

# Nucleolar Epigenomics by Dynamic Molecular Combing and Nucleolar Matrix Analysis

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER  
NATURWISSENSCHAFTEN (DR. RER. NAT.) DER FAKULTÄT FÜR  
BIOLOGIE UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT  
REGENSBURG

vorgelegt von  
KARINA ZILLNER  
aus München, Deutschland

im Jahr 2013

# Eidesstattliche Erklärung

Hiermit erkläre ich, Karina Martina Zillner, geboren am 26. August 1984 in München, an Eides statt, die vorliegende Promotionsarbeit mit dem Titel:

## **"Nucleolar Epigenomics by Dynamic Molecular Combining and Nucleolar Matrix Analysis"**

selbstständig verfasst zu haben unter alleiniger Verwendung der angegebenen Hilfsmittel. Die aus anderen Quellen direkt oder indirekt übernommenen Daten oder Konzepte sind unter Angabe des Literatur- bzw. Personenzitats gekennzeichnet.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

**Das Promotionsgesuch wurde eingereicht am:**

4. Februar 2012

**Die Arbeit wurde angeleitet von:**

PD Dr. Attila Németh

Regensburg, Februar 2013

---

(Karina Zillner)

# Contents

<b>Summary</b>	<b>III</b>
<b>List of Figures</b>	<b>V</b>
<b>List of abbreviations</b>	<b>VII</b>
<b>1 Introduction</b>	<b>1</b>
1.1 DNA modifications . . . . .	1
1.1.1 5-methylcytosine . . . . .	2
1.1.2 Oxidative products of 5-methylcytosine . . . . .	4
1.1.3 Overview of DNA modification analyses . . . . .	6
1.1.4 The importance and difficulty of locus-specific modification analysis of repetitive DNA . . . . .	7
1.2 Nucleolar genome content . . . . .	7
1.2.1 Satellite-2 repeats . . . . .	8
1.2.1.1 Genomic organization of satellite-2 DNA . . . . .	8
1.2.1.2 Methylation of satellite-2 repeats . . . . .	8
1.2.2 Ribosomal RNA genes . . . . .	9
1.2.2.1 Genomic organization of ribosomal DNA . . . . .	9
1.2.2.2 Epigenetics of ribosomal DNA . . . . .	15
1.2.2.3 Tip5 as a key regulator of rDNA silencing . . . . .	18
<b>2 Results</b>	<b>23</b>
2.1 Single molecule epigenomics by Dynamic molecular combing . . . . .	23
2.1.1 Suitable conditions for <i>in vitro</i> combing of $\lambda$ DNA . . . . .	23
2.1.2 Psoralen-combing . . . . .	25
2.1.2.1 <i>In vitro</i> psoralen-combing on $\lambda$ phage DNA . . . . .	25
2.1.2.2 Psoralen-combing on genomic DNA . . . . .	30
2.1.3 Chromatin-combing . . . . .	34
2.1.4 Epi-combing . . . . .	35
2.1.4.1 <i>In vitro</i> epi-combing on $\lambda$ phage DNA . . . . .	37
2.1.4.2 Epi-combing on genomic DNA . . . . .	40
2.2 DNA methylation analysis of satellite-2 DNA . . . . .	43
2.2.1 Satellite-2 DNA methylation analysis by metaphase immuno- FISH on metaphase chromosome spreads . . . . .	43
2.2.2 Satellite-2 DNA methylation analysis by epi-combing . . . . .	43
2.3 DNA methylation analysis of ribosomal RNA genes . . . . .	46
2.3.1 rDNA methylation analysis by metaphase-immunoFISH on metaphase chromosome spreads . . . . .	46
2.3.2 Epigenetic analysis of ribosomal DNA repeats reveal epi- genetic clusters and transitions . . . . .	49

2.3.3	Comparison of ribosomal DNA methylation in Werner Syndrome cells and primary fibroblasts . . . . .	55
2.4	Association of ribosomal DNA with the nuclear matrix . . . . .	64
2.4.1	Tip5 targets rDNA to the Nucleolar Matrix . . . . .	64
2.4.2	AT-hooks of Tip5 display MAR binding with comparable affinities . . . . .	65
2.4.2.1	DNA binding features of potential MAR binding domains of Tip5 . . . . .	65
2.4.2.2	Sequestering of rDNA to the nucleolar matrix requires a functional Tip5 molecule . . . . .	71
2.5	Extended AT-hooks as a novel DNA binding motif . . . . .	73
<b>3</b>	<b>Discussion</b>	<b>75</b>
3.1	Single molecule epigenomics by Dynamic molecular combing . . . . .	75
3.1.1	Psoralen-combing for Sequence-specific Single Molecule Chromatin Analysis . . . . .	75
3.1.2	Epi-combing . . . . .	77
3.2	DNA methylation patterns of tandem repeat arrays . . . . .	80
3.2.1	DNA methylation analysis of satellite-2 repeats . . . . .	80
3.2.2	DNA methylation analysis of ribosomal DNA . . . . .	81
3.2.2.1	DNA methylation clustering and transition of ribosomal DNA repeats . . . . .	81
3.2.2.2	Epigenetic status of non-canonical rDNA repeats . . . . .	83
3.3	Large-scale organization of ribosomal DNA chromatin is regulated by Tip5 . . . . .	85
3.3.1	Targeting of rDNA to the nuclear matrix . . . . .	85
3.3.2	DNA binding features of Tip5 and its functional consequences . . . . .	86
3.3.3	A biological role for NoRC binding to DNA . . . . .	88
<b>4</b>	<b>Materials and Methods</b>	<b>90</b>
4.1	Materials . . . . .	90
4.1.1	Technical devices . . . . .	90
4.1.2	Chemicals and Reagents . . . . .	91
4.1.3	Consumables . . . . .	93
4.1.4	Software and online tools . . . . .	94
4.1.5	Buffers, Solutions and Media . . . . .	94
4.1.6	Enzymes . . . . .	96
4.1.7	Kits . . . . .	97
4.1.8	Standard DNA and protein weight markers . . . . .	98
4.1.9	Antibodies . . . . .	98
4.1.9.1	Primary antibodies . . . . .	98
4.1.9.2	Secondary and tertiary antibodies . . . . .	99
4.1.10	Bacterial strains and mammalian cell types . . . . .	100
4.1.11	Plasmids . . . . .	101
4.1.12	Oligonucleotides . . . . .	102
4.2	Methods . . . . .	103
4.2.1	Working with nucleic acids . . . . .	103
4.2.1.1	Isolation of nucleic acids . . . . .	103
4.2.1.2	DNA purification . . . . .	103



---

4.2.1.3	Checking the quality of nucleic acids . . . . .	104
4.2.1.4	Polymerase chain reaction . . . . .	104
4.2.1.5	Restriction digest of DNA . . . . .	107
4.2.1.6	Ligation of DNA fragments . . . . .	108
4.2.1.7	DNA Sequencing . . . . .	108
4.2.2	Protein Methods . . . . .	108
4.2.2.1	Denaturing polyacrylamide gel electrophoresis . . .	108
4.2.2.2	Protein Quantification . . . . .	109
4.2.2.3	Semi-dry Western Blot and protein detection . . .	110
4.2.2.4	Purification of GST-tagged proteins . . . . .	111
4.2.3	Mammalian cell culture . . . . .	112
4.2.3.1	Maintenance . . . . .	112
4.2.3.2	Transfection of mammalian cells . . . . .	113
4.2.3.3	Detection of senescent cells . . . . .	113
4.2.4	Nucleolus isolation . . . . .	113
4.2.5	Microscopy methods . . . . .	115
4.2.5.1	Fixation of adherently growing cells . . . . .	115
4.2.5.2	Immunofluorescence . . . . .	116
4.2.5.3	Fluorescence <i>in situ</i> hybridization - FISH . . . . .	116
4.2.6	Dynamic molecular combing variants . . . . .	119
4.2.6.1	DNA plug preparation of adherent cells . . . . .	119
4.2.6.2	DNA isolation for combing . . . . .	119
4.2.6.3	The combing process . . . . .	120
4.2.6.4	<i>In vitro</i> psoralen-combing . . . . .	120
4.2.6.5	Chromatin-combing . . . . .	122
4.2.6.6	Epi-combing . . . . .	122
4.2.6.7	Fluorescence <i>in situ</i> hybridization for all combing variants . . . . .	123
4.2.7	Nuclear matrix preparation . . . . .	124
4.2.8	Microscale thermophoresis . . . . .	125
4.2.8.1	Microscale thermophoresis to quantify DNA:protein binding interactions . . . . .	125
4.2.8.2	Preparation of DNA template . . . . .	127
4.2.8.3	Preparation of the titration series . . . . .	128
4.2.8.4	Measurement and analysis of the binding affinity .	129
<b>5</b>	<b>Appendix</b>	<b>131</b>
5.1	Curriculum vitae . . . . .	131
5.2	Scientific contributions . . . . .	133
5.3	Grants, awards and fellowships . . . . .	135
<b>6</b>	<b>Acknowledgements</b>	<b>138</b>
	<b>Bibliography</b>	<b>139</b>

## Summary

With the sequencing of the human genome essentially complete, a comprehensive view is now needed on how the genetic information is established in different cell types and developmental stages and how it is inherited through cell divisions. Nowadays, the focus lies on chemical modifications to DNA and histone proteins that form a complex regulatory network to modulate genome function which is called the epigenome of a cell.

DNA modifications, particularly the methylation of DNA in the context of CpG, is such an important epigenetic mark that regulates gene expression and therefore, adds extra information to the genome of mammalian cells. DNA methylation has an important role in development, differentiation and genome stability and aberrant DNA methylation profiles are found in many diseases such as cancer [36], autoimmune diseases [86], neurodevelopmental disorders and neurodegenerative diseases [87].

Thus, deciphering the information that is encoded in DNA modifications is necessary for fully understanding not only the mechanisms of epigenetic regulation but also their dysfunctions that cause different diseases. Therefore, various methods have been established for the analysis of DNA modifications (mainly DNA methylation). However, it is very difficult or impossible to examine the epigenetic linkage of two distant chromosomal loci, i.e. the associated inheritance of epigenetic modifications at different genomic loci through cell divisions with established experimental approaches. Additionally, crucial information about the genomic localization is lost when repetitive DNA is analyzed. Therefore, novel single molecule methods become increasingly important. In this thesis, different single molecule methods are established, that either enable the locus-specific analysis of nucleosomal patterns or DNA modifications.

The psoralen-combing assay applies the Dynamic molecular combing method on psoralen crosslinked DNA. The technique allows sequence-specific, genome-wide single molecule analysis of chromatin structures based on their psoralen accessibility, at the megabase scale with kilobase resolution. Proof-of-principle experiments are performed demonstrating that the method can be potentially applied for single molecule characterization of chromatin structures, particularly at tandem repeat arrays.

Epi-combing is described as a novel sequence-specific single molecule method to analyse DNA modifications by combining Dynamic molecular combing with DNA modification immunodetection. It can be applied for the investigation of DNA methylation patterns at large, highly repetitive regions of genomic DNA, as well as for epigenetic linkage analysis of kilobase-sized, modified DNA bearing chromosomal regions. Proof-of-principles were performed on native and *in vitro* modified  $\lambda$  DNA and these experiments demonstrate that the detection of DNA modifications on single DNA fibers is specific, sensitive as well as selective.

In addition, epi-combing was successfully applied to the analysis of tandem repeat arrays of human genomic DNA, namely satellite-2 DNA and ribosomal DNA. Both sequence classes build the core of the nucleolar genome [95] and are highly repetitive, which makes their DNA methylation status difficult to analyse with conventional technologies. By examining satellite-2 DNA with epi-combing in HCT116 cells, heavily methylated stretches were revealed, which display an average length of 76 kb, however, satellite-2 DNA stretches of 200 kb length were found, too.

The analysis of ribosomal DNA by epi-combing further revealed novel insights into the organization of DNA methylation patterns on these repetitive transcription units. The standard concept is that ribosomal RNA genes are composed of arrays of identical repeats that are clustered in a mainly head-to-tail orientation on specific chromosomal loci, the "nucleolar organizer regions" (NORs).

It is assumed that an entire NOR is either completely active containing unmethylated rDNA or completely inactive bearing methylated DNA, which is also propagated throughout the cell cycle.

Data presented in this work confirm the epigenetic clustering of many repeats, however transitions of DNA methylation status between individual repeats could also be found which challenges the standard view on NOR activity. In addition, the epigenetic state of non-canonical rDNA repeats could be revealed in this work. By epi-combing, the silencing of those aberrant repeats via DNA methylation could be proven, and reduced DNA methylation levels of non-canonical repeats were shown in malignant and senescent cells.

Ribosomal RNA genes are regulated by different epigenetic mechanisms including chromatin remodelling besides DNA methylation. Another major regulator of ribosomal DNA is Tip5, the large subunit of the Nucleolar remodelling complex (NoRC). NoRC is a multifunctional chromatin-dependent regulator of rRNA genes, which regulates nucleosome positioning, transcriptional repression, epigenetic silencing and replication timing. However, little is known about its role in large-scale spatial organization and distribution of actively transcribed and inactive rRNA gene copies. Nuclear matrix isolations enable a simple biochemical characterization of large-scale chromatin organization and therefore, this method was applied to reveal the function of Tip5 in large-scale chromatin organization of the rDNA locus.

Since the TAM domain and AT-hooks of Tip5 are predicted MAR binders [2], Tip5 may mediate the anchoring of rDNA to the nuclear matrix and thus, separate silenced rDNA repeats from active ones. In the presented work, experiments were performed that prove the Tip5-dependent reorganisation of rDNA in the nuclear matrix by overexpressing of the protein. Furthermore, the DNA binding ability of Tip5 was shown in microscale thermophoresis experiments that demonstrate the potential of AT-hook binding to rDNA and MARs. Additional findings show that the combination of AT-hooks together with the TAM domain of Tip5 are sufficient for nuclear matrix targeting and anchoring, however, the entire Tip5 protein is necessary for rDNA specificity.

In summary, by establishing and applying new methods such as epi-combing and microscale thermophoresis, this work provides novel insights into the epigenetic regulation of mammalian rRNA genes, a process that must be tightly balanced due to the highly energy-demanding metabolic activity of the cell.

# List of Figures

1.1	Mechanisms for DNA methylation mediated repression . . . . .	3
1.2	Standard Model of ribosomal RNA gene organisation . . . . .	10
1.3	Principle of Dynamic molecular combing (DMC) . . . . .	12
1.4	Structural analysis of the human rDNA locus by DMC . . . . .	13
1.5	Discrimination between naked and nucleosomal DNA by psoralen intercalation . . . . .	16
1.6	NoRC and its role in rDNA silencing . . . . .	20
1.7	Modular structure of Tip5 and its binding partners . . . . .	21
2.1	Different concentrations of $\lambda$ DNA stained with Yoyo-1 . . . . .	24
2.2	Fluorescence hybridization of $\lambda$ DNA using differently labelled probes	24
2.3	Psoralen-combing on $\lambda$ DNA . . . . .	25
2.4	Reverse crosslink of psoralen treated $\lambda$ DNA . . . . .	27
2.5	Detection of psoralen-biotin on combed $\lambda$ DNA . . . . .	29
2.6	Comparison of different crosslinking times of psoralen-biotin to $\lambda$ DNA and subsequent detection of PB and DNA . . . . .	31
2.7	Psoralen-biotin combing on naked genomic DNA . . . . .	32
2.8	Psoralen-biotin combing on genomic DNA after cell lysis by 1% Sar- cosyl . . . . .	33
2.9	Psoralen-biotin combing on genomic DNA after cell lysis by 0.5% NP40	34
2.10	Combing of native chromatin fibers . . . . .	35
2.11	Principle of epi-combing . . . . .	36
2.12	Control digestion of modified $\lambda$ DNA with MspI and HpaII . . . . .	37
2.13	<i>in vitro</i> 5mC-combing on $\lambda$ phage DNA . . . . .	39
2.14	Detection of DNA methylation on $\lambda$ phage DNA using a AMCA- coupled tertiary antibody . . . . .	40
2.15	<i>in vitro</i> 5hmC-combing on $\lambda$ phage DNA . . . . .	41
2.16	Two examples of DNA preparations stained by Yoyo-1 . . . . .	42
2.17	Diversity of 5mC signals on human genomic DNA . . . . .	42
2.18	DNA methylation analysis of satellite-2 DNA by metaphase immuno- FISH . . . . .	44
2.19	Satellite-2 DNA is present in long, methylated stretches in HCT116 cells . . . . .	45
2.20	Co-existence of inactive and active NORs during metaphase of female blood lymphocytes . . . . .	47
2.21	Correlation of UBF and rDNA methylation reveals gradual shades in NOR activity in female blood lymphocytes . . . . .	48
2.22	Colocalization of UBF and rDNA methylation reveals gradual de- creases in NOR activity in male blood lymphocytes . . . . .	50
2.23	Locus-specific detection of ribosomal DNA by epi-combing . . . . .	51
2.24	Epigenetic clustering of ribosomal DNA in Imr90 cells . . . . .	53

---

2.25	Epi-combing analysis illustrates transitions of epigenetic states between neighbouring repeats . . . . .	54
2.26	DNA methylation analysis of senescent cells by epi-combing reveals hypomethylation at non-canonical repeats . . . . .	56
2.27	DNA methylation analysis of HCT116 cells by epi-combing . . . . .	57
2.28	DNA methylation analysis of MCF7 cells by epi-combing . . . . .	58
2.29	DNA methylation analysis of AG05283 cells by epi-combing . . . . .	60
2.30	DNA methylation analysis of AG13077 cells by epi-combing . . . . .	61
2.31	DNA methylation analysis of AG12797 cells by epi-combing . . . . .	62
2.32	DNA methylation analysis of AG06300 cells by epi-combing . . . . .	63
2.33	Overexpression of Tip5 leads to enrichment of rDNA in the NM . . . . .	66
2.34	GST-purification of the AT-hook peptides of Tip5 and HMGA1 . . . . .	67
2.35	Binding behaviour of HMGA1 to ribosomal DNA in three different capillary types determined by Microscale thermophoresis . . . . .	68
2.36	Binding of AT-hooks on ribosomal DNA measured by Microscale thermophoresis . . . . .	69
2.37	Binding behaviour of different AT-hook peptides on PRDII measured by Microscale thermophoresis . . . . .	70
2.38	Overexpression of different GFP-Tip5 peptides . . . . .	72
2.39	Identification of a novel DNA binding motif . . . . .	74
3.1	Discrimination of open and closed rDNA chromatin by psoralen-combing . . . . .	76
3.2	Comparison of read length and resolution between different single molecule analysis methods . . . . .	79
4.1	Nucleolus isolation quality control sheet . . . . .	114
4.2	Principle of Microscale thermophoresis . . . . .	126

# List of abbreviations

$\lambda$ DNA	$\lambda$ bacteriophage DNA
2M	High-Salt fraction of nuclear matrix preparations
5azadC	5-aza-2'-deoxycytidine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
<i>E. coli</i>	Escherichia coli
$^{\circ}\text{C}$	degree Celsius
$A_x$	absorbance at x nm
aa	amino acid
ACF	ATP-utilizing chromatin assembly factor
Amp	ampicillin
APS	ammonium persulfate
ATP	adenosine-5'-triphosphate
bp	base pair
BSA	bovine serum albumine
C-terminal	carboxy terminal
CDK	cyclin-dependent kinase
CHD	Chromodomain
CHR	Chromatin fraction of nuclear matrix preparations
CHRAC	Chromatin accessibility complex
CP	Cytoplasmic fraction of nuclear matrix preparations
CpG	cytosine-phosphatidyl-guanosine
CSK	cytoskeleton
D $\alpha$ G	Donkey-anti-goat
D $\alpha$ H	Donkey-anti-human
Da	Dalton
DMC	Dynamic molecular combing
DMSO	Dimethylsulfoxide
DNA	deoxyribonucleic acid
Dnmt	DNA-cytosine-5-methyltransferase
Dnmt1	DNA-cytosine-5-methyltransferase 1
Dnmt2	DNA-cytosine-5-methyltransferase 2
Dnmt3a	DNA-cytosine-5-methyltransferase 3a
Dnmt3b	DNA-cytosine-5-methyltransferase 3b
dNTP	2'-deoxynucleotide triphosphate

EC50 .....	Equilibrium constant
EDTA .....	ethylenediaminetetraacetate
EMSA .....	Electrophoretic mobility shift assay
EtOH .....	ethanol
FCS .....	fetal calf serum
FISH .....	Fluorescence in situ hybridization
g .....	gram
G $\alpha$ M .....	Goat-anti-mouse
G $\alpha$ R .....	Goat-anti-rat
G $\alpha$ Rb .....	Goat-anti-rabbit
gDNA .....	genomic DNA
h .....	hour
H3K4me2 .....	dimethylation at lysine 4 of histone H3
H4K16ac .....	acetylation at lysine 16 of histone H4
HDAC .....	histone acetyltransferase
HM .....	Hybridization mixture
HMT .....	histone methyltransferase
HP1 .....	heterochromatin binding protein 1
Hr XX .....	Human ribosomal DNA part XX
IFN .....	Interferon
Ig .....	immunoglobulin
IGS .....	Intergenic spacer
IPTG .....	isopropylthiogalactoside
ISWI .....	imitation of switch
K <sub>D</sub> .....	Dissociation constant
kb .....	kilobase pair
kDa .....	Kilodalton
l .....	litre
LB .....	Luria-Bertani
M .....	molar
MAR .....	Matrix attachment region
MBD .....	Methyl-CpG-Binding domain
min .....	minute
MST .....	Microscale thermophoresis
MW .....	molecular weight
N-terminal .....	amino terminal
NM .....	Nuclear Matrix
nm .....	nanometer
NOR .....	Nucleolar organizer region
NoRC .....	Nucleolar remodeling complex
NT .....	Nick translation
nt .....	nucleotide

---

PAA .....	polyacrylamide
PB .....	Psoralen-biotin
PBS .....	phosphate buffered saline
PCR .....	polymerase chain reaction
PHD .....	plant homeodomain
PIC .....	pre-initiation complex
Pol I .....	RNA polymerase I
PRDII .....	DNA dodecamer comprising a five base pair A, T tract from the PRDII element of the IFN $\beta$ enhancer element
qPCR .....	quantitative PCR
r <sub>cf</sub> .....	relative centrifugal force
rDNA .....	ribosomal DNA
RNA .....	ribonucleic acid
rpm .....	revolutions per minute
RT .....	room temperature
SAM .....	S-adenosyl-L-methionine
Sat2 .....	Satellite 2 repeats
SDS .....	sodium dodecyl sulfate
SDS-PAGE .....	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec .....	second
SNF2H .....	sucrose non-fermentor 2 homolog protein
StrA488 .....	Alexa488-coupled Streptavidin
TAM .....	Tip5/ARBP/MBD
Taq .....	Thermus aquaticus
TBE .....	Tris borate EDTA
TE .....	Tris EDTA
TEMED .....	N,N,N',N'-tetramethylethylenediamine
TET .....	ten-eleven translocation (tet methylcytosine dioxygenase)
Tip5 .....	TTF-I interacting protein 5
Tris .....	Tris(hydroxymethyl)aminomethane
TTF-1 .....	Transcription termination factor-1
U .....	unit
UBF .....	Upstream binding factor
UCE .....	upstream control element
WB .....	Western Blot
WCE .....	whole cell extract
WS .....	Werner Syndrome

The common abbreviations are used for prefixes, chemical and physical units, elements, essential amino acids as well as bases occurring in DNA.





