 Å was demonstrated. By exploiting the high specificity of ISPLA and the 100% efficiency of the "Green Splinter" approach, we unraveled conformation of the entire chromosome 11 in 3D perspectives, to visualize chromatin clusters in relationship to structural hallmarks of the nucleus.

Sensitivity of generated ChrISP signals

Omitting the RCA step could be perceived as a contradictory step towards high signal sensitivity generated by the 1000-fold rolled amplification of the conventional rolling circle reaction. However, using *in vitro* approaches the "Green Splinter approach" enabled the estimate that only 8 x 100 bp probes in 0,1μm^3 area were required to generate a continuum signal necessary for signal detection.

Uniqueness generated by ChrISP analysis of an individual chromosome

When examining the conformation of an entire chromosome, represented by probes covering the entire chromosome 11 territory, some non-trivial issues emerged. The generation of this probe was based on FACS-purified chromosomes followed by their fragmentation and ensuring amplification [50]. Three potential problems arise with the use of such unique probes. First, the repeat elements found within the purified chromosome 11 can also be found on other chromosomes to compromise interpretation upon FISH analysis, unless competing repeat element probes are included as quenchers. Second, the purified chromosome 11 probe is not completely free of sequences from other chromosomes and third, the representation of the unique sequences might not be uniform along chromosome 11. To find an answer to above-mentioned issues, the chromosome 11 probe was sequenced. The sequencing result revealed moderate overlap of sequences from other chromosomes. Moreover, the probes were uniformly distributed along the chromosome 11 validating the *ChrISP* technology.

Summary and perspective

To validate chromatin interactomes by a technique independent of the "C" techniques, DNA FISH have been the method of choice. This is because DNA FISH is able to score for proximities between chromatin fibres in single cells to generate frequency data in large cell populations. However, DNA FISH has a limitation represented by the diffraction of fluorophores to generate resolution that is usually 250 – 300 nm in the X Y plane and 500 – 700 nm in the Z plane. Even though advancement in microscopy have improved the resolution of fluorophores down to 100 – 150 nm, it is still not sufficient to provide fundamentally novel insight into chromatin structures. To overcome those technical limitations, we custom-tailored ISPLA – an unprecedented technique for proximity detection – to develop a novel technique Chromatin In Situ Proximity (ChrISP). The idea underlying this aim was to be able to quantitatively visualize and score for chromatin proximity in single cells, not only between fibers but also among structural hallmarks and chromatin fibers at a resolution far below any commercially available microscopes. As described above, two different moieties need to be in sufficient proximity to generate a ChrISP signal. Hence, ChrISP could be virtually applied on any combination providing greater flexibility for numerous adaptations. ChrISP will thus score for any proximity available between endogenous protein epitopes and modified DNA fiber, modified DNA fiber and another DNA fiber, protein epitope and another protein epitope. We believe that ChrISP will pave the way to assess stochastic biological processes at a high resolution within the nucleus of a single cell to generate novel horizons. This current paper has fulfilled aim number 2: **To innovate a new technique that is able to identify higher order chromatin structures in single cells at an unrivalled resolution**.

Paper III: The H3K9me2/3 mark provides spatial identity and protects against large scale reorganization of a human chromosome

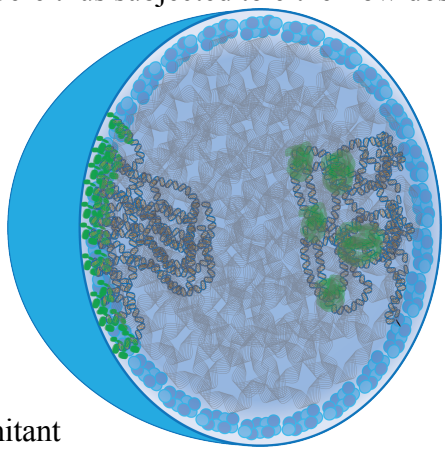
In the previous methods reports, it was demonstrated that ChrISP can be used to identify chromatin clusters within a single nucleus at an unprecedented resolution < 170Å. To further map such clusters and their regulation we expanded our aim to include an adaptation of the ChrISP technique to map chromatin marks to particular chromatin clusters.

Identification of chromatin hubs enriched with H3K9me2 mark

The millions of base pair of sequences tangled in chromatin clusters highlighted in the previous report fitted the description of the so-called Large Organized Chromatin K9 Modification (LOCKs). Such regions were mapped to the nuclear membrane and were expected to cluster together to increase the stability of the repressed mark [14]. To examine this supposition, we applied the ChrISP to map the proximity between the unique sequences of chromosome 11 and H3K9me2 protein. The resulting signals closely followed those of the chromatin clusters in the preceding report to underscore that chromatin clusters within chromosome at the nuclear membrane and LOCKS are likely one and the same. Moreover, the ChrISP signals were visualized to protrude from the bulk of the chromosome 11 territory to provide unprecedented insight into the organization of such LOCK structures. Thus, the LOCKs emerges as finger-like structures from the bulk of chromosome 11 to approach the nuclear membrane.