# 1 Introduction

The sequencing of the human genome is essentially complete, however this sequence is only the basis of how the genetic program is made manifest that finally leads to the phenotype of an organism. Superimposed is another layer of "heritable" gene regulatory information, that is called epigenetics. Epigenetics is the study of heritable changes other than those in the DNA sequence and this information is stored as chemical modifications at cytosine bases (DNA modifications) and histone proteins that package the genome. By regulating chromatin structure and DNA accessibility, these chemical marks influence how the genome is read in different developmental stages, tissue types and disease states (reviewed in [9]).

The sum of such epigenetic information in the cell represents the epigenome, i.e. the different functional states of the genome, which can be inherited through cell division. Epigenomics, the merged analysis of epigenetics and genomics, aims to understand the genetic regulation and its contribution to cellular growth and differentiation, disease and ageing (reviewed in [21]). Normal cellular function relies on correct epigenome establishment and many reports show the causality between epigenetic factors and diseases. To fully understand complex human diseases, epigenome-wide studies may deliver essential information. However, the epigenome is specific to tissues and cell types at different developmental stages, and they change over time making the analysis of the epigenome far more complex than deciphering the human genome. Thus, to achieve this, techniques must be developed that are high throughput, quick and cheap.

Therefore, the challenges of the post-genomic era include comprehensive understanding of epigenome structure, function and other dynamics which will yield a richer understanding on how the genetic code is established across a diverse background of developmental stages, cell types and diseases.

## 1.1 DNA modifications

Modification of DNA bases was found in the genome of many different organisms including bacteria, protozoa, fungi, plants and vertebrates and it has diversified roles. In prokaryotes it serves as a protection against degradation of the DNA by restriction enzymes, regulates gene expression, DNA repair, cell cycle and pathogenicity [25] [27] [43, 80]. In addition, DNA modification in eukaryotes has a role also in development, differentiation and genome stability [12].

There are DNA modifications such as phosphorothioates in bacteria, the deamination of cytosine to uracil in innate and adaptive immunity as well as base J in certain parasites such as trypanosomes [16, 34, 148, 147, 98]. Furthermore, ribonucleotides occur at least in yeast to regulate mate switching [144].

However, the covalent attachment of a methyl group to DNA bases is the most

frequently studied DNA modification. There are different types of DNA methylation depending on the site of the modification. N4-methylcytosine (4mC), N6-methyladenine (6mA), as well as C5-methylcytosine (5mC) have been shown to be widespread in bacteria, whereas 6mA and 5mC can also be found in eukaryotes [113].

Enzymatic oxidation products of 5mC to C5-hydroxymethylcytosine (5hmC), C5-formylcytosine (5fC) and C5-carboxylcytosine (5cC) are found in eukaryotes. The mechanism of oxidation and its physiological relevance in altering the binding sites of genome regulator proteins that specifically recognize DNA modifications have been partially revealed [60, 140, 156, 70].

All these chemical marks add extra information to the genome and they have the potential to change the transcriptional state of a gene, acting in combination with histone posttranslational modifications and chromatin regulator proteins, altogether establishing the epigenome of a cell (reviewed in [168]).

Noteworthy, 5mC is commonly referred to as 'DNA methylation' despite the fact that methylation naturally occurs also at other sites of DNA bases. Following this simplified nomenclature, 'DNA methylation' and '5mC' are used as synonyms also in this work.

### 1.1.1 5-methylcytosine

#### Role of 5-methylcytosine

DNA methylation was found in mammals as early as DNA was identified as the genetic material [3],[83]. In 1948, modified cytosine was first discovered in a preparation of calf thymus using paper chromatography. It was proposed that it was 5-methylcytosine (5mC) because it separated from cytosine in a way similar to thymine separating from uracil.

In animals, the extent of genomic DNA methylation is highly variable. At the lowest extreme is the nematode worm *Caenorhabditis elegans* which lacks detectable 5mC levels and *Drosophila melanogaster* only has low levels of 5mC. Most other invertebrates have a relatively high level of 5mC in large domains separated by domains of unmethylated DNA, showing thus mosaic methylation patterns. The highest levels of 5mC are observed in vertebrates being dispersed over the entire genome, referred to as global DNA methylation pattern (reviewed in [12]).

DNA methylation has an important role in development, differentiation and genome stability, whereas atypical DNA methylation profiles are found in many diseases such as cancer [36], autoimmune diseases [86], neurodevelopmental disorders and neurodegenerative diseases [87].

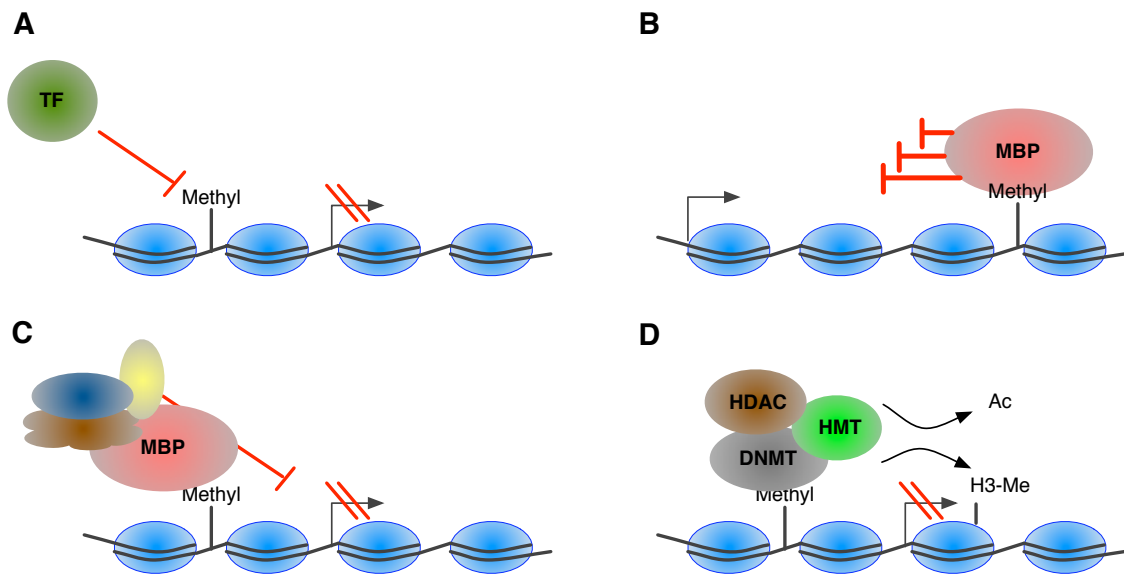
In addition, DNA methylation is essential for silencing retroviral elements, regulating tissue-specific gene expression, genomic imprinting, and X chromosome inactivation.

Approximately 45% of the mammalian genome consists of transposable and viral elements that are silenced by bulk methylation [123]. The vast majority of these elements are inactivated by DNA methylation or by mutations acquired over time as the result of the deamination of 5mC [146]. If expressed, these elements are potentially harmful as their replication and insertion can lead to gene disruption and DNA mutation [91], [72] [49], [142], [157].

## Transcriptional regulation

Many researchers suggested that DNA methylation might regulate gene expression, however, it was not until the 1980s that several studies showed that DNA methylation was involved in cell differentiation and gene regulation. Nowadays, it is well recognized that DNA methylation, in concert with other regulators, is a major epigenetic factor influencing gene activity.

There are different possible mechanisms for DNA methylation to regulate gene expression (see Figure 1.1).



**Figure 1.1: Mechanisms for DNA methylation mediated repression**

**A:** By blocking activators such as transcription factors from binding target sites, DNA methylation directly inhibits transcription **B:** DNA methylation in the gene body may inhibit transcriptional elongation, either directly or by the surrounding chromatin structure **C:** DNA methylation is directly recognized by methyl-CpG-binding proteins which recruit co-repressor proteins. **D:** DNA methyltransferases (DNMTs) are physically linked to histone deacetylases (HDACs) and histone methyltransferases (HMT) which modify the surrounding to a repressed chromatin structure. (after [67])

The inhibition of transcription can either be directly or indirectly caused by DNA methylation.

When DNA methylation is present in the cognate sequence of a transcription factor binding site, transcription is directly blocked since no binding of the activator can occur (see Figure 1.1 A). Another direct mechanism was shown when DNA methylation was found in the gene body, exons as well as introns (see Figure 1.1 B). Here, the elongation of the RNA polymerase was impeded either by DNA methylation itself or by a repressive surrounding chromatin structure. Though, there are also evidences showing a high level of DNA methylation is associated with gene expression in dividing cells [93]. Therefore, how DNA methylation in the gene body exactly regulates gene expression remains under investigation.

In addition, to the direct repressive effects of DNA methylation, the mark can be recognized by different classes of proteins such as methyl-CpG-binding proteins that

recruit co-repressors to the site to silence transcription (see Figure 1.1 C). Furthermore, DNA methyltransferases (DNMTs) serve a dual role, an enzymatic and non-enzymatic role, in repressing transcription (see Figure 1.1 D). First, they add the methyl-group to the base by an enzymatic reaction to yield 5mC, and second, they are physically linked to histone deacetylases and histone methyltransferases, thus silencing transcription by a biochemical interaction to modify near chromatin (reviewed in [67]).

In summary, DNA methylation and histone modifications work closely together to regulate gene expression.

### **CpG sites and CpG islands**

In mammals, DNA methylation is not uniformly distributed along cytosines of the genome but is mainly present in the context of the dinucleotide CpG. This palindromic dinucleotide leads to DNA methylation on both sides of the DNA strands, thus revealing a possible mechanisms to inherit this DNA modification over cell divisions.

However, recent studies found evidence of non-CpG methylation in mouse and human embryonic stem cells [74]. In addition, experiments with the murine frontal cortex has revealed that there is a significant percentage of methylated non-CpG sites. Though, the role of non-CpG methylation is still unclear.

Although DNA methylation at CpG dinucleotides is essential, mammalian genomes are depleted of CpG sites because of the potential of 5mC that can deaminate to thymine, thus causing genetic mutations.

These underrepresented CpG sites are spread across the genome where they are heavily methylated. Though, higher levels of CpG sites can be found in so called CpG islands, often present at gene promoters.

CpG islands are stretches of DNA roughly 1000 base pairs long that have a higher CpG density than the rest of the genome but often are unmethylated. The majority of gene promoters (about 70%), reside within these CpG islands. Interestingly, CpG islands are highly conserved between mice and humans, especially when they are associated with promoters. This conservation of CpG islands in evolution implies that these regions possess functional importance such as to promote gene expression by regulating the chromatin structure and transcription factor binding.

However, the role of CpG islands in regulating gene expression is still not fully understood. CpG islands, especially those associated with gene promoters, are mostly unmethylated. For example, although CpG islands in gene bodies and intragenic regions can show tissue-specific DNA methylation patterns, the CpG islands at promoter regions rarely show tissue-specific DNA methylation. Instead, regions called CpG shores which are located as far as 2kb from CpG islands, have highly conserved patterns of tissue-specific DNA methylation (reviewed in [93]).

#### **1.1.2 Oxidative products of 5-methylcytosine**

5mC was generally considered to be a relatively stable epigenetic modification, however studies showed that global erasure of 5mC takes place in specific embryonic stages and that DNA methylation patterns can be dynamically regulated upon cell

differentiation. These findings suggest enzymes that are capable of erasing or modifying existing DNA methylation. Ten-eleven translocation (TET) proteins were recently identified in mammalian cells as members of a family of DNA hydroxylases that possess enzymatic activity toward 5mC. These TET proteins have the ability to convert 5mC into 5-hydroxymethylcytosine (5hmC) and other oxidative products such as 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) through three consecutive oxidation reactions (reviewed in [156]).

Although the biological significance of Tet-mediated oxidation of 5mC is unclear, the relative abundance of these modified bases suggests that they may either serve as additional epigenetic marks or as intermediates in the process of demethylation. For example, 5hmC is found in significant levels in embryonic stem cells and also some adult tissues such as mouse cerebellar DNA [70] [140]. Interestingly, 5hmC is present in brain tissues but not in other metabolically active non-proliferating cells [70]. Therefore, as potentially stable base, 5hmC might influence chromatin structure and local transcriptional activity by recruiting selective 5hmC-binding proteins or excluding methyl-binding proteins (MBPs) that normally recognize 5mC, leading to replacement of factors that normally bind 5mC. Indeed, it was already demonstrated that the methyl-binding protein MeCP2 has a lower affinity to 5hmC [143]. Furthermore, 5hmC, but not 5mC, was shown to be enriched at many intergenic cis-regulatory elements, such as active enhancers and insulator-binding sites. Interestingly, 5hmC was found to be enriched in the gene body of highly transcribed genes as well as promoters that are repressed by Polycomb repression complex 2. These findings taken together suggest a potential role in transcriptional regulation by 5hmC, both activation and repression (reviewed in [156]).

Alternatively, 5hmC may play a role in passive or active DNA demethylation.

It was shown that the maintenance DNA methyltransferase Dnmt1 does not recognize 5hmC during DNA replication. Therefore, conversion of 5mC to 5hmC may exclude Dnmt1 from binding and therefore, passive demethylation is facilitated in proliferating cells (reviewed in [156]). Already a minor reduction in the fidelity of maintenance methylation would result in an exponential decrease in DNA methylation over cell divisions. Additionally, 5hmC might be an intermediate in replication-independent active demethylation pathways. As a first example, 5mC oxidation derivatives can serve as substrates for mammalian DNA glycosylases or deaminases. Recent studies showed that thymine DNA glycosylase (TDG) is able to excise 5fC or 5caC in the context of the dinucleotide CpG. Thus, subsequent repair of the resulting abasic site by the base excision repair pathway (BER) would generate unmethylated cytosine (reviewed in [156]).

In an alternative second way, 5hmC may first be deaminated by the AID (activation induced deaminase)/APOBEC (apolipoprotein B mRNA editing enzyme complex) family of cytidine deaminases to produce 5-hydroxymethyluracil (5hmU). This is followed by 5hmU:G mismatch repair through the action of DNA glycosylases such as TDG and eventually, the BER pathway.

A glycosylase independent third pathway could be the iterative oxidation of 5hmC to 5caC followed by a decarboxylation step. However, the existence of the putative decarboxylase still needs to be uncovered (reviewed in [156]).

In summary, the discoveries of the sixth base, 5hmC and of additional 5mC oxidation derivatives (5fC and 5caC) in the mammalian genome alter our perception on how DNA methylation status may be regulated in mammalian cells. Neverthe-

less, open questions about the different roles of these DNA modifications in the cell remain which results in a strong interest in quantifying and mapping genomic distribution of the modified cytosine bases.

However, commonly used approaches for DNA methylation studies cannot discriminate 5hmC from 5mC. A full appreciation of the biological significance of the modified cytosine bases requires the development of tools that allow 5hmC, 5mC, 5fC, 5caC and cytosine to be distinguished unequivocally.

### 1.1.3 Overview of DNA modification analyses

Deciphering the information that is encoded in DNA modifications is necessary for fully understanding not only the mechanisms of epigenetic regulation but also their dysfunctions that cause different diseases. In addition, DNA methylation marks may also serve as biomarker for diagnostics.

Therefore, various methods have been established for the analysis of DNA modifications (predominantly for 5mC) and their features are summarized in several comprehensive reviews. A series of analytical techniques such as thin layer chromatography [145, 158], high performance liquid chromatography [71] and capillary electrophoresis-based techniques [38, 39] allow the quantification of the overall methylation level of genomic DNA.

Another set of methods enables the analysis of specific genomic loci. They can be divided into single molecule analysis tools and ensemble methods, the latter one measuring the average of DNA methylation on a population of DNA molecules. All common high-throughput methods so far require sample pretreatment which can be either bisulfite conversion [40], methylation-sensitive enzyme restriction digestion [11] or affinity purification [35, 149]. Depending on the specific research need, the pretreatment is combined with different sample readouts such as two-dimensional gel electrophoresis, microarray platforms or sequencing. High-throughput microarray and sequencing analyses introduce further variability in the characterization of the epigenome, because in these assays both the probe preparations and the bioinformatics platforms differ.

Bisulfite sequencing was perhaps the most widely used method to investigate DNA methylation in the past two decades. Although the assay provides information about the modification of single molecules, it is limited to the analysis of short, few hundred base pair DNA fragments, and cannot distinguish between different modifications of cytosine residues such as methylation and hydroxymethylation [57].

Noteworthy, it is very difficult or impossible to examine the epigenetic linkage of two distant chromosomal loci, i.e. the associated inheritance of epigenetic modifications at different genomic loci through cell divisions with the aforementioned experimental approaches. Additionally, crucial information about the genomic localization is lost when repetitive DNA is analyzed. Therefore, novel single molecule methods become increasingly important (reviewed in [168]).

### 1.1.4 The importance and difficulty of locus-specific modification analysis of repetitive DNA

More than half of the human genome consists of repetitive DNA and a big proportion of DNA methylation takes place there. The repeats belong to the following major types: (1) retrotransposons that are dispersed throughout the genome, such as short interspersed nuclear elements (SINEs) including Alu repeats, furthermore long interspersed nuclear elements (LINEs) and long terminal repeat- (LTR-) retrotransposons; (2) satellite repeats that can be mainly found at centromeric as well as subtelomeric and telomeric chromosomal regions and major satellites; (3) ribosomal DNA repeats that are clustered on the short arms of acrocentric chromosomes [65]. It has been demonstrated that repeats can create hotspots for recombination events and DNA methylation at these sites is crucial for genome integrity and stability by formation of heterochromatin and assembly of nuclear compartments. In addition, non-coding RNAs transcribed from DNA repeats serve an important role in heterochromatin assembly of large chromosomal domains [97, 79, 104, 102]. Accordingly, aberrant DNA methylation of specific repetitive DNA elements was shown to be a feature of various human diseases [154, 33], e.g. different malignancies [26, 127, 105, 73], Alzheimer disease [15] or facioscapulohumeral dystrophy [46]. Hence, investigating the DNA methylation status of repetitive sequence elements may serve as important diagnostic markers. However, ensemble methods, that are the most commonly used ones also in repetitive DNA analysis, provide only limited information since they cannot distinguish between individual repeats or alleles and are not able to visualize epigenomic patterns of repeats [161, 150, 56].

The major drawback of these methods is the limited sequence read length, which prevents in most cases the accurate mapping of the genomic localization of repetitive sequences. This means that a large proportion of DNA methylation is present at genomic regions which can be investigated in detail only with single molecule analysis approaches (see review [168]).

## 1.2 Nucleolar genome content

The largest subnuclear compartment, the nucleolus, was first observed more than 100 years ago by light microscopic studies. The nucleolus is a nuclear organelle where rRNA transcription, processing and early steps of ribosome assembly take place. In addition, further roles of the nucleolus were described recently that involve many cellular processes such as senescence, RNA modification, cell cycle regulation and stress sensing.

To serve this multifunctional role of the nucleolus, it displays not only a specific nucleolar proteome but also specific DNA domains that are associated with the nucleolus (nucleolus-associated chromatin domains: NADs) [95] in addition to ribosomal DNA.

For example, three decades ago a frequent association of the centromeres of chromosome 1 and 9 and the heterochromatin of chromosome Y with nucleoli was observed. The recent high-resolution and high-throughput analysis of NADs showed that a specific subset of the genome is associated with nucleoli and that these are enriched in different sequence features such as satellite repeats, members of the zinc finger,



olfactory, defensin and immunoglobulin protein-coding gene families as well as active 5S rRNA genes and tRNA genes.

Ribosomal DNA is flanked in linear sequence by satellite and low-copy D4Z4 repeats and due to this physical proximity, these sequences belong to the core of the nucleolar genome. Interestingly, D4Z4 repeats, satellite DNA and ribosomal DNA belong to the class of tandem repeat arrays, the latter two described in more detail below (reviewed in [97]).

### **1.2.1 Satellite-2 repeats**

The sequence composition of NADs suggests that specific satellite repeats are important for the formation of heterochromatin at the nucleolar periphery.

#### **1.2.1.1 Genomic organization of satellite-2 DNA**

The classical satellite definition stems from the characterisation of the most abundant repeat family of density gradient fractions. Satellite-2 DNA was one of the repetitive sequences that were first discovered in humans. Initially, it was described as sequences that are digested by *Hinf*I into a large number of very small fragments (10-80bp) [110], [41].

Megabase-sized pericentromeric regions of chromosomes 1 and 16 contain most of these sequences with the higher order satellite-2 repeat units being larger than 1 kb and built from two related units of 23 and 26 bp. They contain the abundant GGAAT and ATC sequences, and a less common sequence that in principle can be methylated, CGAAT [63],[54].

Satellite DNA was assumed to be junk DNA without any function or selfish genes that try to accumulate in the genome. However, the fact that they are concentrated in distinct heterochromatin such as the pericentromeric regions which are important for chromosome segregation together with the observations of distinct behaviour of satellite-2 containing regions in genetic alterations led to the conclusion that these sequences participate in crucial functions such as DNA segregation and genome stability. However, due to the complexity of the highly repetitive nature, they were not extensively studied.

#### **1.2.1.2 Methylation of satellite-2 repeats**

Although the GC content ( 40%) of these sequences is lower than in CpG islands (>50%), the proportion of CpG dinucleotides is similar to the statistically expected value which is characteristic to CpG islands. The importance of DNA methylation at these tandem-repeated sequences was analyzed by different disease studies.

For example, mutations in the DNA methyltransferase *Dnmt3b* is associated with human ICF (immunodeficiency, centromeric heterochromatin instability, facial anomalies) syndrome lead to a selective loss of DNA methylation at heterochromatic regions such as hypomethylation of satellite-2 DNA at pericentromeric heterochromatin [159, 99, 153, 54, 53].

Pericentromeric satellite hypomethylation in ICF syndrome apparently promotes

decondensation of heterochromatin as well as other cytogenetic abnormalities. In addition, satellite hypomethylation may also play a role in higher order gene regulation [159],[52] and genome stability [159] [109].

Furthermore, hypomethylation of satellite-2 DNA is linked to various diseases such as cancer [141] [116] and is associated with a worse prognosis. Investigations in breast adenocarcinomas, non-neoplastic breast tissues, and normal somatic tissue demonstrated that malignant tumors show hypomethylation at satellite-2 compared to non-neoplastic breast tissue samples and normal controls. However, these differences were not reported in centromeric satellite DNA [62].

Taken together, the data demonstrate a crucial role of DNA methylation at satellite-2 DNA in genome stability and prevention of various diseases.