Inhibition of the G9a/Glp function leads to pleiotropic repression events

Since H3K9me2 mediate those hubs formation, we devised two different strategies to knock-down the function of G9a/Glp, the enzyme responsible for the generation of H3K9me2 from H3K9me1 [51]. HCT116 cells were thus subjected to either low dose of BIX-01294, a specific inhibitor of G9a/GLP methyltransferases that inhibits H3K9me2 production or treatment with siRNA targeting the G9a mRNA. Our expectations that these treatments would antagonize the LOCK-specific chromatin hubs were fulfilled. However, to our surprise, loss of G9a function also led to chromosome wide changes in chromatin proximities. We thus concluded that the loss of G9a function led to loss of nuclear membrane-associated chromatin clusters concomitant with large-scale reorganization of chromosome 11 beyond the nuclear periphery. Figure 6, depicts the ChriSP signal location in control and G9a treated cells.



Active or repressed chromatin?

To ascertain whether or not such novel features would be linked to transcriptional changes, we devised two strategies. In the first, the cells were labelled with bromouridine for 20 minutes followed by fixation and ChrISP analysis of proximities between newly transcribed RNA within chromosome 11 using specific antibodies against bromouridine and the chromosome territory probe. In the second strategy, the distribution of RNA polymerase phosphorylated within serine 2 to mark elongating polymerases, was similarly analysed. Both strategies revealed that in the normal context, transcription occurred in large structures distributed at the surface of chromosome 11. However, upon inhibition of the G9a function, either by the specific inhibitor or knock down using siRNA, it revealed that the large transcriptional factories visualized in the control cells were re-distributed to the nuclear membrane. As we have not been able to document significant overall changes in the transcriptome by inhibiting the G9a function, we argue that the character of the transcriptional machinery was reprogrammed by the loss of H3K9me2. Thus the transcriptional pattern of the bulk of the chromosome appeared to be reduced to transcription factories beyond detection in our ChrISP protocol. Conversely, the nuclear membrane-associated chromatin experienced the emergence of large transcription factories. Our interpretation of these data is that the chromosome-wide transcription factories of the control cells were dissolved while the transcriptional units, potentially the so-called euchromatic islands [14], coalesced to form larger transcription factory structures. One prediction from this interpretation is the loss and gain, respectively, of coordinated transcription in response to inhibition of the G9a function. Finally, the loss of large, chromosome-wide transcription complexes upon G9a inhibition suggests that the emergence of chromosome-wide chromatin clusters represents clusters of repressed domains. It is currently unclear what chromatin mark such domains harbor and why they would coalesce in the absence of another repressive mark, ie H3K9me2.

We conclude that the removal of the repressive H3K9me2/3 mark paradoxically leads to repression of chromatin and a rewiring of ongoing transcription as visualized by the re-distribution of active RNA polymerase II and nascent transcripts.

Physical attachment of chromosome 11 to the nuclear lamina

When ChrISP was applied on chromosome 11 and Lamin A/C proximity, both controls and treated cells didn't reveal any changes in the proximity demonstrating that the large-scale rearrangement of chromosome 11 is independent of its anchoring to the membrane. Interestingly, a dynamic pattern emerged visualized by lamin A/C, which can be physically proximal to chromatin fibres even at some distance from the nuclear membrane.

Summary and perspectives

Chromatin marks underlie the organization of higher order chromatin structures, as exemplified by the so-called Large Organized Chromatin K9 Modification (LOCKs) structures at the nuclear periphery. This structure depends on the formation of H3K9me2/3 and its interaction with the nuclear membrane. Reasoning that the ChrISP-mediated visualization of clustered chromatin at the nuclear periphery shown in the previous report represented LOCKs we further modified the ChrISP technique to include an analysis of the proximity between chromatin fibres and particular chromatin marks. By further optimization of the ChrISP protocol it was possible to demonstrate that the H3K9me2 mark actually decorates a cluster of chromatin fibres specifically at the nuclear membrane. Moreover, using the ChrISP technique we could demonstrate that such structures emerge as "fingers" from the chromosome impinging on the nuclear membrane. To confirm our observation, we treated the colon cancer cells with G9a/GLP methyltransferases inhibitor or knocking down G9a transcript using siRNA, which significantly reduced the H3K9me2/3 signal. Interestingly, not only were the cluster structures at the nuclear periphery dissolved, but new repressive chromatin clusters emerged elsewhere in chromosome 11. Using yet another modification of the ChrISP protocol we could show that this dramatic reorganization of the chromosome structure was accompanied by widespread re-distribution of nascent transcripts and active RNA polymerase II. Finally, we couldn't see any physical detachment or lost of proximity between chromosome 11 and the nuclear membrane once controls were subjected to G9a inhibitor. Our study conclude that epigenetic marks such as H3K9me2/3 can prevent large scale rewiring of chromatin structures under developmental stage or environmental cue to maintain spatial identity of chromosomes. This current paper has fulfilled aim number 3: **To use the novel ChriSP technique to identify the spatial identity of interphase chromosomes and its relationship with chromatin marks and transcriptional competence**.