## 1.2.2 Ribosomal RNA genes

### 1.2.2.1 Genomic organization of ribosomal DNA

Each human ribosome contains in addition to numerous ribosomal proteins four ribosomal RNA (rRNA) molecules playing indispensable roles in translation. To produce sufficient amounts of rRNA for the highly abundant ribosomes, the genes encoding rRNA (rDNA) are present in multiple copies in the genome. The 5S rRNA molecule is mainly transcribed from a cluster of repeated 2.2 kb genes on chromosome 1q42 by RNA polymerase III [138]. However, the focus of this work lies on the rRNA gene which is transcribed in the nucleolus by RNA polymerase I (Pol I). Therefore, the term "rDNA" as well as "rRNA genes" will be used as the transcription unit transcribed by Polymerase I.

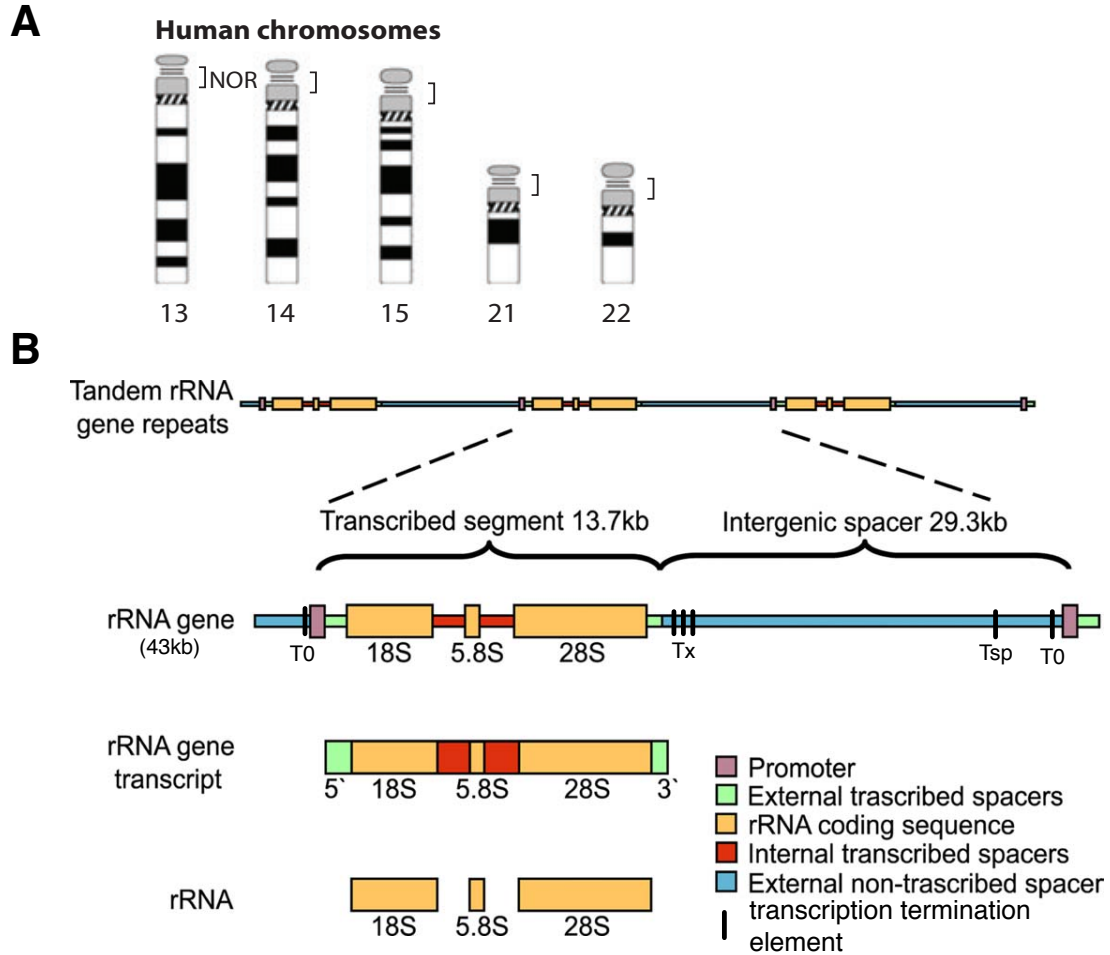
### Ribosomal RNA genes and nucleolar organizer regions

The standard concept is that arrays of mammalian rRNA genes are composed of identical repetitive transcription units that are clustered on specific chromosomal loci. The sequences that encode 18S, 5.8S and 28S rRNA are present in a single transcription unit and transcribed by Pol I to yield a 47S rRNA precursor transcript which is processed to generate one molecule each of 18S, 5.8S and 28S rRNA. Human diploid cells contain about 400 copies of rRNA repeats which are organised in mainly tandem-repeated arrays on the short arms of the five acrocentric chromosomes 13, 14, 15, 21 and 22 in a telomere-to-centromere-orientation (see Figure 1.2 A) [85, 10]. These clusters of rDNA repeats are called "nucleolar organizer regions" (NORs) and pulse-field gel electrophoresis of genomic DNA digested with enzymes that do not cut human rDNA revealed that most human NORs are composed of  $\sim 70$  copies of rRNA genes. Additionally, it was demonstrated that they only consist of rDNA since this restriction digestion led to only one major rDNA band of 3 Mb as well as several minor bands of 1 and 2 Mb. Noteworthy, NORs are isolated in linear sequence from other regions through heterochromatin formation of satellite repeats [85] and low-copy number D4Z4 repeats [77] adjacent to rDNA. This positioning separates rRNA genes from genes transcribed by Pol II and Pol III.

Each ribosomal DNA repeat unit <sup>1</sup> consists of a  $\sim 13$  kb long sequence encoding the

<sup>1</sup>GenBank accession number:U13369

precursor rRNA and of an Intergenic Spacer (IGS) of  $\sim 30$  kb which contains the transcription regulatory elements (see Figure 1.2 B). The standard model of rRNA gene clusters is that they are arranged in a telomere-to-centromere orientation, where all transcribed segments are orientated in the same direction. This is also called a head-to-tail orientation, where the 5'-end of the coding region is depicted as head and the 3'-end is described as tail.



**Figure 1.2: Standard Model of ribosomal RNA gene organisation**

(A) Brackets indicate the localisation of rDNA clusters on five human acrocentric chromosomes (out of [85]). (B) Shown are tandemly repeated rDNA repeats consisting of Intergenic Spacer (blue) and rRNA genes (yellow). An rDNA repeat is transcribed to yield one molecule 18S, 5.8S and 28S each. Black bars indicate transcription termination elements,  $T_{sp}$ ,  $T_0$  being upstream of the transcription start site and  $T_x$  being located at the 3'-end of the transcribed region. Picture is modified after [112] and not true to scale.

Pol I starts transcription at the rDNA promoter and synthesizes the long precursor containing internal and external transcribed spacer sequences in addition to the 18S, 5.8S and 28S coding regions. The gene promoter has two important regulatory sequences, the upstream control element (UCE) and the core element [117]. The UCE has a modulatory role and the core element is essential for accurate transcription initiation. The spacing between these two elements is very important as well as the relative orientation. Several transcription-termination elements ( $T_x$ ) are

located at the 3'-end of the transcribed region of the rDNA as well as two elements located upstream of the rRNA gene transcription start site ( $T_0$  and  $T_{sp}$ ) [67].

It was found that essential components of the Pol I transcription machinery such as UBF and Pol I remain associated with NORs during mitosis. These components contain acidic/argyrophilic domains which can be visualised by silver nitrate. It is believed that NORs associated with transcription factors during mitosis, termed "competent" NORs, are the only NORs being actively transcribed during interphase while the others, non-competent NORs remain silent throughout the interphase. This provides an explanation for the apparently constant number of transcribed NORs after cell cycling [130].

### Non-canonical ribosomal DNA

rDNA belongs to the class of repetitive tandem array DNA which is difficult to sequence and annotate. So far, one ribosomal RNA gene was entirely sequenced and therefore, it was assumed for a long time that the arrangement of this exemplarily represents all repeats in the genome. However, using the Dynamic molecular combing method (DMC) in combination with a subsequent fluorescence hybridization step, this knowledge about rDNA was proven wrong ([90]).

DMC is a technology for the direct visualization and analysis of single DNA fibers. It provides clear visualization of kinetics of DNA replication as well as large scale DNA rearrangements, which are difficult to detect with other techniques (e.g. inversions and duplications).

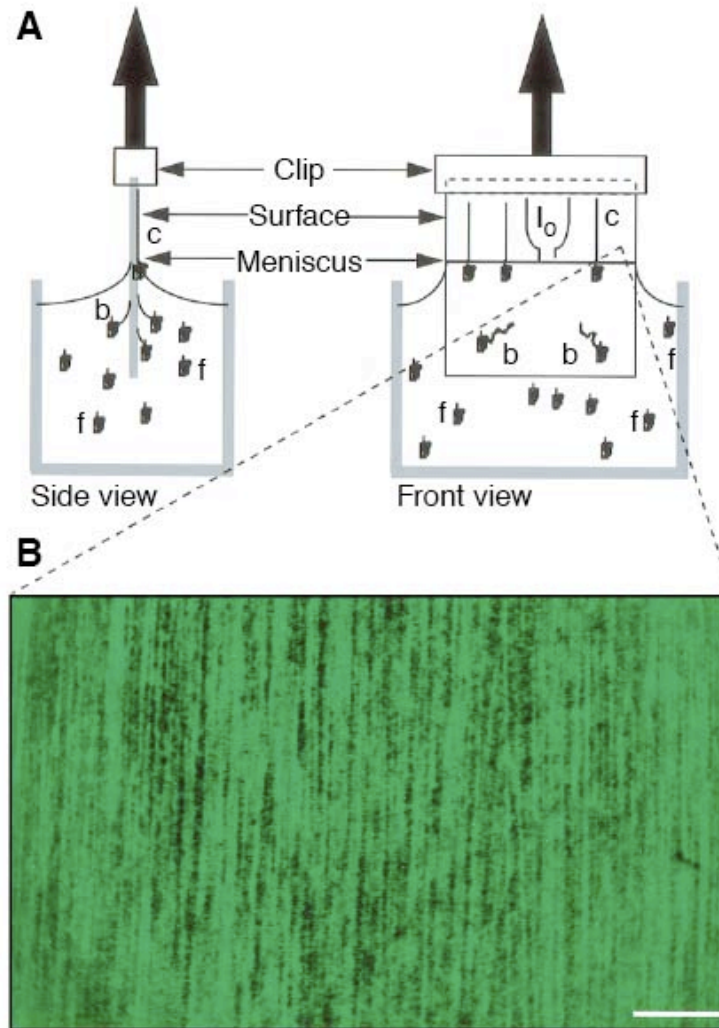
In principal, DMC consists of four steps (see Figure 1.3):

(i) preparation of cover slip surfaces coated with trichlorosilane, (ii) preparation of a DNA solution from cells embedded in low-melting agarose plugs, (iii) incubation of the surface in the DNA solution for 5 min, (iv) extraction of the glass cover slip out of the solution at a constant vertical speed.

In this process, a silanized coverslip is dipped vertically into a solution of high-molecular weight DNA and the surface is incubated for 5 min in the solution. During this incubation, DNA molecules bind to the surface by their extremities only. Then, the coverslip is pulled out slowly (300  $\mu\text{m}/\text{sec}$ ) with a mechanical device while the receding meniscus (air - solution - glass interface) exerts a restoring force that unwinds DNA coils and aligns and stretches DNA fibers. Because of their hydrophobicity, the silanized glass surfaces dry instantly as they are pulled out of the solution and are thereby irreversibly fixed on the surface [29, 90].

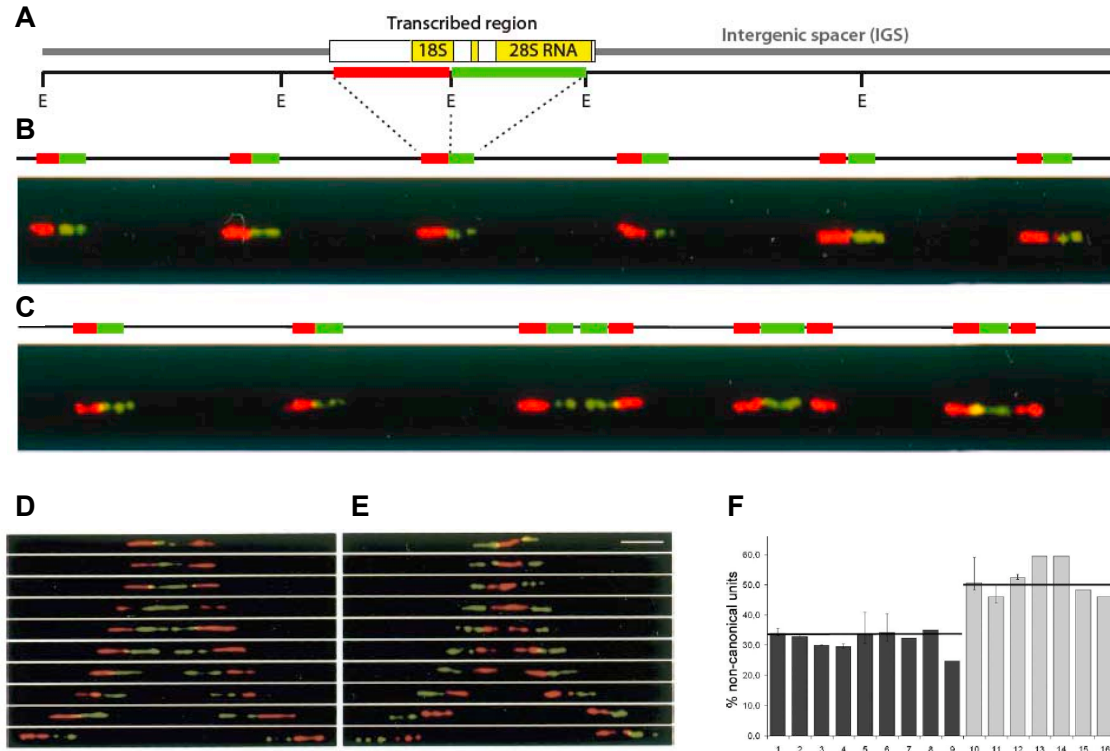
This results in parallel DNA fibers, aligned in a single direction all over the surface and a constant stretching factor of 2 kb/ $\mu\text{m}$ . DMC enables the locus-specific detection on single molecules and is therefore perfectly suited for the analysis of tandem repeated arrays such as ribosomal DNA. The standard organization assumed for rDNA units is a transcribed region followed by a non-transcribed spacer, as described in 1.2.2.1.

When the rRNA-coding regions were analyzed by fluorescence hybridization on combed molecules of DNA with two specific probes that cover their entire length,



**Figure 1.3: Principle of Dynamic molecular combing (DMC)**

**A:** High-molecular weight DNA is prepared using low-melting agarose plugs and finally, the solution is poured into a teflon reservoirs. Silanized cover slips are incubated for 5 min in the DNA solution during which freely floating molecules (f) bind by their extremities to the silanized surface (b). Afterwards, the surface is pulled out vertically at a constant speed ( $300 \mu\text{m}/\text{sec}$ ) using a molecular combing apparatus. The horizontal air-water meniscus exerts a constant and localized downward-vertical force and thereby aligns and stretches the coiled DNA molecules (c). Loops can be observed when DNA fibers are bound by both extremities ( $l_0$ ). **B:** One field of view is shown of human genomic DNA which is observed with an epifluorescence microscope. DNA fibers span several fields of view corresponding to hundreds of micrometers ( $1 \mu\text{m} = 2 \text{ kb}$ ) and the density of the DNA fibers is close to maximum (bar:  $25 \mu\text{m}$ ). Figure taken from [90]



**Figure 1.4: Structural analysis of the human rDNA locus by DMC**

**A:** Schematic representation of an rDNA repeat. Restriction analysis with EcoRI yields four distinct fragments spanning the coding region and the IGS (recognition sites of EcoRI indicated with E). The orientation of individual repeats can be distinguished by two colour-fluorescent hybridization on human rDNA using two hybridization probes (5'-end of coding region in Texas red; 3'-end of the coding region in FITC). Gaps between the signals indicate the non-hybridizing spacer segments. **B:** Shown are six canonical rDNA repeats arranged in a head-to-tail orientation. **C:** Image displays human rDNA containing non-canonical repeats. Two canonical units (left) are found next to three palindromic units with each half of the coding region joined by 3'-ends and (3'-3' palindromes) separated by a short spacer segment. **D:** - **E:** Variability in rDNA spacer length **D:** Series of hybridization signals showing the variability in rDNA spacer length for 3'-3' palindromes **E:** rDNA spacer length for 5'-5' palindromes **F:** Percentage of non-canonical repeats in human cell lines. Dark grey are control cell lines and light grey are Werner syndrome patient cells' (numbers in parentheses indicate number of repeats analyzed): (3) AG05283 (1640); (4) AG13077 (1549); (9) IMR90 (6357); (13) AG12797 (1275); (14) AG06300 (1275); Horizontal bars: average percentage. (out of [20] and [85])

this structure was confirmed in many cases. Additionally, unorthodox patterns were observed as well (see Figure 1.4).

DMC with subsequent two-colour hybridization of rDNA was performed using two adjacent fluorescent probes, a red 5'-probe (length 5.9 kb) and a green 3'-probe (length 7.1 kb), covering the coding region of rRNA genes (see Figure 1.4 A). For a large proportion, the basic head-to-tail arrangement (canonical) with an intervening IGS of rDNA was observed. (see Figure 1.4 B). However, with great average variability in spacer length ( $34.2 \pm 5.4$  kb) and single spacers ranging from 9 to 72 kb.

Another difference from canonical patterns were observed in about one-third of analyzed repeats. Clusters of genes with novel structures were found and called non-canonical repeats (see Figure 1.4 C), typically palindromic structures. These typical palindromes can be distinguished between 3'-3' palindromes (the peripheral 5' coding sequences are joined by central 3'-sequences) as in Figure 1.4 C and 5'-5'-palindromes (with 3' peripheral regions joined by central 5' regions).

The analysis of palindromic structures demonstrated a wide variety in the size of the central hybridizing region ranging from compact structures (gap < 4 kb) to large gaps (> 36 kb) between adjacent inverted repeats. This variation in spacer length can be displayed as 'Eiffel tower' forms (see Figure 1.4 D and E).

Analysis of thousands of ribosomal RNA gene hybridizing sequences revealed nearly one-third in palindromic arrangements, in control cells from both adult and fetal samples and independent of age (see Figure 1.4 F, dark bars). Additionally, samples of patients with the disease Werner Syndrome (WS) were also under investigation because of the well-characterized lesion involving a RecQ DNA helicase [162] and the resulting genomic instability. Since the WRN protein is a nucleolar protein, the pathological change may be observed at rDNA. Werner syndrome is a genetic disorder that is characterized by the rapid appearance of features associated with ageing. Individuals with WS generally develop normally until the third decade of life, when premature ageing phenotypes and a series of age-related disorders begin. Interestingly, rDNA of the WS cells displayed a significant increase in the level of non-canonical repeats because the mean incidence of such units is very close to 50% (see Figure 1.4 F, light grey bars). This increased level of palindromic structures in WS samples could also be related to an increase in the rDNA methylation state of fibroblast cultures from WS patients, which may be related to inactivation of transcription units following rearrangements [78]. Also described were additional, slow migrating linear ribosomal DNA species resulting from decreased EcoRI digestion, which were more easily detected in the most senescent passages of fibroblasts from older individuals and WS patients.

As summary, DMC experiments showed that there are other and more complex arrangements of ribosomal RNA genes in addition to the standard model. If those repeats are functional or silenced pseudogenes is not yet proven and still under investigation. One would predict that these repeats are non functional and silenced to avoid base pairing of antisense transcripts to pre-rRNA, which could seriously disturb ribosome biogenesis. The fact that those repeats have the same sequence but different orientations makes them difficult to analyze with other methods than DMC. So far, the view is proven that rDNA genes are more heterogeneous and dynamic than previously assumed.

### 1.2.2.2 Epigenetics of ribosomal DNA

As described, there are about 400 copies of rRNA gene repeats in a diploid cell. Though, not all of them are active at the same time but active and inactive repeats appear in a cell. Particularly, about half of the repeats seem stably silenced in differentiated cells. rDNA repeats exist in distinct epigenetic states that can be distinguished by a specific chromatin structure that is maintained throughout the cell cycle.

#### Analysis of different chromatin states using psoralen

The epigenetic chromatin states can be distinguished by the treatment with the chemical psoralen. Psoralen (4,5,8-trimethylpsoralen, Trioxsalen) photo-crosslinking was used to investigate chromatin structure already more than 30 years ago [23, 22]. Psoralens are tricyclic organic molecules, which intercalate in helical DNA and form covalent crosslinks between pyrimidines of opposite strands upon irradiation with ultraviolet A (UVA) light (366 nm). Psoralen reacts photochemically with naked DNA and produces interstrand crosslinks. In contrast, the formation of covalent psoralen:DNA adducts is inhibited in chromatin, where the presence of nucleosomes does not permit psoralen intercalation.

In principle, two different bands can be distinguished on agarose gel electrophoresis experiments. The active, psoralen-accessible, heavily crosslinked DNA fragments migrate slower (s-band), whereas the inactive, only partially crosslinked fragments migrate faster (f-band) (Figure 1.5). Consequently, psoralen crosslinking can be used to selectively label genomic regions, which are not properly assembled into nucleosomes or devoid of them (out of [169]).

Psoralen cross-linking assays in a variety of organisms have shown that two classes of rRNA genes coexist in growing cells. These active and silent rDNA clusters can be distinguished by their pattern of DNA methylation, specific histone modifications, and distinct nucleosome positions (reviewed in [85]).

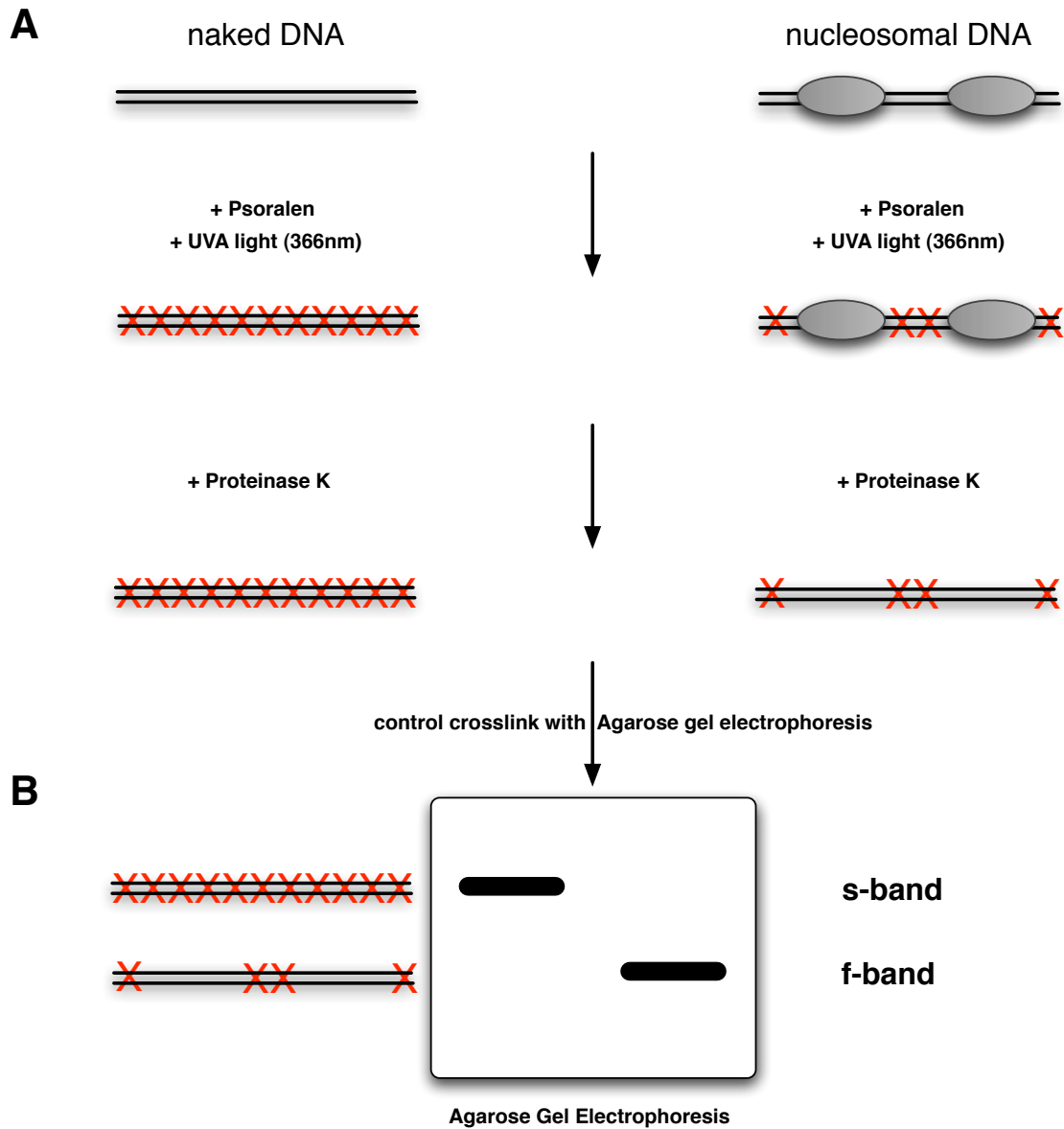
#### Open chromatin state of ribosomal DNA

It is assumed that active rDNA repeats are clustered and complete NORs are active which is propagated throughout the cell cycle. Therefore, active NORs remain undercondensed in mitosis (secondary constriction) and may stay associated with Pol I and Pol I-specific transcription factors such as UBF and TTF-I. Generally, open ribosomal chromatin exhibits euchromatic features and is permissive to transcription. Open chromatin is nucleosome-depleted and therefore, accessible for psoralen resulting in a slow migrating s-band. Active genes are hypomethylated as well as associated with acetylated histones H4 and H3 and with H3K4me3, all markers for active chromatin (reviewed in [85]).

#### Closed chromatin state of ribosomal DNA

In closed chromatin, different markers for inactive DNA can be found. In general, it has a heterochromatic conformation and is transcriptionally refractive, thus in-





**Figure 1.5: Discrimination between naked and nucleosomal DNA by psoralen intercalation**

**A:** Psoralen intercalates in naked DNA whereas intercalation is not possible in nucleosomal DNA. Upon irradiation, psoralen is crosslinked with DNA.

**B:** After Proteinase K digestion, different migration behaviours are analyzed by agarose gel electrophoresis resulting in a slow migrating band (s-band) and a fast migrating band (f-band) reflecting nucleosome-depleted and nucleosomal DNA, respectively.

active NORs are indistinguishable from surrounding heterochromatin. Because of this, psoralen cannot intercalate in nucleosomal DNA which displays a fast migrating band in agarose gel electrophoresis. In addition, inactive repeats are devoid of Pol I and Pol I-specific transcription factors. The promoter of hypermethylated, silent genes is associated with methylated H3K9, H3K20, and H3K27 indicating inactive chromatin (reviewed in [85]).

Finally, there is a strong correlation of inactive rDNA repeats and DNA methylation, which is described in more detail below.

### DNA methylation at the rRNA gene locus

Immunofluorescence analysis and biochemical fractionation demonstrated that the enzymes that catalyze DNA methylation (Dnmt1, Dnmt3a and Dnmt3b) localise in the nucleolus. It was shown that all three Dnmts predominantly localise with inactive rDNA promoters at the perinucleolar heterochromatin whereas UBF is associated with active promoters [122]. Furthermore, cells lacking functional Dnmt1 show a structurally disorganised nucleolus [35]. All these data indicate that DNA methylation plays an important role in rDNA regulation.

Indeed, sequence analysis of the rRNA transcription initiation region of the human rRNA gene showed that it is highly enriched in CpG sites since the promoter harbours a CpG island which encompasses 19 CpGs in the upstream control element (UCE) and six CpGs in the core promoter element (CPE). Therefore, the human rDNA promoter contains 25 CpGs whereas the mouse and rat rDNA promoters contain only one and five CpGs, respectively. The methylation status of each CpG was investigated within different *cis* elements in the rDNA promoter of human livers and their potential alterations in hepatocellular carcinomas (HCCs) [42].

The results demonstrate that the rRNA promoter, particularly in the UCE, is heavily methylated in human liver cells and significantly hypomethylated in hepatocellular carcinomas. It was further shown that DNA methylation has an inhibitory effect on rRNA gene expression in human cells and the promoter activity is dependent on the extent of methylation. While methylation at a single site in the human rDNA promoter only weakens promoter activity, methylation of multiple sites in the UCE and the core promoter element are needed to completely abolish promoter activity. This finding stands in contrast with CpG methylation in murine rDNA promoters. There, a single DNA methylation at position -133 abrogates UBF binding and rDNA transcription [118]. Thus, the data clearly demonstrate that methylation density at the promoter plays an important role in repressing human rRNA transcription. It was also investigated whether the down-regulation of promoter activity was solely due to DNA methylation in the promoter or whether it is also influenced by methylation at specific sites in the coding region. Experiments demonstrate that CpGs within the coding region can be methylated in an all-or-nothing manner and that it is inversely correlated with transcriptional activity [18, 96]. However, when only one CpG residue in the external transcribed spacer 1 was methylated, no reduction in the promoter activity was observed.

As a summary, symmetric methylation of CpGs residing in the promoter led to transcriptional repression showing an inverse relationship between promoter activity and CpG methylation status. This is mediated by binding of MBD2 to methylated DNA, followed by histone modifications, finally leading to gene silencing [42].

The importance of rDNA promoter methylation is also shown in the pathological states of mild cognitive impairment and Alzheimer's disease [106]. The results indicate that rDNA hypermethylation in the parietal and prefrontal cortex is most strongly associated with early stages of Alzheimer's disease pathology. Reduced rRNA gene transcription leads to depletion of ribosomes, a signal for Alzheimer's disease and mild cognitive impairment.

### 1.2.2.3 Tip5 as a key regulator of rDNA silencing

Silencing of ribosomal DNA is not only dependent on DNA modification, but is in addition a complex interplay with different proteins such as histone modifiers and chromatin remodellers, finally leading to changes in nucleosome structure. One major player of rRNA gene silencing in human and mouse cells was shown to be the nucleolar remodelling complex NoRC which consists of two subunits. The large subunit of NoRC is Tip5 (TTF-I interacting protein) which interacts with the protein TTF-I and thereby, gets recruited to the rDNA promoter. Tip5 is a 205 kDa nucleolar protein and shares a number of functional domains with the largest subunits of the human ATP-dependent chromatin remodelling complexes such as ACF. All complexes contain the ISWI/SNF2H ATPase as the small subunit which is the catalytic core acting together with the respective large subunit [136]. NoRC is a multifunctional chromatin-dependent regulator of rRNA genes, which regulates nucleosome positioning, transcriptional repression, epigenetic silencing and replication timing.

### rRNA gene silencing via NoRC

NoRC physically interacts with TTF-I, DNA methyltransferases, histone deacetylases (HDACs) and histone methyltransferases (HMTs) thereby targeting enzymatic activities to the rDNA promoter to form heterochromatin (reviewed in [85]). Extensive analysis of the order of events revealed a hierarchical order and mutual dependence of chromatin remodelling, histone modification and DNA methylation that operate in a common mechanistic pathway for repressing transcription [119]. Initially, NoRC is recruited to rDNA by interaction of its large subunit, Tip5, with TTF-1 which is bound to the promoter-proximal transcription termination element  $T_0$ . Importantly, NoRC function requires the association of Tip5 with promoter-associated RNA (pRNA) that originates from a Pol I promoter located in the intergenic spacer 2 kb upstream of the pre-rRNA transcription start site [81]. The following scenario was proposed to describe the function of pRNA in rDNA silencing: The intergenic transcripts are of low abundance and usually do not accumulate *in vivo* because they are either degraded or processed into 150 to 250 nucleotide long RNAs that are protected by NoRC binding. Tip5 recognizes the secondary structure of pRNA, and the interaction of Tip5 with pRNA changes the structure of both pRNA and NoRC in an induced fit mechanism [82].

Subsequently, NoRC interacts with the Sin3 corepressor complex which contains HDAC1 and HDAC2 and leads to deacetylation of Histone H4 of nucleosomes at the rRNA genes. This deacetylation might act as a signal for SNF2H-mediated nucleosome remodelling. The remodelling shifts the promoter-bound nucleosome downstream of the transcription start site to a position which is unfavourable for

transcription complex formation. Furthermore, the remodelling might be required for chromatin opening thereby allowing CpG methylation by DNA methyltransferases at the rDNA promoter. This methylation pattern might also impair binding of transcriptional activators and thus, silences the rRNA gene. [119]. These modifications alter the nucleosomal surface which then might recruit other proteins such as HP1, finally leading to heterochromatin formation. Figure 1.6 summarizes the role of NoRC in rRNA gene silencing.

In an alternative model, pRNA is bound to the promoter-proximal transcription termination element  $T_0$  via formation of a triple helix, thereby recruiting Dnmt3b to methylate the promoter region and finally, leading to transcriptional silencing [121]. Future experiments will reveal the different roles that intergenic transcripts play in rRNA gene silencing.

As a summary, NoRC serves at least two functions: first, as a remodelling complex that alters the position of the nucleosome at the rDNA promoter and second, as a scaffold that co-ordinates the activities of several complexes that modify histones, methylate DNA and establish a "closed" chromatin state [48].

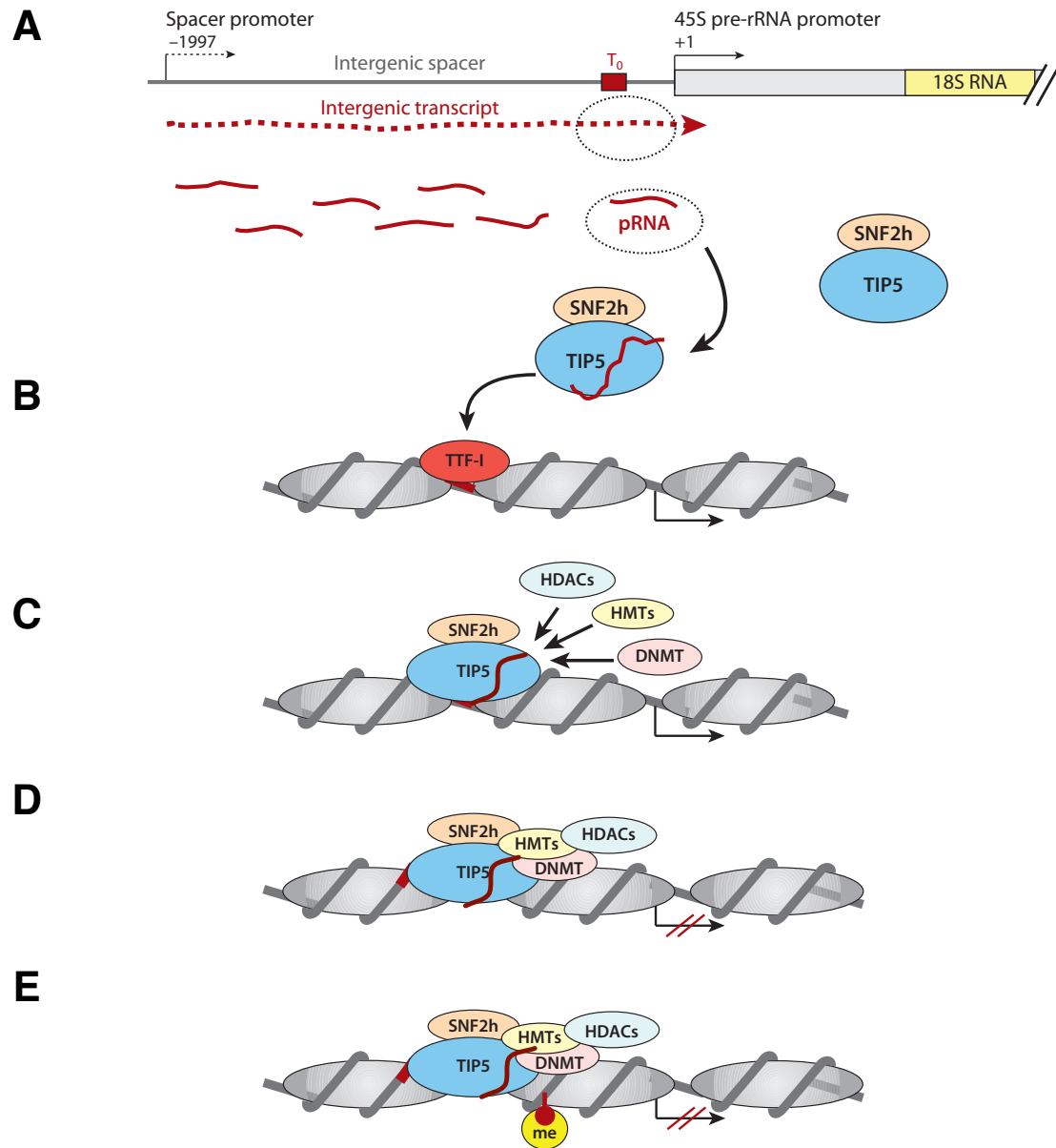
### Modular organisation of Tip5

In order to act as a scaffold in rDNA inactivation, Tip5 must consist of several protein - as well as nucleic acid binding domains. Indeed, conserved domains that are important for Tip5 function include a tandem plant homeodomain (PHD) finger/bromodomain at the C-terminal part of Tip5, several AT-hooks and the Tip5/ARBP/MBD (TAM) domain shown in Figure 1.7.

The PHD finger is a 50-80 amino acid zinc finger-like motif with a unique Cys4-His-Cys3 pattern that has been identified in many proteins implicated in chromatin-mediated transcriptional control [136]. The bromodomain, an approximately 100 amino acid sequence is found in many chromatin-associated proteins and interacts specifically with histone H4 acetylated lysine 16 (H4K16ac) thus, anchoring Tip5 to acetylated chromosomes (reviewed in [48]). This interaction is required for subsequent deacetylation of H4K5, H4K8 and H4K12. The PHD finger with the adjacent bromodomain is a co-operative unit that was shown to bind chromatin modifiers and to target them to gene promoters. In Tip5, the bromodomain co-operates with the PHD domain to recruit SNF2H and other chromatin-modifying enzymes to the rDNA promoter (reviewed in [48]).

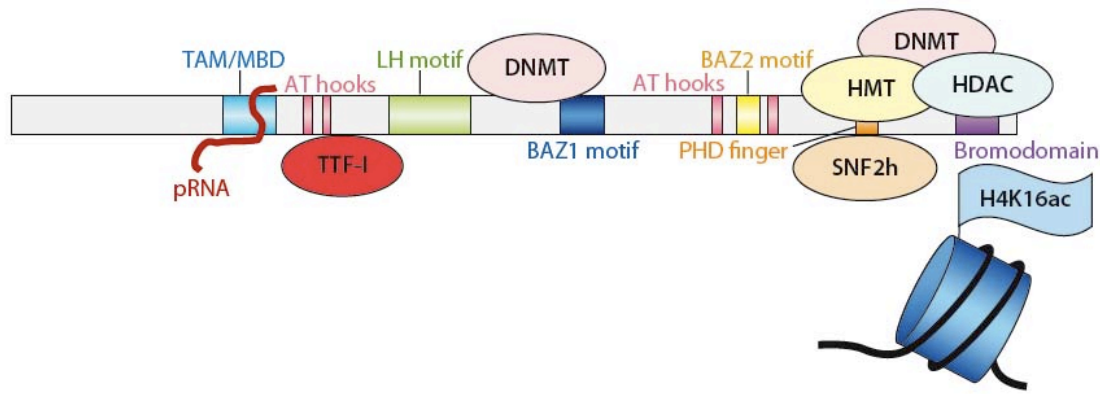
The TAM domain has sequence homology to the methyl-CpG binding domain (MBD) which is present in proteins that bind methylated DNA and are involved in gene silencing and heterochromatin formation. It was shown that the TAM domain is able to bind methylated DNA when fused to two AT-hooks [136]. Additionally, it was shown that the TAM domain can bind RNA and thus, mediates the interaction of NoRC with the pRNA [81].

In addition, the large subunit Tip5 harbours multiple AT-hooks that bind to the minor groove of DNA and co-regulate transcription by modifying the architecture of DNA, thus enhancing the accessibility of promoters to transcription factors.



**Figure 1.6: NoRC and its role in rDNA silencing**

NoRC function exemplarily shown for mouse rDNA. **(A)** Intergenic transcripts (dotted line) are synthesised from a spacer promoter located -1997 upstream of the transcription start site. Processing results in transcripts of 150-300 nt that match the rDNA promoter (pRNA) which bind Tip5 via its TAM domain. **(B)** NoRC is recruited to the promoter by TTF-I. **(C)** NoRC interacts with the Sin3 corepressor complex and HMTs leading to deacetylation of H3 and H4 and to methylation of H3K9, H3H20 and H3K27 respectively. **(D)** SNF2H may recognise these histone modifications and shift the promoter-bound nucleosome 25 nt further downstream. **(E)** The action of SNF2H alleviates DNA methylation by DNA methyltransferases (Dnmts) which impairs PIC assembly (pictures modified after [85]).



**Figure 1.7: Modular structure of Tip5 and its binding partners**

This scheme illustrates the modular organisation and localisation of sequence motifs in Tip5 that have been associated with functions in chromatin structure. The domains of Tip5 (coloured boxes) that interact with proteins involved in the epigenetic control of rDNA expression. The C-terminal part contains a PHD finger/ bromodomain that interacts with HDACs and H4K16ac. Dnmt1 interacts with the internal part of Tip5. The TAM domain is required for binding of small intergenic transcripts (pRNA). TTF-I binding and the four AT-hooks are also indicated. (Picture out of [85]).

### Large-scale organization of ribosomal DNA chromatin

NoRC is a chromatin-dependent regulator of rRNA transcription, which takes part in the balancing of this highly energy-demanding metabolic activity of the cell (reviewed in [85]). However, little is known about its role in large-scale spatial organization and distribution of actively transcribed and inactive rRNA gene copies. The synthesis of 47S pre-rRNA from active rDNA takes place at the fibrillar center/dense fibrillar component (FC/DFC border) of the mammalian nucleolus, whereas inactive rDNA is localized within the FC or outside of nucleoli (see review [129]). It has been demonstrated that changes in the ribosome synthesis activity result in alterations of nucleolar architecture when cells are treated with different inhibitors of ribosome biogenesis or serum starved [50],[45],[24]. Part of this morphological alteration in nucleolar structure may be correlated to rDNA chromatin movements, which may accompany changes in the transcriptional activity of rRNA genes.

In addition to the direct visualization of nuclear morphology by cell biology methods, biochemical characterization is also possible. DNaseI accessibility of genes correlates with their transcriptional activity. For example, nuclear matrix isolation uses extensive DNase I digestion and therefore, enables a simple biochemical characterization of large-scale chromatin. The nuclear matrix was originally defined as a component of nuclei that resists extensive DNase I digestion and salt extraction [7]. It contains mainly intermediate filament proteins like lamins, heterogeneous nuclear ribonucleoprotein particles, specific non-histone chromatin proteins and associated DNA, which represents the matrix-attachment regions (MARs) of the genome. MARs, which are supposed to anchor chromatin loops to the nuclear matrix constitutively or transiently, have been implicated in the regulation of gene expression and replication (reviewed in[100]).

Since the TAM domain and AT-hooks of NoRC are predicted MAR binders [2],

Tip5 may mediate the anchoring of rDNA to the nuclear matrix and thus, separate silenced rDNA repeats from active ones. Importantly, specific enrichment of rDNA in nuclear matrix preparations has been demonstrated by using biochemical [101] and cell biology methods [30]. To contribute to transcriptional repression by large-scale organization of rDNA chromatin, Tip5, however, must be capable of modular binding of DNA stretches, the MARs.

## 2 Results

### 2.1 Single molecule epigenomics by Dynamic molecular combing

In addition to the detection of genomic anomalies such as rearrangements, DMC may also be proven useful to detect epigenetic modifications such as DNA methylation or DNA hydroxymethylation when combined with an immunofluorescence step using specific antibodies.

Furthermore, nucleosomal and nucleosome-free DNA can be distinguished by adding the chemical psoralen to the procedure.

Establishing these variations of the Dynamic molecular combing method, which will be called "epi-combing" and "psoralen-combing", will greatly enhance the understanding of the epigenetic regulation of repetitive DNA at the single molecule level.

#### 2.1.1 Suitable conditions for *in vitro* combing of $\lambda$ DNA

To establish the methods epi-combing and psoralen-combing, the proof-of-principle experiments were performed under defined conditions by using  $\lambda$  DNA.

At first, different concentrations of  $\lambda$  DNA were combed on silanized coverslides to find the most suitable density of DNA fibers for conducting the *in vitro* experiments. Tested concentrations of  $\lambda$  DNA were 0.5 ng/ $\mu$ l and 1 ng/ $\mu$ L and fibers were stained using the fluorescent dye Yoyo-1 (Figure 2.1).

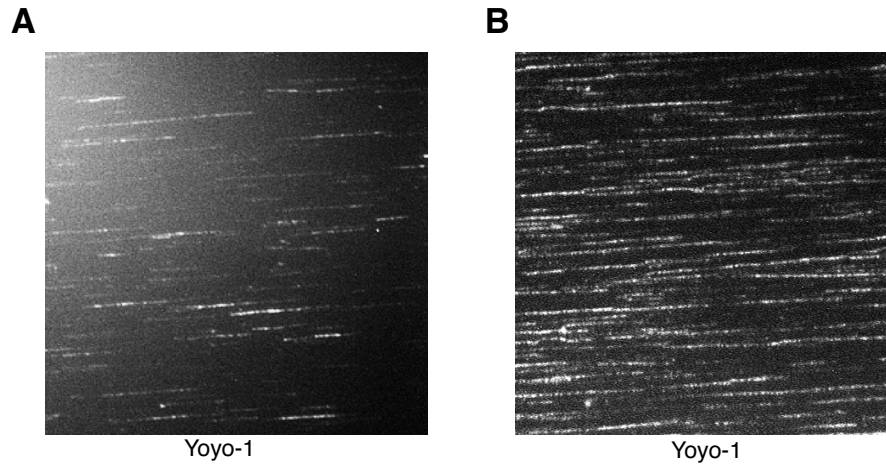
In both cases, single DNA fibers were aligned in parallel and no drilled superfibers were detected (Figure 2.1). Therefore, the higher DNA concentration was used for further experiments, since the density of signals is higher at 1 ng/ $\mu$ L  $\lambda$  DNA and thus, more signals can be analyzed in one field of view.