Paper IV: Rainbow chromatin in situ proximity analysis of intra-chromosomal topologies in single cells

Higher order chromatin conformations underlie not only long-distance regulation of gene expression, but also define the repressive status of lineage-specific genes. Moreover, higher order chromatin structures and their transcriptional status may facilitate translocation events implicated in human cancer [52]. It is thus most desirable that such information can be teased out at the single cell level. However, despite immense efforts during several decades we know very little of such features. The reason is of course that current techniques analyzing higher order chromatin conformations are hampered by either poor resolution of individual cells or higher resolution in only large cell populations. We reasoned that the ChrISP technique could be employed once again to cover this gap. However, the ChrISP technique as it has been described above can only visualize distances less than 162Å. To expand on the ChrISP technique a new strategy was devised to generate the "Rainbow" ChrISP or rChrISP that translated different distances between chromatin fibres into different colors that can be visualized using a conventional confocal microscope.

A method to translates distances between chromatin fibres into different colours

The rChrISP technique was thus modified to include not only one splinter, but three different splinters. Each distance would represent a footprint of the actual size of the modified splinter bringing the proximity probes together. The three splinter molecules displayed atomic lengths of 11.9, 18,5 and 27,2 nm respectively and labeled with 3 different fluorescent dyes 488 (Green), Cy5 (Far Red/pink/cerise) and Cy3 (Red). The plan was that these colors would visualize different distance between chromatin fibres of an interphase chromosome, here represented by chromosome 11.

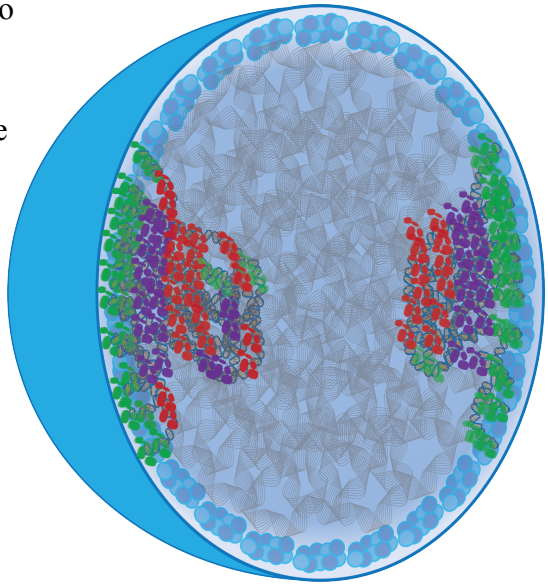


Figure 7, Highlighting the possibility of Rainbow ChrISP to uncover the 3D perspective of chromosome 11 in 3 different colors.

Specificity of the signal in relationship to the splinter length

Sequential addition of splinters was a rational and critical step during the development of the technique. Reasoning that the long splinter could easily overtake all other splinters and dominate the mosaic color pattern of the chromosome. Therefore, we performed all possible combinations to ensure that each distance is saturated and covered by the most appropriate splinter avoiding any false positive signal.

Summary and perspective

Although the conventional ChrISP technique has unraveled new insight into the dynamics of chromatin conformations in interphase cells, it is based on analysis of only one distance between chromatin fibres. To analyze the structures of chromosomes in interphase cells and how such features regulate the transcriptome and hence cell fate will likely require the exploration of many more levels of chromatin fibre proximities. To this end, we developed the rChrISP version, which is based on a multiplex proximity assay for the simultaneous visualization of the 3D topology and interactions of single unique fiber sequences in a particular chromosome with three different levels of resolution. This was achieved by designing two additional fluorescent splinters with atomic lengths of 185 and 272 Å. When combining these two splinters with the previous well-established "Green splinter" providing a resolution of less than 170Å, the aim and expectation was that each splinter would bridge distinctive proximity distances. By differential labelling with fluorescent dyes, it was possible to translate three different distances between chromatin fibres into a color code. However, to achieve this aim it was essential to add the splinters in a sequential manner. The result was further processed to generate images demonstrating globular structures of intermediate proximity surrounded more and less tight chromatin structures following a gradient from the nuclear membrane. Such observations may form a platform for better understand chromatin conformations and their link to regulation of transcriptome during normal development and disease progression. This current paper has fulfilled aim number 4: **To translate conformations within a single chromosome into a color code to uncover novel aspects of chromatin organization**.