To have a locus-specific analysis in future experiments, it is crucial to establish fluorescence hybridization on  $\lambda$  DNA. Two probes were tested, one labelled with digoxigenin, the other one with biotin to make a variety of detection methods possible.

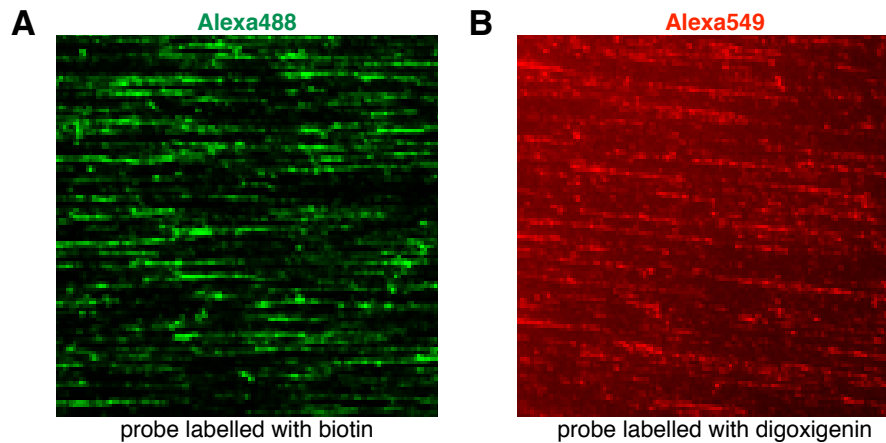
DNA fibers were detected in both cases (dUTP-biotin: Figure 2.2A, dUTP-digoxigenin: Figure 2.2B), thus making several detection methods possible depending on the combination of fluorescence dyes.

These experiments were the basis for future *in vitro* experiments using  $\lambda$  DNA as template.





**Figure 2.1: Different concentrations of  $\lambda$  DNA stained with Yoyo-1**  
**A:** 0.5 ng/ $\mu$ L  $\lambda$  DNA was combed and stained with the fluorescent dye Yoyo-1 **B:** Same as A, but illustrated are DNA signals resulting from double the concentration of  $\lambda$  DNA (1 ng/ $\mu$ L)



**Figure 2.2: Fluorescence hybridization of  $\lambda$  DNA using differently labelled probes**  
**A:**  $\lambda$  DNA is hybridized with a probe labelled with biotin and immunodetected by three layers of antibodies (StrA488 - Rb $\alpha$ -Str-biotin - StrA488) **B:**  $\lambda$  DNA is hybridized with a probe labelled with digoxigenin and immunodetected by two layers of antibodies (S $\alpha$ dig-Rho - G $\alpha$ S-A549)

### 2.1.2 Psoralen-combing

Psoralen-combing is based on the hypothesis that heavily crosslinked DNA is not accessible for fluorescence hybridization, thus a labelled probe will not hybridize to crosslinked DNA and no FISH signal will be visible.

This approach is a novel tool for the sequence-specific and genome-wide analysis of chromatin on single molecules.

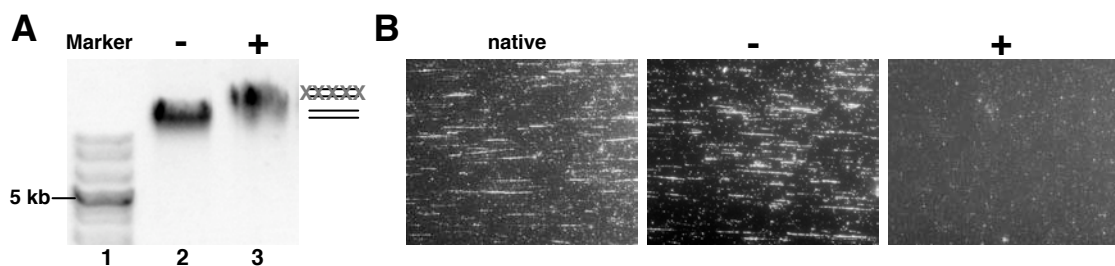
#### 2.1.2.1 *In vitro* psoralen-combing on $\lambda$ phage DNA

The proof-of-principle experiments have been performed on  $\lambda$  phage DNA, which was subjected first to psoralen treatment and photo-crosslinking with UV light. Subsequently, the DNA has been stretched on silanized coverslips and visualized with fluorescent hybridization either with or without reversal of the psoralen crosslinking at 254 nm UVB light (see Figure: 2.3 and 2.4, respectively).

#### Crosslinking of psoralen and DNA abolishes fluorescence hybridization

$\lambda$  DNA is irradiated four times 5 min each with fresh psoralen added before each irradiation (see 4.2.6.4). In addition, a mock control was included, which was treated solely with ethanol instead of psoralen.

The success of the crosslink between DNA and psoralen is controlled by agarose gel electrophoresis, where 400 ng of each sample was analyzed on a 0.7 % agarose gel and visualized by UV illumination (see Figure 2.3A)



**Figure 2.3: Psoralen-combing on  $\lambda$  DNA**

**A:** Heavily crosslinked DNA with psoralen (+ ; lane 3) shows a slower migrating band (s-band) compared to DNA that was treated solely with ethanol (mock control; - ; lane 2) (f-band) **B:** Fluorescence hybridization on combed  $\lambda$  DNA that was untreated (left panel), mock-treated (middle panel) or treated with psoralen (right panel). DNA fibers are detected solely in the native and mock treated DNA, not in the psoralen treated DNA.

The heavily crosslinked DNA could be clearly distinguished as a slower migrating band compared to the mock control. The crosslink between psoralen and DNA was successful, therefore, the DNA samples were combed and subsequently hybridized with a biotin-labelled probed for  $\lambda$  DNA. In addition to ethanol and psoralen treated DNA, native, untreated  $\lambda$  DNA was included as positive control. The signals were analyzed under a fluorescent microscope (see Figure 2.3B). Single DNA fibers were

detected in the untreated and the mock treated sample, whereas no fibers but un-specific background was visible in the psoralen-treated control.

Since no signals were detected on the slide of the psoralen-treated DNA, it seemed that the psoralen:DNA interaction abrogates hybridization detection as predicted. However, it is also possible that the psoralen-DNA interaction interferes with the attachment of the DNA fibers to the silanized glass surface resulting in no DNA fibers on the surface. To eliminate this possibility, the psoralen crosslink was reversed by UV radiation making the detection possible and proofing that combing of crosslinked DNA works in principal.

### **Reverse crosslink of psoralen:DNA adducts rescues fluorescence hybridization**

To reverse the crosslink of psoralen with DNA, the prepared microscopy slides were irradiated with UV light of 254 nm wave length [128] using a Stratalinker device. After adhering the cover slip onto the microscope slide, they were exposed to different amounts of energy to reveal the correct dose of UV light.

At low doses, the energy amount may not be sufficient to reverse the crosslink and the detection by fluorescence hybridization is still inhibited. When the energy amount is too high, the UV light will induce breaks in the double helix. This nicking can also result in thymine dimers, which interfere with hybridization. Therefore, the fluorescence hybridization will not yield signals of single DNA fibers, since the DNA on the slide got destroyed.

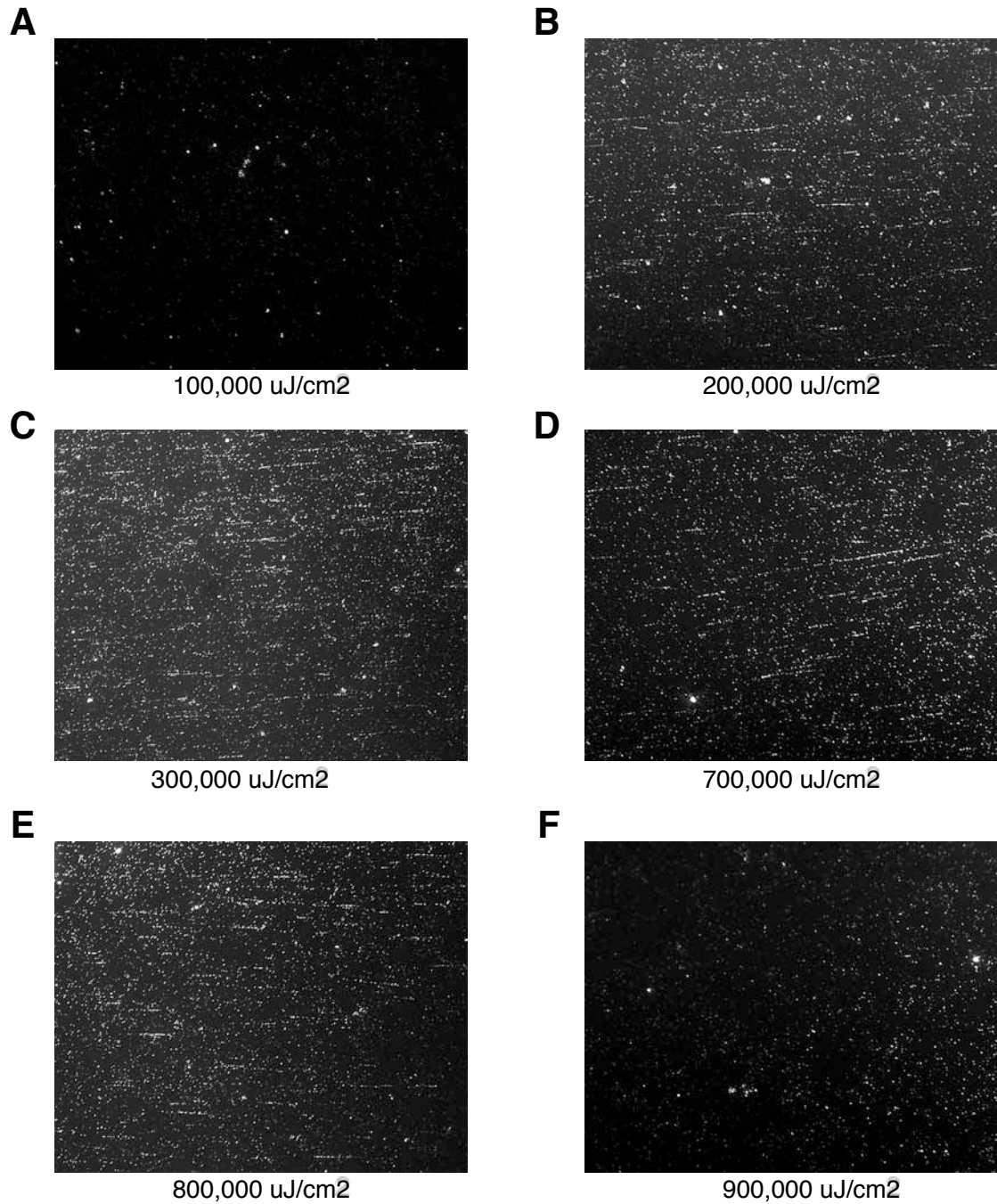
Typically, nicking of DNA begins at  $150,000 \mu\text{J}/\text{cm}^2$  [135], however reversing the crosslink requires also high energy doses. Hence, a broad range of energy doses ranging from  $100,000 \mu\text{J}/\text{cm}^2$  to  $900,000 \mu\text{J}/\text{cm}^2$  must be tested to find the correct dose for reversing the crosslink between DNA:psoralen adducts (see Figure 2.4).

After visualization of the probe, no hybridization signals appeared when energy amounts of  $100,000 \mu\text{J}/\text{cm}^2$  and  $900,000 \mu\text{J}/\text{cm}^2$  were used. In the first case, the amount of energy was not enough to reverse the crosslink and to make fluorescence hybridization possible. In the latter case ( $900,000 \mu\text{J}/\text{cm}^2$ ), the dose of energy was apparently too harsh and DNA was degraded, again yielding no fluorescent DNA signals. Interestingly, energy doses between  $200,000 \mu\text{J}/\text{cm}^2$  to  $800,000 \mu\text{J}/\text{cm}^2$  were all able to decross the covalent bond between psoralen and DNA, thus enabling the hybridization with  $\lambda$  DNA and the labelled probe. In these cases, DNA fibers could be detected on the slide while  $700,000 \mu\text{J}/\text{cm}^2$  resulted in the strongest signals of fluorescence (see Figure 2.4D).

This experiment gave evidence that the combing process between crosslinked DNA and the silanized glass surface is not inhibited. As a summary, the reversible psoralen-dependent masking of the signals demonstrates that the method can be potentially applied for the single molecule characterization of chromatin structures.

### **Psoralen-biotin enables direct visualization of crosslinked DNA**

The previous experiments have proven, that DNA crosslinked with psoralen can be combed and that the fluorescence hybridization is abrogated upon crosslinking with psoralen. However, this method only enables the indirect visualization of



**Figure 2.4: Reverse crosslink of psoralen treated  $\lambda$  DNA**

**A - F:** Psoralen treated  $\lambda$  DNA was combed, fixed and irradiated with different amounts of energy at 254 nm wavelength **A** : Energy dose ( $100,000 \mu\text{J}/\text{cm}^2$ ) is not sufficient to reverse the DNA:psoralen crosslink. **B - E:** Different energy amounts that all enable decrosslinking and subsequent detection by fluorescent hybridization **F:** The high energy dose ( $900,000 \mu\text{J}/\text{cm}^2$ ) led to degradation of DNA, thus inhibiting detection of DNA fibers.

nucleosome-depleted DNA by showing no signal. A better solution would be the direct visualization of nucleosome-depleted DNA by using labelled psoralen. Thus, the method psoralen-combing was established on  $\lambda$  DNA using psoralen which is biotinylated (psoralen-biotin = PB), (see Figure 2.5). Therefore, the crosslink of psoralen-biotin with  $\lambda$  DNA was repeated as described earlier (see first paragraph 2.1.2.1). However, irradiation procedure was reduced from 4x5min to 2x5min to protect DNA from damage.

The crosslink with PB also worked with the reduced irradiation time of 2x5min, since a slow migrating (PB-treated) and a fast migrating (ethanol-treated) band can be distinguished in an agarose gel electrophoresis (Figure 2.5A). When detecting PB (see 4.2.6.7 starting with DNA detection), only background signals as expected were detected in the mock control. However, when the PB-treated DNA was analyzed, only strong background instead of single fibers could be detected (Figure 2.5B). In theory, detection with the unspecific DNA stain Yoyo-1 should not show any difference between combed mock control or PB-treated DNA, since same concentration of DNA in both cases were combed. While the staining of the combed mock control displayed single fibers stretched and aligned in parallel, the PB-treated sample showed few fibers (Figure 2.5C).

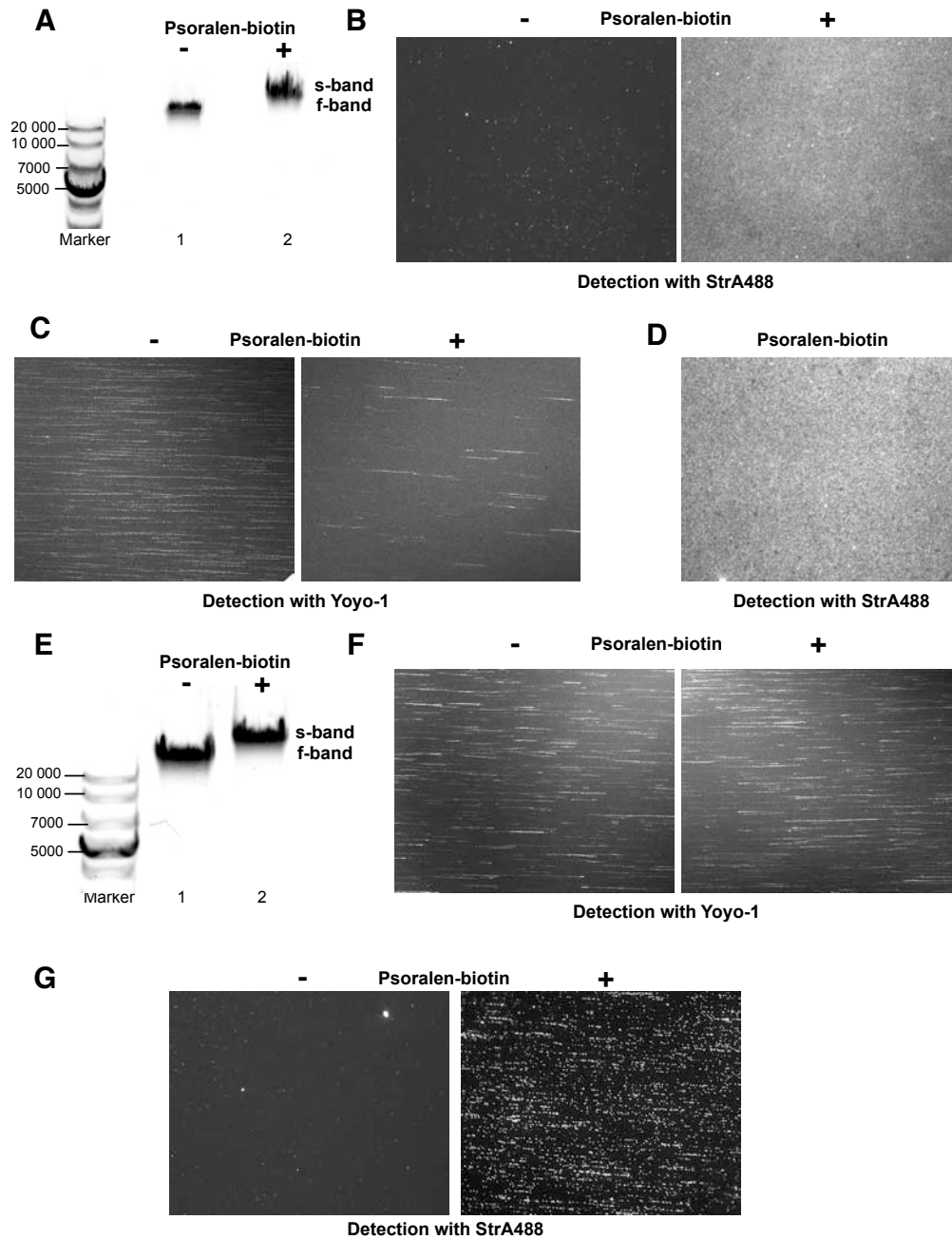
One possibility is that the chemical psoralen-biotin itself binds to silanized surfaces and hence, yields the high background. To test this, a solution without DNA but with a comparable amount of PB was combed (Figure 2.5D). The picture shows the same background as in the panel of the PB-treated DNA (Figure 2.5B, right). This experiment proved that PB itself interferes with DNA detection by unspecific binding to the silanized glass surface. This explains also why the detection with Yoyo-1 did not work since the DNA was not accessible to the dye.

This means that the remaining reagent must be removed from the solution after irradiation and cross-link which was done by DNA precipitation. Since the precipitation protocol might disturb the crosslink of the PB:DNA adduct, it was checked by agarose gel electrophoresis whether the s- and f-bands can be still distinguished. After precipitation, the two distinct bands were separated, demonstrating that the crosslink between PB and DNA was not destroyed (Figure 2.5E). Combing these samples and staining them with Yoyo-1 leads to the detection of single DNA fibers in both cases, showing no difference between the mock and the PB-treated sample (Figure 2.5F). When the purified DNA was combed, background is found in the mock control, whereas specific signals of single DNA fibers are successfully displayed in the PB-treated sample.

These experiments demonstrate that the protocol must be adapted for psoralen-biotin by purification of the DNA, otherwise, the chemical PB itself will interfere with the detection.

Figure 2.3 shows that fluorescence hybridization is abrogated at DNA crosslinked with psoralen. In this experiment, the DNA was irradiated 4x5min whereas in the experiment with psoralen-biotin, DNA was irradiated 2x5min. The direct visualization of psoralen-biotin is demonstrated, however it is possible that 2x5min irradiation is not sufficient to interfere with fluorescence hybridization.

Hence, irradiation times of 2x5min and 4x5min must be compared. After irradiating diluted DNA, ethanol-treated and PB-crosslinked DNA samples were purified and the DNA was combed. After visualization of PB (see 4.2.6.7 starting with



**Figure 2.5: Detection of psoralen-biotin on combed  $\lambda$  DNA**

**A:** Agarose gel electrophoresis showing the faster migrating ethanol-treated mock control (f-band) and the heavily crosslinked psoralen-treated DNA (s-band). Shown are 400ng of DNA each. **B:** Detection of psoralen-biotin on combed  $\lambda$  DNA with StrA488. In both cases, background is detected. **C:** Detection of psoralen-biotin on combed  $\lambda$  DNA by Yoyo-1. While single DNA fibers can be detected in the mock control, only few fibers are detected in the PB-treated DNA. **D:** The chemical psoralen-biotin binds to the silanized surface and creates high background. **E:** After DNA purification, the faster migrating ethanol-treated mock control (f-band) and the heavily crosslinked psoralen-treated DNA (s-band) are still distinct in agarose gel electrophoresis. Shown are 400ng of DNA each. **F:** After removal of PB by DNA purification, mock control and crosslinked DNA show the same pattern of Yoyo-1 staining. **G:** Purified DNA was combed and detected with StrA488. Unspecific background is detected in the mock control, whereas DNA fibers resulting from crosslink with psoralen biotin are found in the right panel.

DNA detection), the slides were crosslinked with 4% paraformaldehyde/1xPBS to fix the antibodies. Then, hybridization mixture that is labelled with digoxigenin was added and the slides were hybridized and finally, the DNA was detected according to the protocol described in subsection 4.2.6.7. This enables the parallel detection of psoralen-biotin and DNA, hence, the direct comparison of different irradiation times (see Figure 2.6).

Figure 2.6 demonstrates the dependence of fluorescence hybridization success on different incubation times with psoralen-biotin. In the case of shortened incubation (2x5min), PB and the hybridization signal can be detected in parallel (Figure 2.6 A), whereas only fluorescence hybridization signals and not PB signals are found in the mock control. Apparently, crosslink of PB with  $\lambda$  DNA for 2x5 min does not abolish hybridization.

In contrast, longer irradiation and incubation times with PB interferes with hybridization that no DNA signal are detected (2.6B). The mock control displays so signals of PB, however fluorescent hybridization signals yield single DNA fibers. The PB-treated sample (4x5min irradiation) again show PB detected by StrA488, though longer irradiation times inhibit detection of DNA fibers by fluorescence hybridization.

These findings show that longer incubation and irradiation with PB leads to more efficient crosslinking between PB and DNA which finally abrogates fluorescence hybridization and thus, detection of hybridization signals.