4 ACKNOWLEDGEMENTS

This work was performed both at the Evolution Biology Centrum department (EBC), Uppsala Universitet and at the department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institutet. I would like to thank the various organizations that provided financial support especially, ***Vetenskapsrådet***, ***Cancerfonden***, ***Barncancerfonden***, ***HEROIC & CHILL*** consortiums, ***VR*** and ***Knut*** and ***Alice Wallenbergs*** stitelse. Without your support our work and research would have been impossible.

I would like to express my sincere gratitude to my main supervisor ***Rolf Ohlsson,*** for the tremendous support during those years. I am also happy that we managed with the "up and downs" we had and quite surprised you put up with my Lebanese – Mediterranean temper ☺. Thanks as well for introducing me to the "epigenetics" world and for understanding my appetite for technology developments and my humdrum when it comes to classical biological conundrum. I also appreciated your contagious enthusiasm and curiosity for science and for always exposing me to new challenges and technical adventures.

My co-supervisor, ***Ulf Landegren***, thanks you for all the vivid and inspiring talks, for the unlimited support and availability. I got imprinted with your passion to technology development from the days as a student in your lab and carried this desire with me the entire period of my Ph.D, thanks as well for the valuable inputs and suggestions on my work. Thank you truly for being an inspiration and a true role model.

My co-supervisor, ***Lazlo Szekely***, thank you, for being available when I needed your tremendous knowledge for microscopy! For all the stimulating sessions regarding fluorophores, lenses, exposure time, tips, tricks, 63x or 100x, 3D…. Thanks for your home made Image J plugins, for the 3D reconstructions discussions while building your 3D printer… For your help with the laser dissection microscope and the knowledgeable input to my super complicated *in situ* riddles.

I also would like to thank as well ***Ola Söderberg*** for all the *in situ* PLA tips and tricks, help and availability when needed. Thanks for the conversation during meetings as well while catching up on campus. Indeed, ISPLA is such an amazing adventure by itself!

I would like also to thank as well, current and previous members of the ***Ohlsson Lab***! We have been a great big happy international family. We shared a lot of good and bad times as well as we all contributed to what we have now. ***Anita Göndör***, Thank you for the continuous discussions, optimizations ideas and help with my projects. ***Marta Imreh***, a.k.a **Piri,** Thank you for always being there! You have managed the lab orders and chemicals in a way that our life became much easier. Thanks for the last second order, for the best ES cells slides, for the help to grow and maintain cells and cultures…. Thanks for always considering us as your children and always make sure we have sandwiches and fruits! ***Chengxi***, you had 6 month head start on your Ph.D ☺ still you didn't…. We shared quite a lot of time and memories… I am quite proud of what you achieved ***Dr. Shi***, good luck with your future and I am sure we meet again!

Your welcome for the '*a la mode'* wardrobe upgrade and thank you for the ChinGlish words you added to my vocabulary. ***Noriyuki***, "*sensei"* ***Sumida***, my looooooooong everlasting Ph.D friend and Chemical Calculator! I bet you could calculate the "amount of ink molecules" used in the entire Thesis! Thanks for being my first senior colleague and always defusing my technical frustration with your endless tolerance… I learned a lot from you and enjoyed every single brain storming session where we always wanted to find a technique for every single issue in science! I am sure that you will prevail soon and I won't hesitate 1 second to work again with you! Good Luck and see you on the other side ☺ ***Xingqi***, your next ☺ yep start writing! Good luck with your future career and your aim to enjoy academia! You will succeed I am sure. ***Moumita,*** ta ta ta ta ta ta ta ta, Madame ***Biswas,*** a pleasure sitting next you! A pleasure teasing you all day round! A pleasure frightening & terrifying you from time to time… Thanks for your critical discussions and correction of my thesis! Thanks for giving me "those eyes™"☺ whenever I pulled your legs! We had a lot in common especially our Anti-Acid Drugs! You will be an amazing teacher one day, enjoy your future and get the best out of it! ***Feri***, I am sure I can pronounce your full 29 letter long name, director and producer ***Farzaneh Shahin Varnoosfaderani*** please stop rolling ☺ movies keep rolling ISPLA. I am pretty sure you will own 90% of the material shown in my dissertation party! Keep that amazing friendly spirit of yours and the nice oriental dishes that you cook and welcome to **H**… You're a really good and sincere friend /colleague, take care and enjoy your freedom! And ***Feri***, your welcome to inherit all my chemicals / slides / antibodies / duties/ work / pipettes / tubes/ or whatever you find with my name labeled on it ☺. ***Li-Sophie***, sorry for not joining your relaxation sessions! I am amazed with your ability to speak several languages especially Cantonese! I could see the reaction on the native speakers faces ☺ good luck and you should consider China as a new option… you will do great science there. ***Anna lewandowska***, good luck with your future and your research career and thanks a lot for the home made backing! I will be loosing a couple of KG soon. **Maria**, Miss **Israelsson,** life is full of great moments and you will find yours soon! Good luck and consider me on your next Nature Paper ☺ ***Honglei***, you tried to teach me how to pronounce your last name ***Zhao*** but you ended up learning how to count ISPLA! All you will see from now on is Green/red/blue and the rest are just illusions! Good luck with your work. ***Manos***, re ***Emmanouil Sifakis*** being an engineer wont save you from simply pushing the button and play around with colors that later on you will call "DATA" welcome to the club… good luck understanding the complexity of our experiments and the "Biology" behind it… I am sure networking online is a completely different universe than our network!