### 2.1.2.2 Psoralen-combing on genomic DNA

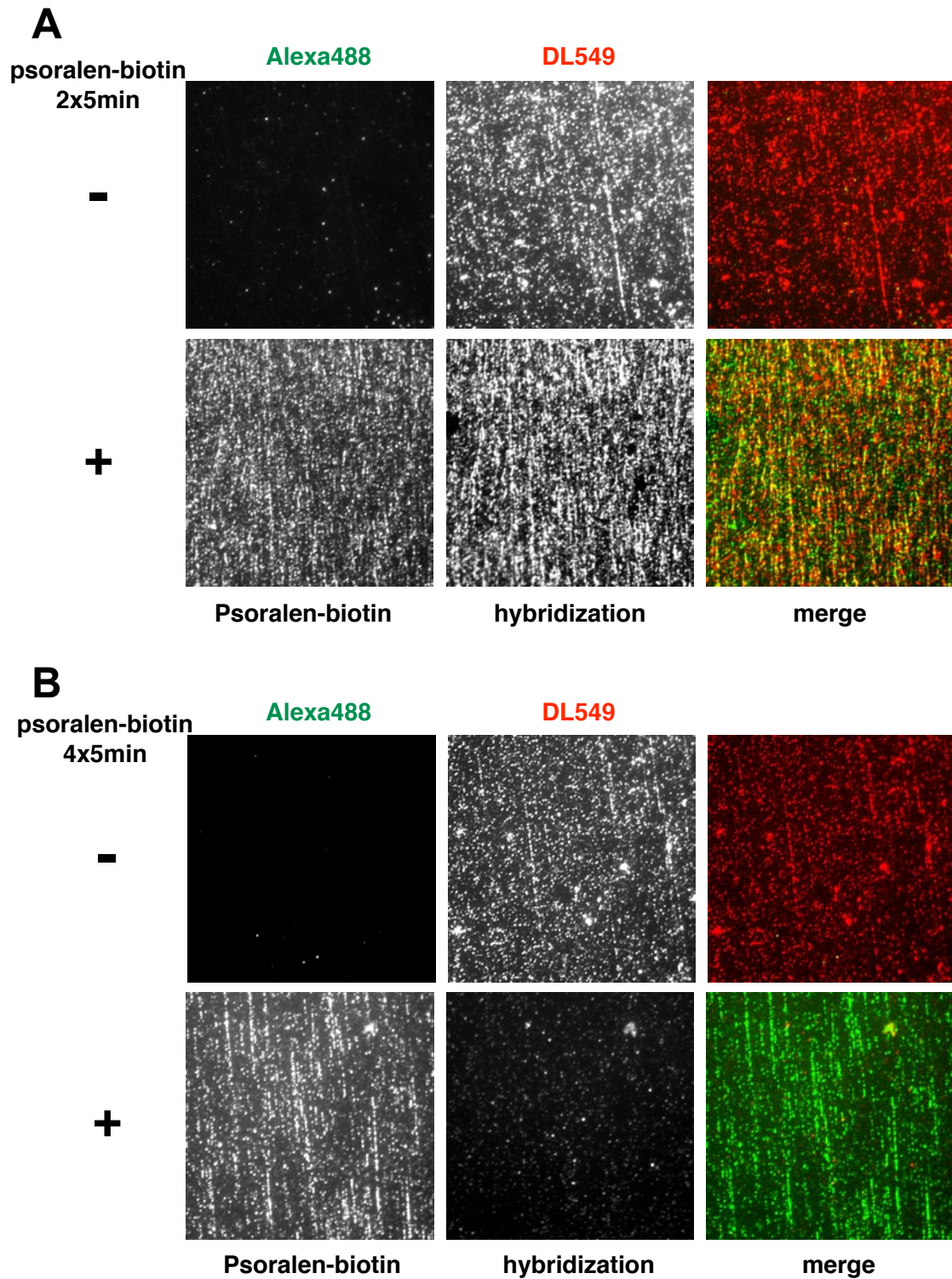
Since preparation, combing process and analysis on  $\lambda$  DNA were established, the principle was tested on genomic DNA. In the *in vitro* experiments,  $\lambda$  DNA was in solution and psoralen could easily intercalate into the double helix. To determine the extent of psoralen intercalation in genomic DNA embedded in an agarose plug, different incubation and irradiation times need to be tested. Under defined conditions, namely naked DNA isolated in a low melting agarose plug, the intercalation of psoralen in genomic DNA was examined.

#### Direct visualization of psoralen-biotin on naked genomic DNA

Nucleosome-free, high molecular weight DNA was stored at 4°C in agarose plugs and treated with different amounts and irradiation times with psoralen according to 4.2.6.4. DNA was combed and after PB was visualized, the signals were analyzed under a fluorescent microscope (see Figure 2.7).

Irradiation of 5 min did not completely crosslink the whole DNA fiber, since gaps in the signal were apparent (see Figure 2.7A). Irradiation of 10 min (Figure 2.7B) or 15 min (Figure 2.7C) in total show a continuous DNA fiber resulting from covalent interaction between psoralen-biotin and naked DNA. By this, it was proven that psoralen can easily intercalate in low melting agarose plugs and it will not present a major problem in future experiments. Noteworthy, the DNA concentration in the



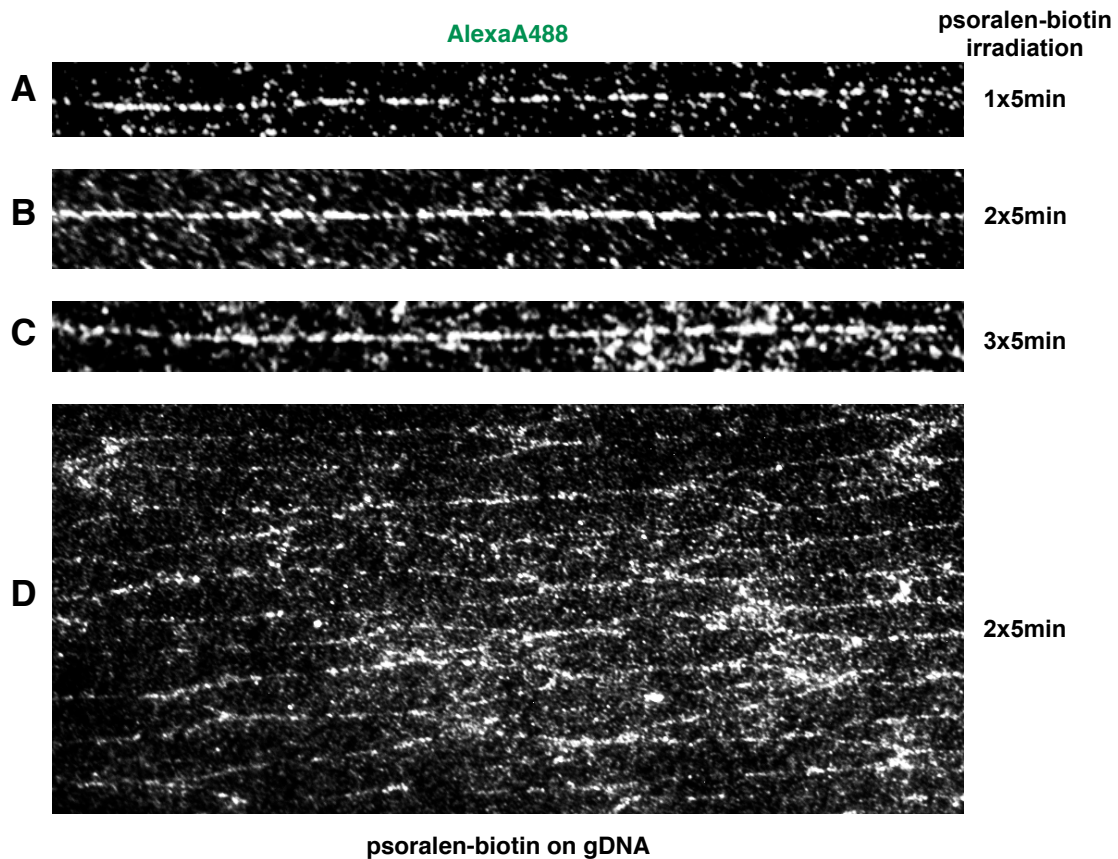


**Figure 2.6: Comparison of different crosslinking times of psoralen-biotin to  $\lambda$  DNA and subsequent detection of PB and DNA**

**A:** Purified as well as combed mock (-) and PB-treated (+)  $\lambda$  DNA irradiated 2x5min. PB is detected in the case of the PB sample, however, the degree of crosslink was not sufficient to abrogate DNA hybridization since DNA fibers can be detected. In the case of mock treated DNA, low background of unspecific StrA488 binding and the specific detection of the digoxigenin probe, resulting in the visualization of DNA fibers, is demonstrated.

**B:** Same as A, but samples are irradiated 4x5min for a more efficient crosslink. Again, PB is detected in the PB-treated sample but not in the mock control. This time, the fluorescence hybridization worked in the case of the mock control, but no DNA fibers are visible in the PB-treated sample.





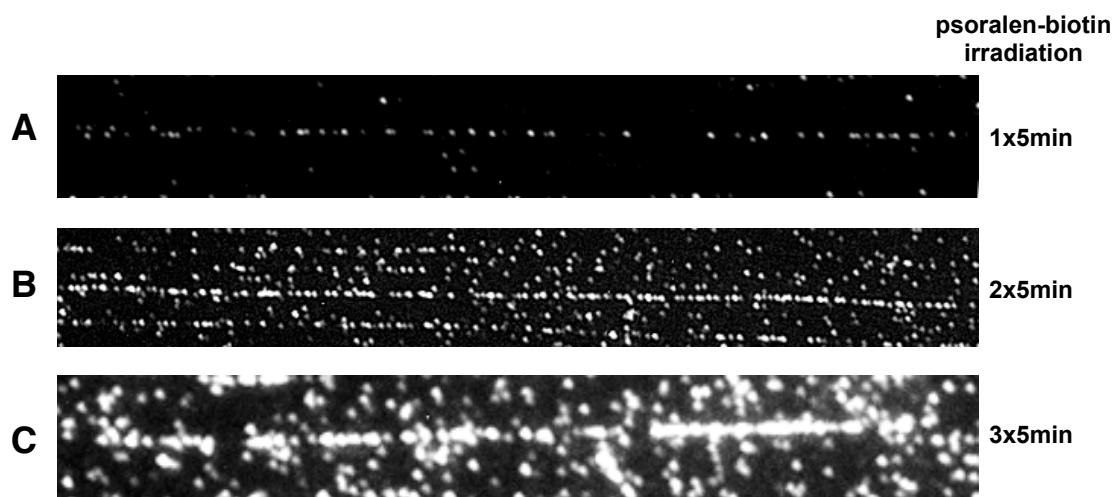
**Figure 2.7: Psoralen-biotin combing on naked genomic DNA**

Genomic DNA in a low melting agarose plug is incubated with PB for 10min at RT and subsequently irradiated by UVA light for the times indicated. Samples are combed and psoralen-biotin crosslinked DNA is visualized by addition of fluorescently coupled Streptavidin. **A:** After incubation, DNA is irradiated for 5 min by UVA light **B:** Same as (**A:**), but with another addition of 1  $\mu\text{g}$  psoralen-biotin followed by irradiation for 5 min (2x5min total) **C:** Same as (**B:**), but with another addition of 1  $\mu\text{g}$  psoralen-biotin followed by irradiation for 5 min (3x5min total). **D:** High concentration of DNA fibers lead to crosslink between different fibers which results in a network of fibers; Sample treatment as in (**B:**)

plug should be kept low since high local density of DNA fibers and psoralen may lead to the formation of crosslinked networks (Figure 2.7D).

### Cell lysis with 1% sarcosyl-EDTA solution results in disruption of nucleosomal DNA

Since the method psoralen-combing was established on naked genomic DNA, native chromatin from cells was to be tested for further analysis. HCT116 cells were harvested, embedded in agarose plugs and irradiated as described in 4.2.6.4 and DNA was isolated as described. Three different irradiation times were tested as indicated, combed and analyzed (see 4.2.6.7 starting with DNA detection) as can be seen in Figure 2.8.



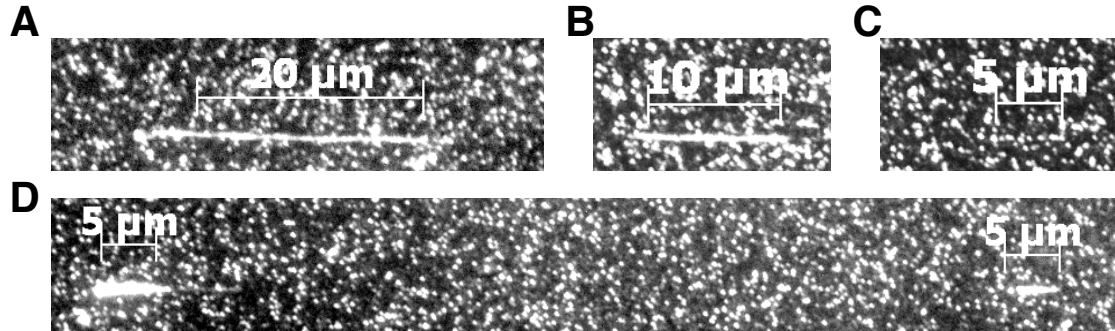
**Figure 2.8: Psoralen-biotin combed on genomic DNA after cell lysis by 1% Sarcosyl**

Before the plug was irradiated for the times indicated, HCT116 cells were embedded in a low melting agarose plug and lysed for 30min. Samples are then combed and psoralen-biotin crosslinked DNA is visualized by addition of fluorescently coupled Streptavidin. **A:** After incubation for 10min with psoralen-biotin, DNA is irradiated for 5 min by UVA light **B:** Same as (**A:**), but with another addition of 1  $\mu$ g psoralen-biotin followed by irradiation for 5 min (2x5min total) **C:** Same as (**B:**), but with another addition of 1  $\mu$ g psoralen-biotin followed by irradiation for 5 min (3x5min total).

The lowest irradiation time results in few signals of psoralen-biotin (Figure 2.8A) whereas 2x5min and 3x5min yield strong signals of psoralen-biotin with few gaps in the signal (Figure 2.8B and C, respectively). Since psoralen-biotin indicates non-nucleosomal DNA, long stretches of nucleosome-depleted DNA were detected. This means that 1% of the detergent sarcosyl already disrupted the native chromatin structure, hence, the combed protocol must be adapted for psoralen-combing and a protocol for mild cell lysis protecting nucleosomal DNA is searched for.

### Cell lysis with 0.5% NP-40 retains nucleosomal DNA and enables visualization of nucleosome-depleted DNA with psoralen-biotin

A mild detergent that keeps nucleosomal DNA intact is NP40 at a concentration of 0.5% (v/v). Therefore, the experiment is repeated with a buffer containing 50mM Tris and 0.5% of the detergent according to the protocol described in 4.2.6.4. After the DNA was isolated, combed and PB was visualized, the signals were analyzed under a fluorescence microscope (Figure 2.9).



**Figure 2.9: Psoralen-biotin combing on genomic DNA after cell lysis by 0.5% NP40**

Before the plug was irradiated for 4x5min, HCT116 cells were embedded in a low melting agarose plug and lysed for 30min in buffer containing 0.5% NP40. Samples are then combed and psoralen-biotin crosslinked DNA is visualized by addition of fluorescently coupled Streptavidin. **A: - D:** Displayed are different examples of psoralen-biotin signals with different lengths ranging from 4μm to 25 μm, representing at least 8 kb to 50 kb, respectively.

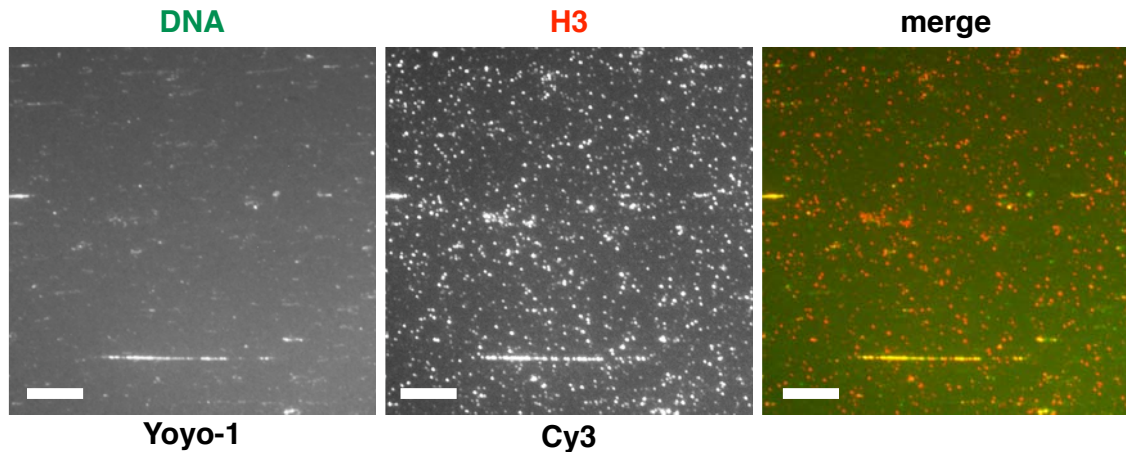
Different signals of psoralen-biotin can be found that display specific signals of psoralen-biotin intercalation in DNA, thus nucleosome-depleted genomic DNA. Signals range from a few kilobases in length up to 50 kb of nucleosome-depleted DNA. Interestingly, two short signals of about 10 kb in length are found, which may be on the same fiber. However, to manifest this observation, detection of DNA fibers with Yoyo-1 is necessary, which shows emission at the same spectra as Alexa488, though.

As summary, cell lysis by 0.5% NP40 is suitable to preserve nucleosomal structure of DNA, which subsequently can be analyzed by psoralen-combing. For a locus-specific analysis of psoralen-combing, a fluorescence hybridization step should be included after detection of psoralen-biotin.

### 2.1.3 Chromatin-combing

Psoralen-combing is a method that enables the direct visualization of nucleosome-depleted DNA by intercalation of psoralen-biotin into that regions and detection by layers of fluorescently antibodies. Combing native, unmodified chromatin fibers would make feasible the direct visualization of nucleosomal DNA, together with other proteins bound to DNA such as transcription factors.

HCT116 were fixed with formaldehyde and the combing protocol was adapted as described in 4.2.6.5 (100,000 cells per plug). Chromatin was detected by immunofluorescence of the histone H3 and DNA was unspecifically stained with Yoyo-1. The corresponding signals were visualized under a fluorescence microscope and can be seen in Figure 2.10.



**Figure 2.10: Combing of native chromatin fibers**

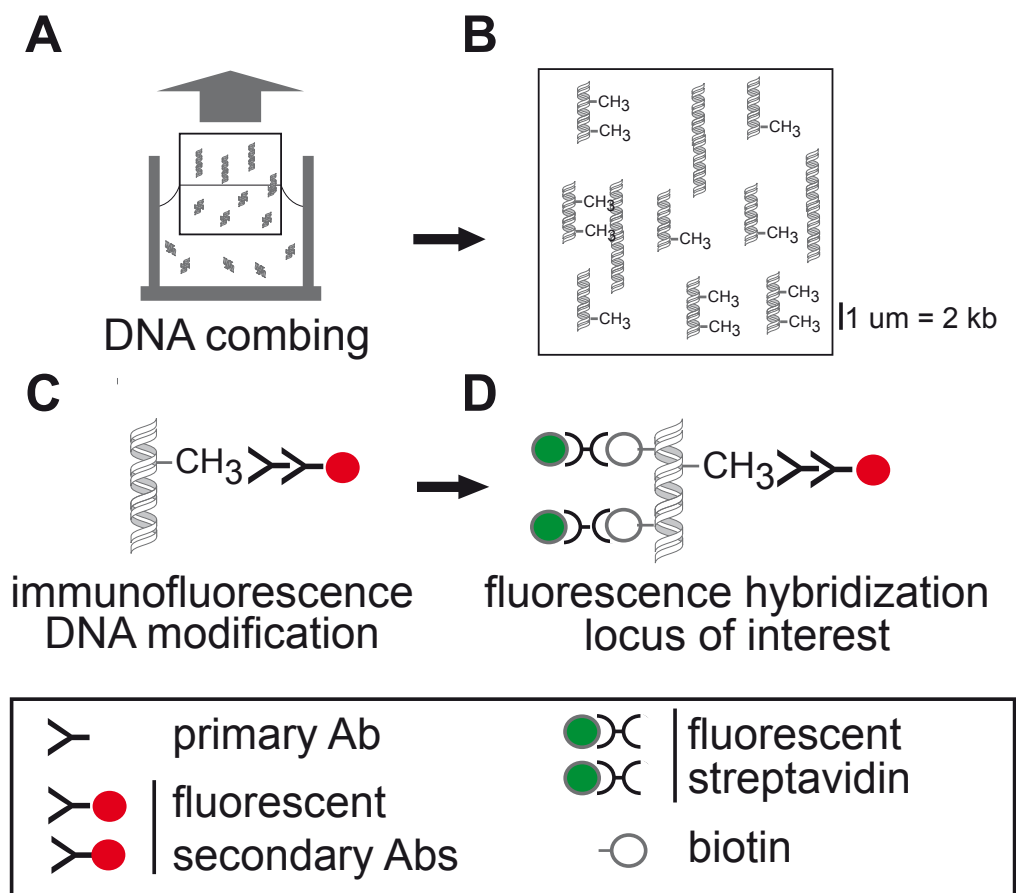
Chromatin is visualized by immunodetection of histone H3 and Cy3-coupled antibodies (middle) whereas DNA is detected by staining with Yoyo-1 (left). Merge (right) illustrates the colocalization of both signals, demonstrating intact chromatin fibers. Scale bar: 10  $\mu\text{m}$

Figure 2.10 shows DNA signals stained with Yoyo-1 (left panel), the histone H3 (middle) and the merge picture (right). DNA and H3 signals colocalize which demonstrates that native chromatin was combed. However, DNA fibers are shorter than in standard combing experiments.

These initial experiments build the basis for chromatin-combing which enables the parallel detection of nucleosomes, nucleosome variants and other DNA bound proteins such as transcription factors to reveal chromatin signatures on single molecules.

### 2.1.4 Epi-combing

The Dynamic molecular combing method can also be adapted to the analysis of DNA modifications such as DNA methylation and DNA hydroxymethylation. In principle, high-molecular weight DNA is isolated according to the standard Dynamic molecular combing protocol (as described in 4.2.6.1). After DNA is combed and fixed on a silanized coverslip, the DNA modification is detected by specific antibodies. To enable a sequence-specific investigation, fluorescence hybridization is performed as final step of the epi-combing method (see Figure 2.11).



**Figure 2.11: Principle of epi-combing**

**A:** High-molecular weight DNA is isolated and combed on a silanized coverslip. The receding air-water meniscus aligns and stretches DNA fibers **B:** Upon combing, single DNA molecules are aligned and stretched in parallel on a coverslip. **C:** To analyze DNA modifications, an immunofluorescence step with specific antibodies is included. **D:** For locus-specific analysis, fluorescence hybridization is performed.