MTC and Colleagues: I am sure I won't be able to name you all! I hope most of you will not be mad if unintentionally unmentioned! But I am sure I will make all of you smile when I tell you my favourite greeting quote: "Same S… Different Day"

I start by thanking ***Ingemar Ernberg***, I really admire your energy and devotion towards continuously organizing national and international seminars, courses, journal clubs, conferences… I have enjoyed and learned a lot from those KI cancer retreats! ***Galina Selivanova***, thanks for the collaboration and scientific discussions every time we see each other in the corridor… ***Lars-Gunnar Larsson***, I am seriously impressed with your hidden talents! Not only a great scientist but also a cool musician and to hang out with. Thanks for the cancer book examination! I really enjoyed the discussion and felt super

relaxed that I even forgotten it was an exam that took 2h! ***Joanna zawacka-pankau***, thanks for the scientific and non-scientific discussions! Good luck with your future projects and molecules… you will always find a new molecule that do MAGIC… you need to simply believe in yourself. All the best for your future career.

The MTC student Association ☺ a.k.a MSA.
Agata Korecka , Paola Martinez Murillo, Habib Abdullah Mohammad Sakil, Soazig Le Guyon , Arnika Wagner , Maria Lisa Knudsen , Marijke Sachweh, Siti Mariam Zakaria, Nicolas Ruffin…(and the unknown student ☺)
Thanks you guys for all the help during all those years. I was proud of all of you and it was always my pleasure working together on issues to improve our student conditions. I had the best time being "boss" and looking forward to work with you again in real life.

Also thanks for Dr. ***Pegah Rouhi***, Dr. ***Rozina Caridha*** and Dr. ***Hamid Sharifi***
You've been amazing! Good luck and I join your clan soon! ***Pegah***, keep walking! Life is too short so keep moving ☺ **Rozina**, now it is time to Party!!! ***Hamid***, don't be late for my party ☺ and we need to practice volleyball again!

Now it is time to thank my Friends both within the scientific bubble and the lucky ones outside! You all have made my life and time quite enjoyable! Thanks for being part of my daily life….

Uppsala! Amazing memories truly!
Erika and ***Bernhard!*** I admire you both ☺ your both a unique and lovely couple! I enjoy your company! ***Erika,*** your simply mysterious and full of surprises! I am super glad to have you as a friend! Please delete the SMS from your phone to get new ones ☺ ***Bernhard***, your approach to daily life is simply astonishing! I really look forward for those discussions every time we meet! Sweden is a lovely place…. And it is always cold here ☺ get used to it and visit us more often!
Stephan, Reelika & Ruben! I've quite enjoyed having you as close friends ☺ yep ***Stephan*** I do! Please accept that… you're an amazing cook! Im not sure why your still wasting your time and not having your own restaurant!!! ***Reelika***, thanks for being so cheerful and for the amazing company! I know your always available once needed… ***Ruben***, your simply awesome! Keep pulling that energy out from your father ☺ he needs to be in shape for his next biking event!
Loic, ***Nora*** & ***Lilea!*** Lovely family and great friends! ***Loic***, I know you've been trying hard to impress me with your French accent and "Galette des Rois" recipe! Keep trying ☺ Keep running after the white rounded sphere you call ball… I'm sure you're a professional football player! ***Nora***, thanks for always being positive and enthusiastic! I know you do all the cooking and ***Loic*** just pretend ☺ Please stop driving him everywhere! Let him get this driving license… we can go together now! Seriously! ***Lilea***, your simply gorgeous! Don't listen much to your father… I know he can overdo it sometimes ☺
René, ***Therese*** & ***Islay!*** When are we going for my favorite mushrooms hunting? ***René***, Thanks for the amazing vegetarian (including FISH) dinners… Really miss those pizzas… ***Therese***, thanks for introducing us to ludvika ☺ it is always a pleasure meeting with all of you! ***Islay***, cute little angel, welcome!

Harris, ***Nikoleta***, ***Philipe! Harris***, hope you will show up this time! Was fun when we managed to meet ☺ or actually when you showed up! I know I'm a bit harsh on you… but time to buy an electrifying watch to remind you about time! I can't be mad from you, trust me I really tried ☺ ***Niki***, I understand you now ☺ seriously ☺ Let us hope we meet much often… we don't have to wait ***Harris*** ☺ ***Philipe***, welcome to the real world! We love you no matter how much you father tries to hide you ☺
Paolo and ***Lauren***! Thanks for the time we spent together and all the nice parties. I am sure we will keep on meeting and enjoying our time! Good luck and see you soon!

Stockholm my new home!
Michael, ***Suzi*** & ***Kalle***, stop moving around Stockholm please! ***Micha,*** thanks for the home made killer booze and the nice gatherings. ***Suzi***, keep that smile it is quite contagious! ***Kalle,*** high five! Super easy-going little angel.
Punit & ***Moumita,*** Dearest friends, amazing company, great colleagues! ***Punit***, thanks for everything you've done to match my picky taste for spices! Thanks for all the great dishes you've cooked and for my favourite "Bread". Thanks for reviewing my thesis! We will keep hanging out and enjoy as always! ***Moumita,*** once again ☺ thanks you for being a good friend as well and listening to my super complicated way of approaching things. I am quite surprised you managed to survive my sense of humour! Hang in there I am not going anywhere ☺
Chengxi & ***Yen***, Good luck for your future! ***Chengxi***, thanks for the friendship outside the lab, will be meeting once again! ***Yen***, hopefully you will manage to teach him some correct Chinese pronunciation! Welcome to the clan and see you soon in China!
Vassilis Galiotos, thanks for always being available when I needed an unscientific escape! We should start going more often for beers and "mussels"… re' stop being so tired every time you run 200m… stop using those excuses to be home alone reading news! You have been with me since ages and still your trying to impress with your "eatable" food… I admit you can cook but still there is a big margin for improvements… See you next week for dinner! ***Lovisa***! Welcome to our world! thanks for the "fake-real" fantastic flowers! I am sure you will survive my humour and will enjoy the company!

Kyriakos Kokkoris, my brother in crime! ***Dr. Kyri*** I have talked, discussed, shared, worked, partied, travelled … with you a total amount of hours definitely higher than that I have spent with my real brother. Thanks for the very first days in Uppsala! I don't think that I would have managed that well without you. We share tons of memory… million minutes of fun and laughter… I am really proud of what you have become now and looking forward to spend more time with you again! I am sure you will prevail and we would sit together around 80 (if we managed that far) and look back those glory days. ***Nathalie***! I always admired your wisdom and understanding for our madness! Thanks for the moments we shared and will share in the near future!

My Family in law
Krzysztof, Maria & ***Bartosz***! I am proud to be a member of your family! Thanks for the support, joy, happiness and hospitality you offer me every time I visit my second home! Dziękuję bardzo za wszystko! zawsze z przyjemnością Was odwiedzam i cieszę się z pobytu z moją drugą rodziną! Ilekroć jestem z Wami, zawsze czuję się jak w domu! Kocham Wam i tęsknię za Wami bardzo.

My Family
No words can express those feelings!
Mansour, ***Sana*** & ***Bassem! Dr.Yammine***, you have been my inspiration, my role model and my idol! Mansour, the amazing father and friend! thank you for your endless love and support! Thank you for all those days where I felt weak and down and you never left me alone! Thank you for putting education as number one priority! Thank you for believing in me! I hope I will always make you proud. ***Sana***, the greatest mother that life can offer! I am so grateful for your patience, infinite love and care. Thank you for putting aside your career and taking amazing care of us and invest your time and effort in our education. ***Bassem***, you played a big role in my life, you have invested in my education and me! You supported me even when I was wrong and you stood always by me. Thank you for everything you have done in the past and will do in the future.

Aucun mot ne peut exprimer ces sentiments!
Mansour, ***Sana*** & ***Bassem***! ***Dr.Yammine***, vous avez été mon inspiration, mon modèle et mon idole! ***Mansour***, le père incroyable et l'ami en même temps! merci pour votre amour et votre soutien illimité! Merci pour tous ces jours où je me sentais faible et vers le bas et vous ne m'avez jamais laissé seul! Merci de mettre l'éducation comme priorité numéro un! Merci de croire en moi! J'espère que je vais toujours vous rendre fiers. ***Sana***, la meilleure mère que la vie peut offrir! Je suis tellement reconnaissant pour votre patience, l'amour et le soin infini. Merci de mettre de côté votre carrière et de prendre un soin incroyable de nous et d'investir votre temps et votre d'effort dans notre éducation. ***Bassem***, vous avez joué un grand rôle dans ma vie, vous avez investi dans mon éducation et en moi même! Vous m'avez soutenu, même si je me suis trompé et vous vous teniez toujours à côté de moi. Merci pour tout ce que vous avez fait dans le passé et vous aller faire à l'avenir.