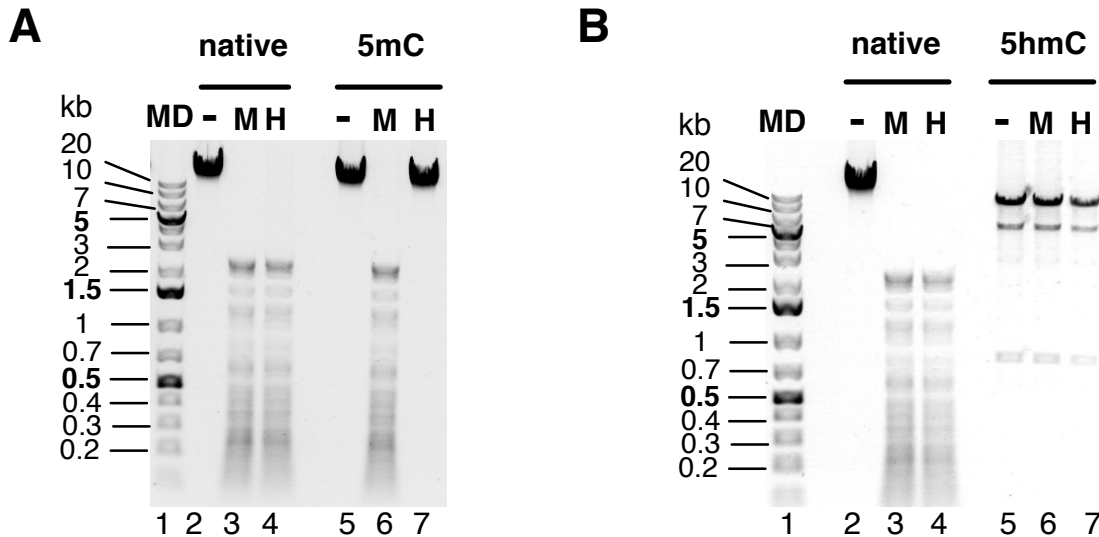


### 2.1.4.1 *In vitro* epi-combing on $\lambda$ phage DNA

The method epi-combing needs to be established under defined conditions, therefore, it was first arranged using  $\lambda$  DNA that is either unmethylated or *in vitro* modified. By incubation of 8.25  $\mu\text{g}$  of  $\lambda$  DNA with 8 U M.SssI as well as 100 nmol SAM over night at 37°C,  $\lambda$  phage DNA gets methylated. The next day, fresh 4U of M.SssI as well as 100 nmol SAM is added and incubated over night which is repeated until the methylation reaction is complete.

Hydroxymethylated  $\lambda$  DNA was obtained by a specific PCR described in 4.2.1.4, which yields a 12.5 kb fragment with hydroxymethylated cytosines.

To check the success of the DNA modification reactions, the samples were digested with MspI or HpaII, two isoschizomers, and analyzed by agarose gel electrophoresis. The enzyme HpaII is DNA modification-sensitive, thus does not cut when the recognition site CCGG is methylated or hydroxymethylated. In contrast, MspI is methylation-insensitive and will cut methylated DNA, however, hydroxymethylated DNA will not be digested by MspI either (see Figure 2.12).



**Figure 2.12: Control digestion of modified  $\lambda$  DNA with MspI and HpaII** Unmodified lambda DNA is cut by both enzymes. Methylated lambda DNA is cut by MspI, however not digested with HpaII. The hydroxymethylated PCR fragment is not digested by any of the two enzymes

### Epi-combing detection system for analysis of DNA methylation is sensitive, selective and specific

The DNA mark 5mC will be detected by a monoclonal antibody (33D3) and by the methyl-binding domain MBD2 that is fused to a human Fc antibody fragment (MBD2-Fc). To test if those proteins are sensitive, they were applied on methylated  $\lambda$  DNA ( $^m\lambda$ ) and checked if all fibers will be recognized. The application of the proteins to hydroxymethylated  $\lambda$  DNA ( $^{hm}\lambda$ ) and unmodified  $\lambda$  DNA will prove if the immunodetection is also specific. To check if the proteins work also selective, a 1:1:1 mixture of  $\lambda$ ,  $^{hm}\lambda$  and  $^m\lambda$  is combed.

1ng/ $\mu$ L of  $\lambda$  DNA is combed and detected according to the protocols described in 4.2.6.6, 4.2.6.7 and 4.2.6.3. DNA was detected using a biotinylated probe according to 4.2.6.7, 5mC was detected either by using MBD2-Fc or by 33D3 with corresponding secondary antibodies.

MBD2-Fc does not detect unmodified or hydroxymethylated DNA and is therefore specific. However, it does not recognize all DNA fibers that are present in the field of view and at this used concentration is not sensitive enough for further experiments. (see Figure 2.13A). In contrast, 33D3 is sensitive since it recognized all DNA fibers. In addition, it is also specific since no signals can be found at  $\lambda$  and  $^{hm}\lambda$  DNA. The fact that the antibody is selective is proven in the case of a 1:1:1 mixture of  $\lambda$ ,  $^{hm}\lambda$  and  $^m\lambda$  DNA, since it recognizes only methylated DNA and not all fibers. Therefore, the monoclonal antibody will be further used.

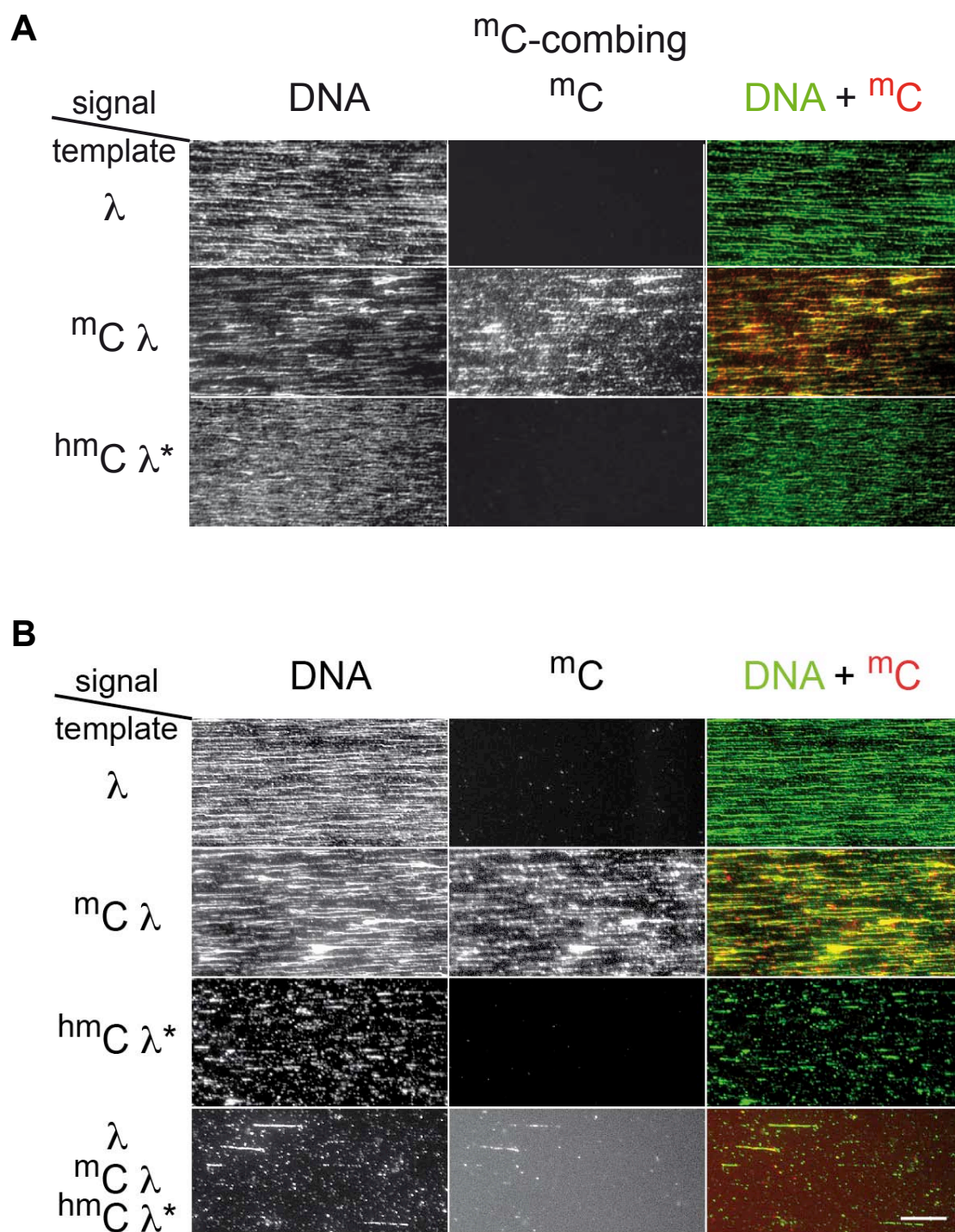
To enable a variety of colours for detection, the detection of 5mC with the aminomethyl coumarin acetate (AMCA) blue-fluorescent dye was tested. This will be useful for a three colour detection system, when DNA patterns should be detected in two colours and the third colour is needed for DNA modification detection. To achieve this,  $^m\lambda$  DNA at a concentration of 1ng/ $\mu$ L was combed and detected according to 4.2.6.6, 4.2.6.7 and 4.2.6.3. However, instead of antibodies coupled to Cy3, a secondary antibody coupled to Alexa350 (A350) and a tertiary antibody coupled to AMCA was used.

Figure 2.14 shows that the detection with other fluorescent dyes is possible. However, Cy3-coupled antibodies are more effective since they require lower concentrations. Therefore, this detection method should only be used where three colour detection is desired.

### **Epi-combing detection system for analysis of DNA hydroxymethylation is sensitive, selective and specific**

Two different commercially available antibodies against  $^{hm}\lambda$  DNA are to be tested, Rb- $\alpha$ -5hmC (Diagenode) and R- $\alpha$ -5hmC (Active motif). In principle, the experiment for sensitivity, specificity and selectivity was repeated as described above with R- $\alpha$ -5hmC (Diagenode) and Rb- $\alpha$ -5hmC (Active motif) and corresponding secondary and tertiary antibodies coupled with Cy3. DNA was again detected by fluorescence hybridization.

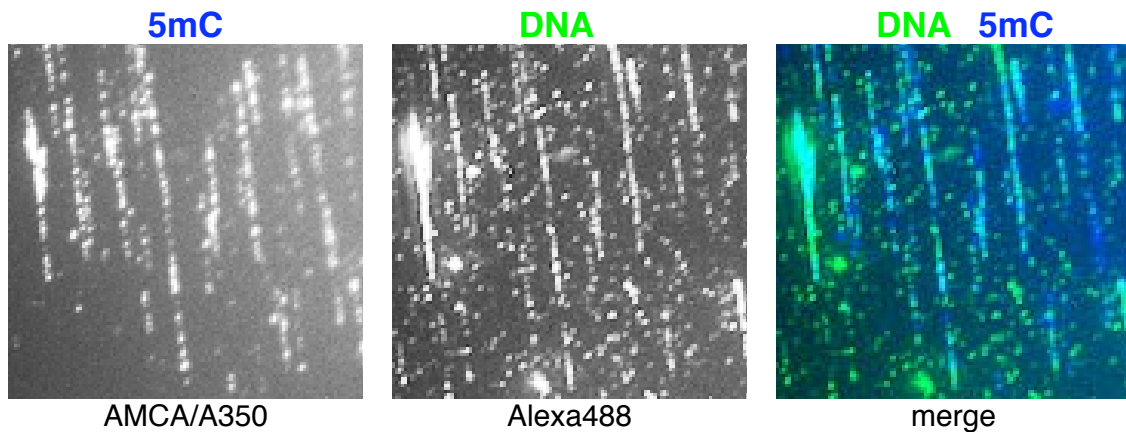
Both antibodies show that they are specific and sensitive, since they do only detect  $^{hm}\lambda$  DNA, and these fibers to 100 % (see Figure 2.15). However, Rb- $\alpha$ -5hmC gave in several experiments more reliable results and therefore, this one is tested for selectivity. Since it does only recognize 12.5 kb short fragments of DNA, the antibody is selective and will be used for future experiments.



**Figure 2.13: *in vitro* 5mC-combing on  $\lambda$  phage DNA**

**A:** MBD2-Fc (1:100 dilution in BlockAid) does not recognize  $\lambda$  and <sup>hm</sup> $\lambda$  DNA. However, not all DNA fibers are detected at this concentration. **B:** In contrast, 33D3 monoclonal antibody (1:50 dilution in BlockAid) recognizes all DNA fibers in the field of view. Furthermore, it does not detect  $\lambda$  and <sup>hm</sup> $\lambda$  DNA as well as methylated signals in the mixture.





**Figure 2.14: Detection of DNA methylation on  $\lambda$  phage DNA using a AMCA-coupled tertiary antibody**

Left panel: Detection of 5mC using M- $\alpha$ -5mC (1:50) and a secondary antibody coupled to A350 and a third layer of antibody coupled to AMCA. middle panel: DNA detection by fluorescence hybridization using biotin-labelled hybridization mixture and StrA488 detection system right panel: merge shows colocalization of signals

#### 2.1.4.2 Epi-combing on genomic DNA

##### Isolation of high molecular weight genomic DNA

To analyze DNA modifications at repetitive DNA and to answer the question about epigenetic linkage, a way to isolate high molecular weight DNA where single DNA fibers are not twisted is established in the lab (see 4.2.6.2, 4.2.6.1).

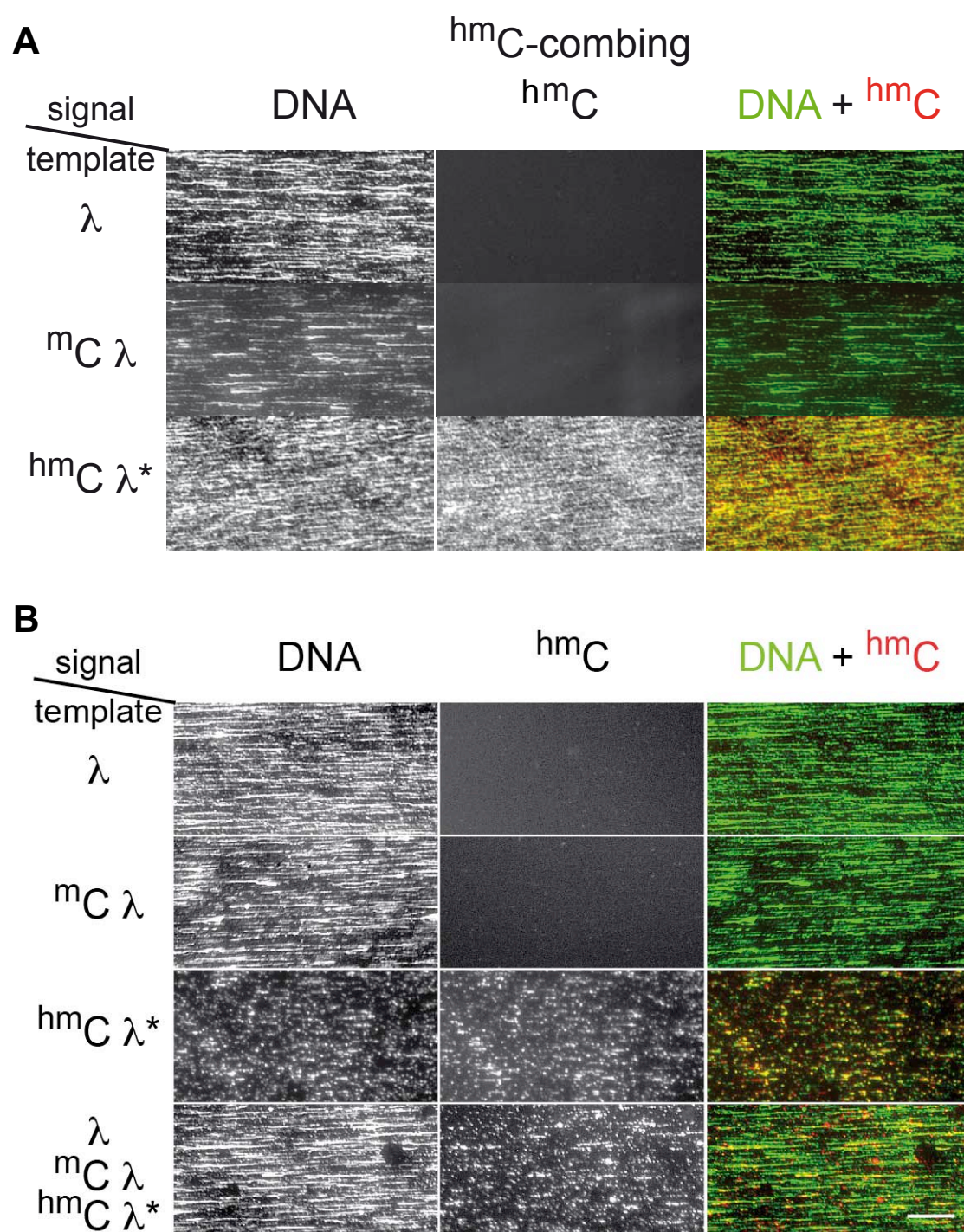
After combing and fixation of the DNA, the quality of the DNA preparation can be checked by staining with Yoyo-1 (see Figure 2.16).

Figure 2.16 displays two different DNA preparations from Imr90 and HCT116, respectively. Imr90 DNA was not stretched as long single fibers but as bundles of twisted fibers, that can be seen as superfibers in Figure 2.16A. More precaution during the preparation process must be taken that the plugs are not shaken or incubated in water baths, since this may cause bundle formation. Figure 2.16 B shows a successful combing of high molecular weight DNA with more than 100 kb in length. In addition, fibers of DNA are aligned in parallel that they can be detected as single fibers.

##### Non-locus specific detection of DNA methylation

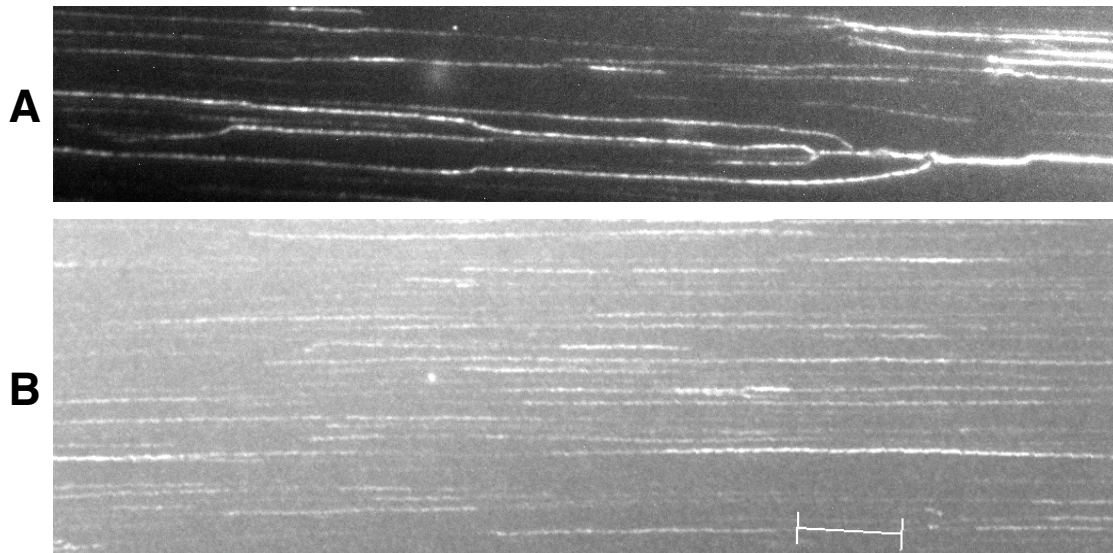
After the establishing of genomic DNA isolation, DNA methylation can be detected as a first step of epi-combing in genomic DNA. DNA is isolated and combed according to 4.2.6.2, 4.2.6.1, 4.2.6.3 and DNA methylation is detected as described in 4.2.6.6.

Figure 2.17 displays different signals of DNA methylation on human genomic DNA. Signals can be distinguished between molecules that have weak DNA methylation



**Figure 2.15: *in vitro* 5hmC-combing on  $\lambda$  phage DNA**

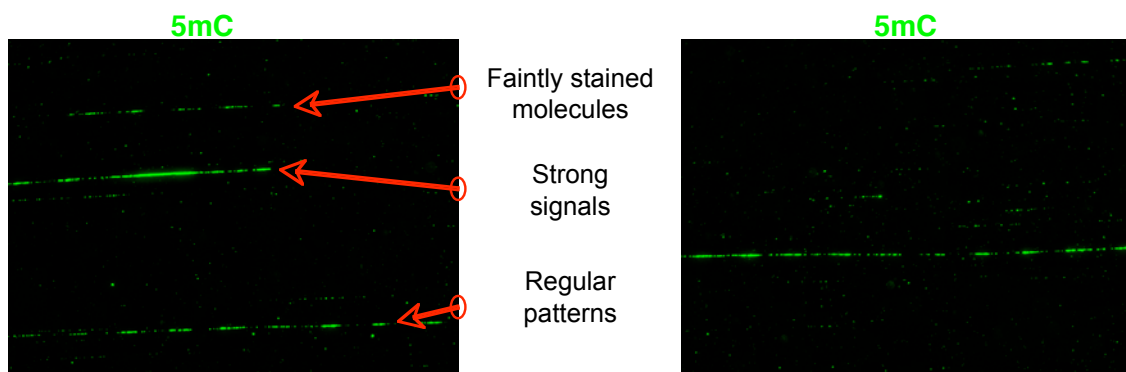
**A:** Rb $\alpha$ 5hmC (Active motif; 1:50 dilution in BlockAid) does not recognize  $\lambda$  and  $m\lambda$  DNA whereas  $^{hm}\lambda$  DNA is detected by the antibody. **B:** Same as A, but R $\alpha$ 5hmC (Diagenode; 1:50 dilution in BlockAid) is used. In addition, this antibody detects in the mixture of modified and native DNA the short 12.5kb fragments of  $^{hm}\lambda$  DNA only.



**Figure 2.16: Two examples of DNA preparations stained by Yoyo-1**

**A:** Displayed are drilled fibers of Imr90 DNA, that are twisted and not aligned in parallel.  
**B:** Illustrated are long stretches of single DNA fibers, which are aligned next to each other and more than 100 kb in length. Scale bar: 10  $\mu\text{m}$  (can be applied to A and B)

staining, regions that are highly methylated and DNA parts that show an alternating pattern of methylated and non-methylated DNA <sup>1</sup>.



experiment performed by Genomic Vision, Bagneux, France

**Figure 2.17: Diversity of 5mC signals on human genomic DNA**

Signals can be distinguished between faintly and strongly stained molecules as well as regular patterns of alternating methylated and non-methylated DNA.

Noteworthy, the detection of DNA methylation on genomic DNA so far is not locus-specific, therefore, a fluorescence hybridization step must be included in future experiments.