Michalina Lewicka – Yammine! Simply my gorgeous lovely wife! Thanks ***Dr. Misia*** for your love, respect and support. I am so proud of you! Sorry for taking that long but we managed to get 3 weddings and 3 parties and 2 Ph.D. You always showed me life from different perspectives! Travelling with you around the world was the most vivid experience of my life! Thank you for being you at your best ☺ thank you for enduring and listening to me while writing this thesis!

5 REFERENCES

1. Wang, D. and S. Bodovitz, *Single cell analysis: the new frontier in 'omics'.* Trends Biotechnol, 2010. **28**(6): p. 281-90.
2. Graf, T. and M. Stadtfeld, *Heterogeneity of embryonic and adult stem cells.* Cell Stem Cell, 2008. **3**(5): p. 480-3.
3. Pederson, T., *The nucleus introduced.* Cold Spring Harb Perspect Biol, 2011. **3**(5).
4. Clamp, M., et al., *Distinguishing protein-coding and noncoding genes in the human genome.* Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19428-33.
5. Pertea, M., and Salzberg, S.L., *Between a chicken and a grape- estimating the number of human genes.* Genome Biology, 2010. **11**(206): p. 1-7.
6. TA., B. *Transcriptomes and Proteomes*. Genomes. 2nd edition 2002; Available from: http://www.ncbi.nlm.nih.gov/books/NBK21121/.
7. Zlatanova, J. and P. Caiafa, *CTCF and its protein partners: divide and rule?* J Cell Sci, 2009. **122**(Pt 9): p. 1275-84.
8. Herold, M., M. Bartkuhn, and R. Renkawitz, *CTCF: insights into insulator function during development.* Development, 2012. **139**(6): p. 1045-57.
9. Yu, W., et al., *Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation.* Nat Genet, 2004. **36**(10): p. 1105-10.
10. Kim, M.Y., T. Zhang, and W.L. Kraus, *Poly(ADP-ribosyl)ation by PARP-1: 'PAR-laying' NAD+ into a nuclear signal.* Genes Dev, 2005. **19**(17): p. 1951-67.
11. Farrar, D., et al., *Mutational analysis of the poly(ADP-ribosyl)ation sites of the transcription factor CTCF provides an insight into the mechanism of its regulation by poly(ADP-ribosyl)ation.* Mol Cell Biol, 2010. **30**(5): p. 1199-216.
12. Bartova, E., et al., *Histone Modifications and Nuclear Architecture: A Review.* J Histochem Cytochem, 2008. **56**(8): p. 711-721.
13. Chen, X., et al., *G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment.* Genes Dev, 2012. **26**(22): p. 2499-511.
14. Wen, B., et al., *Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells.* Nat Genet, 2009. **41**(2): p. 246-50.
15. Bartova, E., et al., *Histone modifications and nuclear architecture: a review.* J Histochem Cytochem, 2008. **56**(8): p. 711-21.
16. Bradbury, E.M., *Chromatin. Series in molecular biology.* J. Mol. Recognit., 1989. **2**.
17. Vignali, M., et al., *ATP-Dependent Chromatin-Remodeling Complexes.* Molecular and Cellular Biology, 2000. **20**(6): p. 1899-1910.
18. Croft, J.A., et al., *Differences in the Localization and Morphology of Chromosomes in the Human Nucleus.* J Cell Biol, 1999. **145**: p. 1119-1131.
19. Cremer, T. and M. Cremer, *Chromosome territories.* Cold Spring Harb Perspect Biol, 2010. **2**(3): p. a003889.
20. Cremer, M., et al., *Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei.* J Cell Biol, 2003. **162**(5): p. 809-20.
21. Cook, P.R., *A model for all genomes: the role of transcription factories.* J Mol Biol, 2010. **395**(1): p. 1-10.
22. Papantonis, A. and P.R. Cook, *Transcription factories: genome organization and gene regulation.* Chem Rev, 2013. **113**(11): p. 8683-705.
23. Guelen, L., et al., *Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions.* Nature, 2008. **453**(7197): p. 948-51.
24. Solovei, I., et al., *Nuclear Architecture of Rod Photoreceptor Cells Adapts to Vision in Mammalian Evolution.* Cell, 2009. **137**: p. 356-368.
25. Misteli, T., *Beyond the Sequence Cellular Organization of Genome Function.* cell, 2007. **128**: p. 787-800.

26. Shrestha, P., I. Shrestha, and K. Kurisu, *Immunohistochemistry: A Review of Practical Procedure.* Nepal Journal of Neuroscience, 2009. **6**: p. 38-41.
27. tsien, R.Y., *The green fluorescent protein.* annu. Rev. Biochem 1998. **67**: p. 509-544.
28. Sun, Y., et al., *Forster Resonance Energy Transfer Microscopy and Spectroscopy for Localizing Protein–Protein Interactions in Living Cells.* Cytometry, 2013. **83A**: p. 780-793.
29. Jayasena, S.D., *Aptamers: An Emerging Class of Molecules That Rival Antibodies in Diagnostics.* Clinical Chemistry, 1999. **45**(9): p. 1628-1650.
30. Lofblom, J., et al., *Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications.* FEBS Lett, 2010. **584**(12): p. 2670-80.
31. Gusev, Y., et al., *Rolling Circle Amplification.* The American Journal of Pathology, 2001. **159**(1): p. 63-69.
32. Soderberg, O., et al., *Direct observation of individual endogenous protein complexes in situ by proximity ligation.* Nat Methods, 2006. **3**(12): p. 995-1000.
33. Clausson, C.M., et al., *Increasing the dynamic range of in situ PLA.* Nat Methods, 2011. **8**(11): p. 892-3.
34. de Wit, E. and W. de Laat, *A decade of 3C technologies: insights into nuclear organization.* Genes Dev, 2012. **26**(1): p. 11-24.
35. Markaki, Y., et al., *The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture: 3D structured illumination microscopy of defined chromosomal structures visualized by 3D (immuno)-FISH opens new perspectives for studies of nuclear architecture.* Bioessays, 2012. **34**(5): p. 412-26.
36. Huang, B., et al., *Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy.* Science, 2008. **319**(5864): p. 810-3.
37. Henriques, R., et al., *PALM and STORM: unlocking live-cell super-resolution.* Biopolymers, 2011. **95**(5): p. 322-31.
38. Tanke, H.J., R.W. Dirks, and T. Raap, *FISH and immunocytochemistry: towards visualising single target molecules in living cells.* Curr Opin Biotechnol, 2005. **16**(1): p. 49-54.
39. Gibcus, J.H. and J. Dekker, *The hierarchy of the 3D genome.* Mol Cell, 2013. **49**(5): p. 773-82.
40. Weibrecht, I., et al., *Visualising individual sequence-specific protein-DNA interactions in situ.* N Biotechnol, 2012. **29**(5): p. 589-98.
41. Gomez, D., et al., *Detection of histone modifications at specific gene loci in single cells in histological sections.* Nat Methods, 2013. **10**(2): p. 171-7.
42. Phillips, J.E. and V.G. Corces, *CTCF: Master Weaver of the Genome.* Cell, 2009. **137**(7): p. 1194-1211.
43. Feil, R., et al., *Developmental control of allelic methylation in the imprinted mouse Igf2 and H19 genes.* Development, 1994. **120**: p. 2933-2943.
44. Kanduri, C., et al., *Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive* Current Biology, 2000. **10**: p. 853-856.
45. Pant, V., et al., *The nucleotides responsible for the direct physical contact between the chromatin insulator protein CTCF and the H19 imprinting control region manifest parent of origin-specific long-distance insulation and methylation-free domains.* Genes Dev, 2003. **17**: p. 586-590.
46. Ame, J.C., C. Spenlehauer, and G. de Murcia, *The PARP superfamily.* Bioessays, 2004. **26**(8): p. 882-93.
47. Caiafa, P. and J. Zlatanova, *CCCTC-binding factor meets poly(ADP-ribose) polymerase-1.* J Cell Physiol, 2009. **219**(2): p. 265-70.
48. D'Amours, D., et al., *Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions.* biochem. J., 1999. **342**: p. 249-268.
49. Scobie, K.N., et al., *Essential role of poly(ADP-ribosyl)ation in cocaine action.* Proc Natl Acad Sci U S A, 2014. **111**(5): p. 2005-10.
50. Lichter, P., et al., *HIGH-RESOLUTION MAPPING OF HUMAN CHROMOSOME 11 BY IN SITU HYBRIDIZATION WITH COSMID CLONES.* Science, 1990. **247**: p. 64-69.

51. Shinkai, Y. and M. Tachibana, *H3K9 methyltransferase G9a and the related molecule GLP.* Genes Dev, 2011. **25**(8): p. 781-8.
52. Roukos, V., B. Burman, and T. Misteli, *The cellular etiology of chromosome translocations.* Curr Opin Cell Biol, 2013. **25**(3): p. 357-64.

6 PUBLICATIONS & MANUSCRIPTS