<sup>1</sup>Experiment performed by Jean-Pascal Capp, Genomic Vision, Bagneux, France

## 2.2 DNA methylation analysis of satellite-2 DNA

### 2.2.1 Satellite-2 DNA methylation analysis by metaphase immuno-FISH on metaphase chromosome spreads

DNA methylation of pericentromeric satellite-2 DNA is important for chromosome segregation and genome stability. The extent of DNA methylation at satellite DNA can be analyzed at low resolution by metaphase-immunoFISH according to the protocol described in 4.2.5.3 with commercially available biotinylated LNA probes for satellite-2 DNA (Exiqon) and metaphase chromosomes were analyzed using a fluorescence microscope (see Figure 2.18)

The biotinylated LNA probe gives strong signals of satellite-2 DNA on chromosomes 1, 9, 15 and 16 (see Figure 2.18A). When analyzing the chromosomes at higher resolution, it can be seen that most of the chromosomes show an overlap of heavy DNA methylation and satellite-2 DNA in pericentromeric regions (see Figure 2.18B and C). However, also hypomethylation was observed at both of the chromosomes 9: One shows partial overlap of 5mC and satellite-2 DNA with some satellite-2 repeats not methylated (Figure 2.18D) whereas the other region does not colocalize with heavy DNA methylation at all (Figure 2.18E).

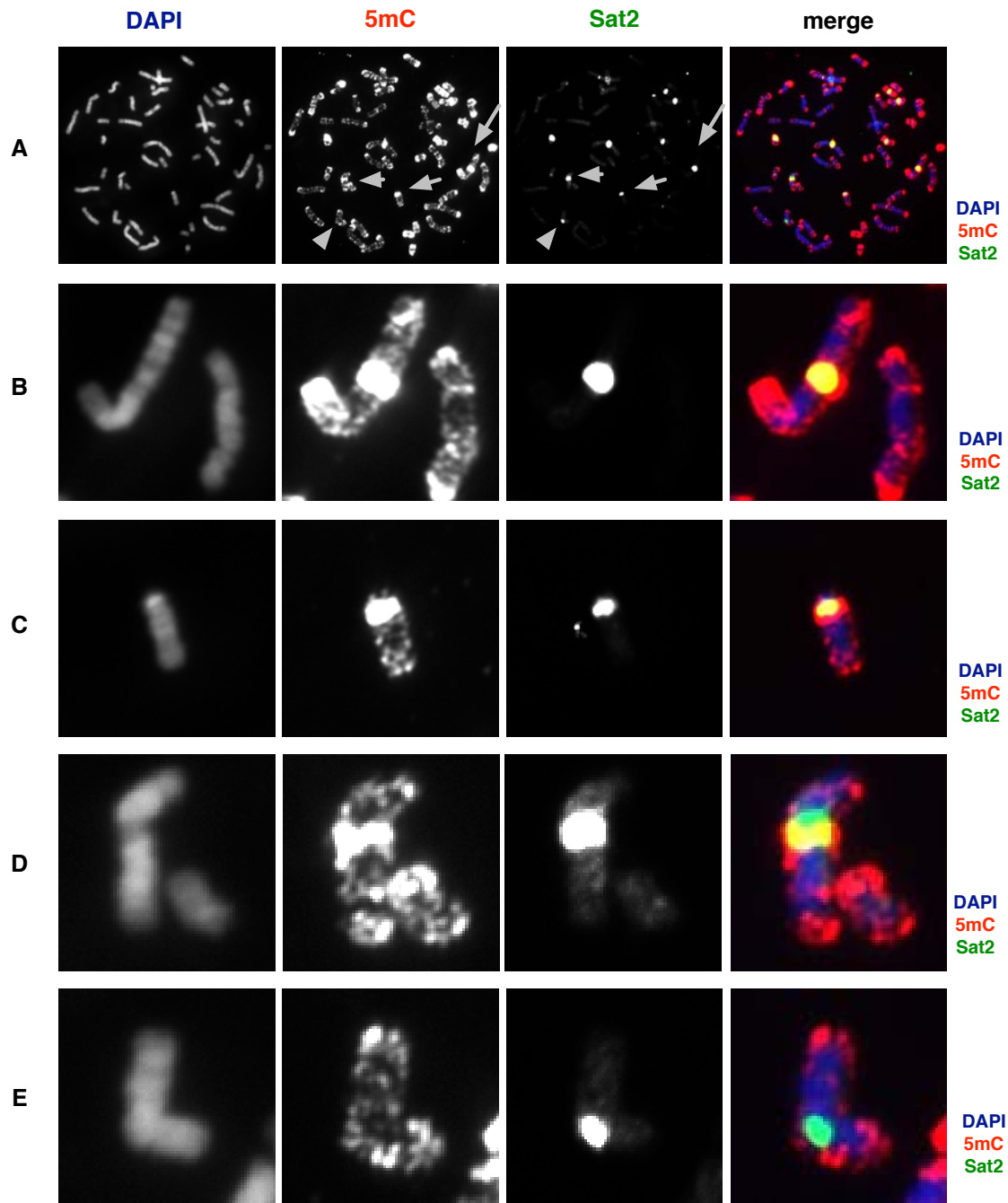
### 2.2.2 Satellite-2 DNA methylation analysis by epi-combing

#### Epi-combing reveals long stretches of methylated satellite-2 DNA

To enable DNA methylation analysis of satellite-2 DNA at a higher resolution, HCT116 DNA was isolated and combed as described in subsections 4.2.6.2, 4.2.6.1 and 4.2.6.3. Afterwards, DNA was detected using the biotinylated LNA probe that was already used in the metaphase immuno-FISH experiment as described above. DNA methylation is visualized according to the protocol described in 4.2.6.6. Examples of methylated satellite-2 signals obtained can be seen in Figure 2.19A, B and C which display a length of 20 kb, 140 kb and 180 kb respectively.

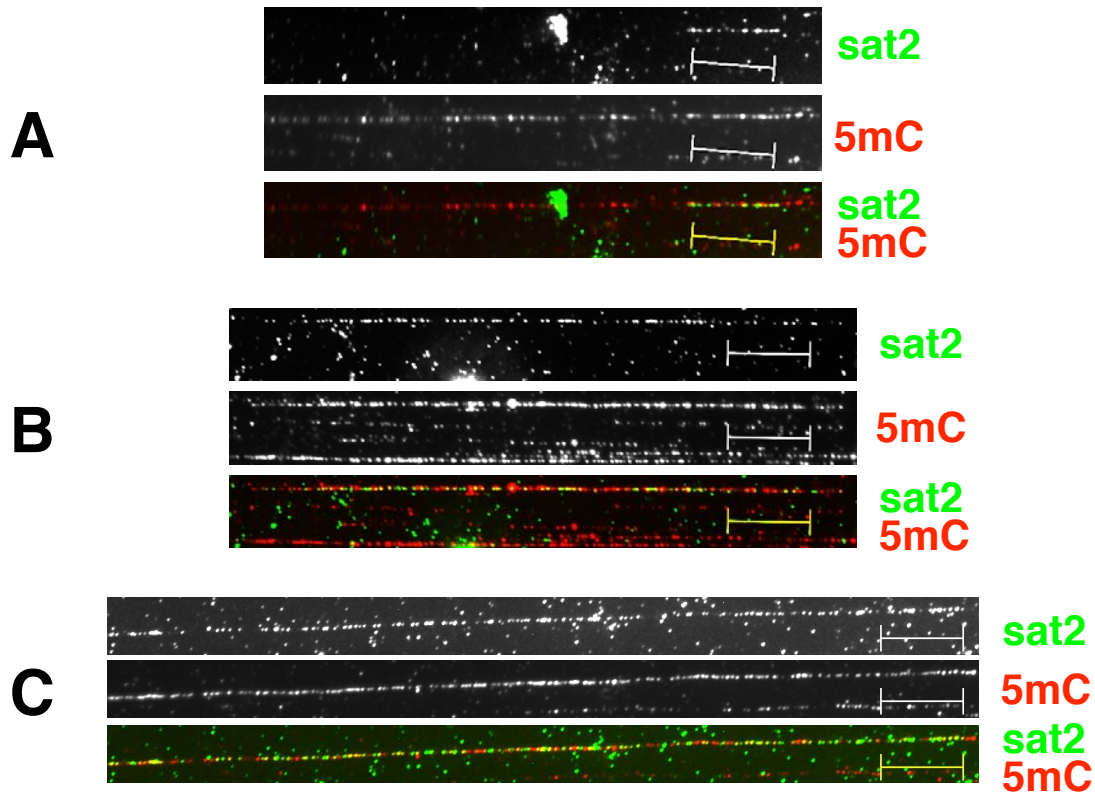
In summary, all analyzed satellite-2 repeats were methylated but the signals differed in their length. Stretches as short as about 2 kb were observed, however stretches of DNA that carry satellite-2 DNA signals over 200 kb in length could also be found. On average, the length of the analyzed repeats were 76 kb with a standard deviation of 55 kb. The results obtained display a high variety between different satellite-2 repeats, however they share DNA methylation as a common feature.





**Figure 2.18: DNA methylation analysis of satellite-2 DNA by metaphase immuno-FISH**

**A:** Metaphase spread stained with DAPI. Strong satellite-2 signals can be found on chromosomes 1,9,15 and 16. In most cases, these signals overlap with heavy DNA methylation. Arrows indicate chromosomes that are shown in magnification (B:-E:) **B:** and **C:** Chromosomes (1 and 15, respectively) of the same metaphase spread at higher resolution, both satellite-2 regions are fully methylated. **D:** Chromosome 9 at a higher resolution showing partial overlap of 5mC and satellite-2 DNA. **E:** The other chromosome 9 with pericentromeric satellite-2 DNA that is not heavily methylated.



**Figure 2.19: Satellite-2 DNA is present in long, methylated stretches in HCT116 cells**

HCT116 cells were harvested and further processed for combing. DNA methylation was detected by monoclonal antibody and layers of corresponding secondary and tertiary Cy3-coupled antibodies. Satellite-2 DNA was detected by fluorescence hybridization using commercial biotinylated LNA oligos against a 26mer of satellite-2 repeats, followed by detection of the LNA probe by StrA488. Scale bar: 10  $\mu\text{m}$  applicable to all images **A:-C:** All signals show strong DNA methylation **A:** Satellite-2 DNA is detected on a long stretch of DNA fiber. The length of the signal is about 20 kb and heavily methylated **B:** Same as A, but the satellite-2 signal corresponds to a stretch of 140 kb length **C:** Same as A, but this time, a 180kb stretch of repeated satellite-2 DNA is observed.

## 2.3 DNA methylation analysis of ribosomal RNA genes

### 2.3.1 rDNA methylation analysis by metaphase-immunoFISH on metaphase chromosome spreads

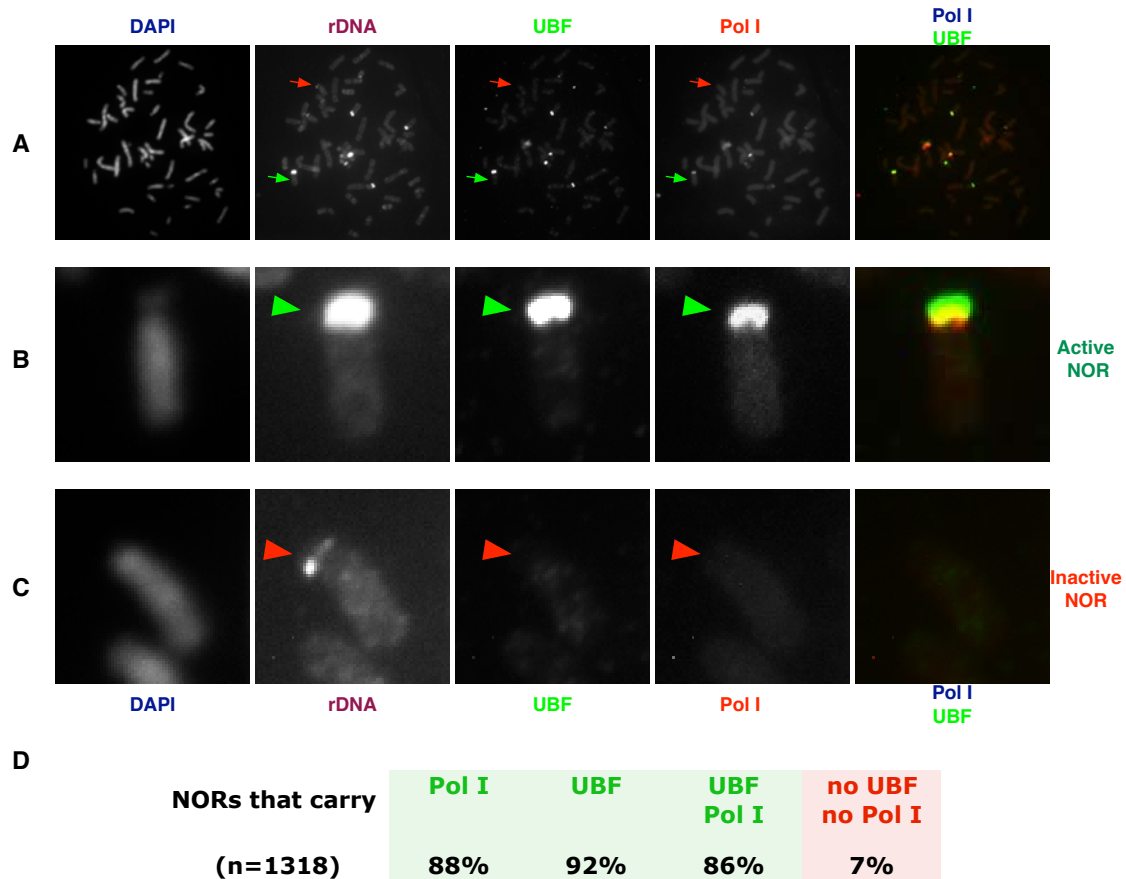
A method to analyse ribosomal DNA and NOR activity is metaphase immuno-FISH, which enables the parallel visualization of ribosomal DNA to identify NORs by fluorescence hybridization together with the immunodetection of proteins such as UBF and Pol I that mark active NORs. To investigate the activity of NORs, UBF and Pol I were detected on metaphase spreads of blood lymphocytes of a female individuals together with rDNA. The metaphase spread was consecutively analyzed using Rb- $\alpha$ -UBF. Then, Pol I was detected using M- $\alpha$ -RPA194. Both steps were performed according to the protocol in subsection 4.2.5.2. Afterwards, rDNA is detected by a fluorescence hybridization as described in 4.2.5.3. Finally, DNA was stained by DAPI and the slide was analyzed using a standard fluorescence microscope. An example for a metaphase spread is shown in Figure 2.20.

Figure 2.20A illustrates an entire metaphase spread with ten NORs seen in the rDNA lane. In addition, three NORs are inactive, since they do not exhibit any signals of UBF or Pol I. One example is marked with a red arrow and is displayed at higher resolution in Figure 2.20C. The majority of signals carry UBF and Pol I and are dubbed as active NORs (green arrow). Figure 2.20B shows the colocalization of Pol I and UBF at the NOR as well as the secondary constriction, the gap in the DAPI staining which is typical for active NORs.

Statistical analysis revealed that 86% of all NORs in blood lymphocytes are active and carry Pol I and UBF, whereas only a minor fraction of 7% carry neither of the active mark. The discrepancy to 100% comes from the fact of transcription competent NORs that carry UBF but not Pol I and therefore, are not part of the active fraction of 86%, where UBF and Pol I are both present.

To analyze the extent of DNA methylation at NORs and to compare it with the active mark UBF, another metaphase-immunoFISH is performed which will detect UBF and DNA methylation together with rDNA. The metaphase spread was analyzed using Rb- $\alpha$ -UBF according to the protocol in subsection 4.2.5.2. Afterwards, rDNA is detected by a hybridization mixture containing directly Cy3-labelled regions of the IGS (Hr4 and Hr6) as described in 4.2.5.3. After fluorescence detection, DNA methylation was detected according to the protocol of 4.2.6.6. Finally, DNA was stained by DAPI and the slide was analyzed using a standard fluorescence microscope. An example for a metaphase spread is shown in Figure 2.21.

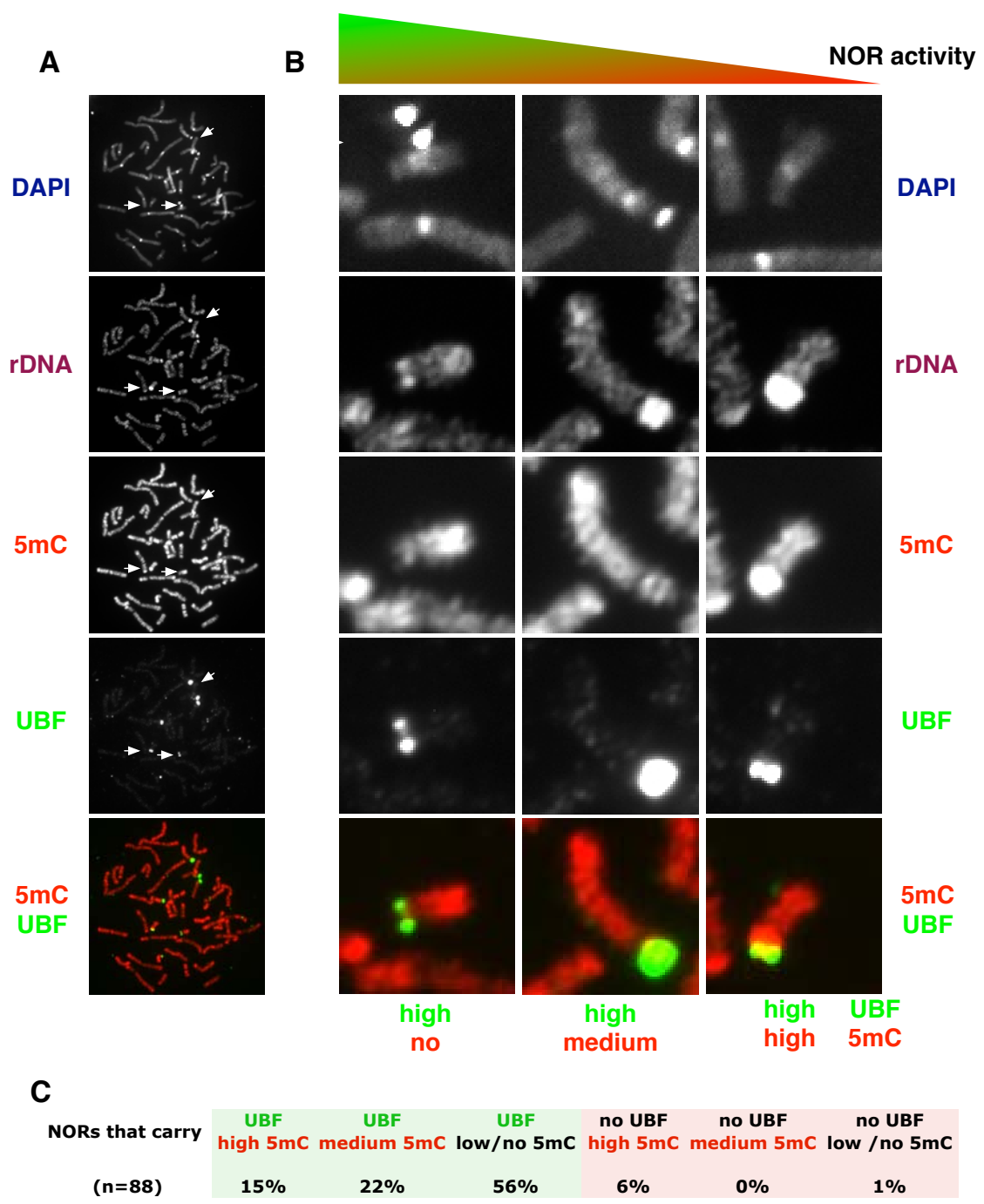
In contrast to ten rDNA signals that represent ten NORs in a diploid cell, only six rDNA signals could be clearly identified. Those NORs exhibit different activities, when DNA methylation and UBF staining are compared (exemplarily shown in 2.21B and marked with arrows in the overview) . The highest activity display NORs that carry UBF but are not methylated. Then, an increase in DNA methylation



**Figure 2.20: Co-existence of inactive and active NORs during metaphase of female blood lymphocytes**

A metaphase spread of female blood lymphocytes was subjected to metaphase-immunoFISH. UBF and Pol I were consecutively immunodetected with layers of A488- and Cy5-labelled antibodies, respectively. Afterwards, NORs are visualized by fluorescence hybridization using Cy3-labelled probes against human ribosomal DNA (Hr4, Hr6). Finally, DNA is stained by DAPI. **A:** The entire metaphase spread reveals the expected number of ten NORs (rDNA panel). In three NORs, no UBF or Pol I was found (red arrow for example). Remaining NORs exhibit active marks of Pol I and UBF (example shown with a green arrow). **B:** Chromosome marked in A with a green arrow shown at higher resolution. The localization of Pol I and UBF indicates an active NOR. **C:** Chromosome marked with a red arrow shown at higher resolution. No UBF or Pol I is found on the NOR, thus the NOR is inactive. **D:** After analysis of 1318 NORs, co-localization of Pol I and UBF can be found in 86% of NORs, and in 7% of analyzed signals, neither UBF nor Pol I was found.





**Figure 2.21: Correlation of UBF and rDNA methylation reveals gradual shades in NOR activity in female blood lymphocytes**

A metaphase spread of female blood lymphocytes was subjected to metaphase-immunoFISH. First, UBF was immunodetected with layers of A488-labelled antibodies. Afterwards, NORs are visualized by fluorescence hybridization using Cy3-labelled probes against human ribosomal DNA (Hr4, Hr6). Finally, DNA methylation was detected with M- $\alpha$ -5mC and with a secondary Cy5-labelled secondary antibody, as well as DNA was stained by DAPI. **A:** In the entire metaphase only six NORs (rDNA panel) can be clearly identified. Different combinations of UBF and 5mC signals can be detected, exemplarily marked with three arrows and shown at higher resolution in B. For **B:** In the first panel, the highest level of NOR activity can be detected, due to the UBF signal and the missing DNA methylation. This activity decreases to a co-localization of UBF with strong DNA methylation. **C:** In female blood lymphocytes, the majority of NORs (56%) display high UBF levels and no DNA methylation, however, also other combinations of NOR activity can be found to a significant level.

decreases the activity of the NOR, although the transcription factor UBF is still present.

However, not all ten NORs present in a cell could be clearly identified which has the potential of a biased analysis. Therefore, the experiment is repeated to prevent a bias in the analysis with metaphase slides of blood lymphocytes of a male individual which is shown in 2.22.

In this case, all ten rDNA signals were identified on the short arms of the acrocentric chromosomes. The overview of the entire slide is shown in Figure 2.22A whereas single chromosomes are exemplarily shown in Figure 2.22B. Again, several different activity states of NORs can be differentiated beginning in the left panel with the NOR of highest activity, since it carries UBF but no DNA methylation. Then, the level of DNA methylation increases although UBF is still present or partially overlaps, representing a decrease in overall NOR activity. Finally, there are also NORs that have a variable DNA methylation profile, but due to the missing UBF signal, correspond to the least active NORs.

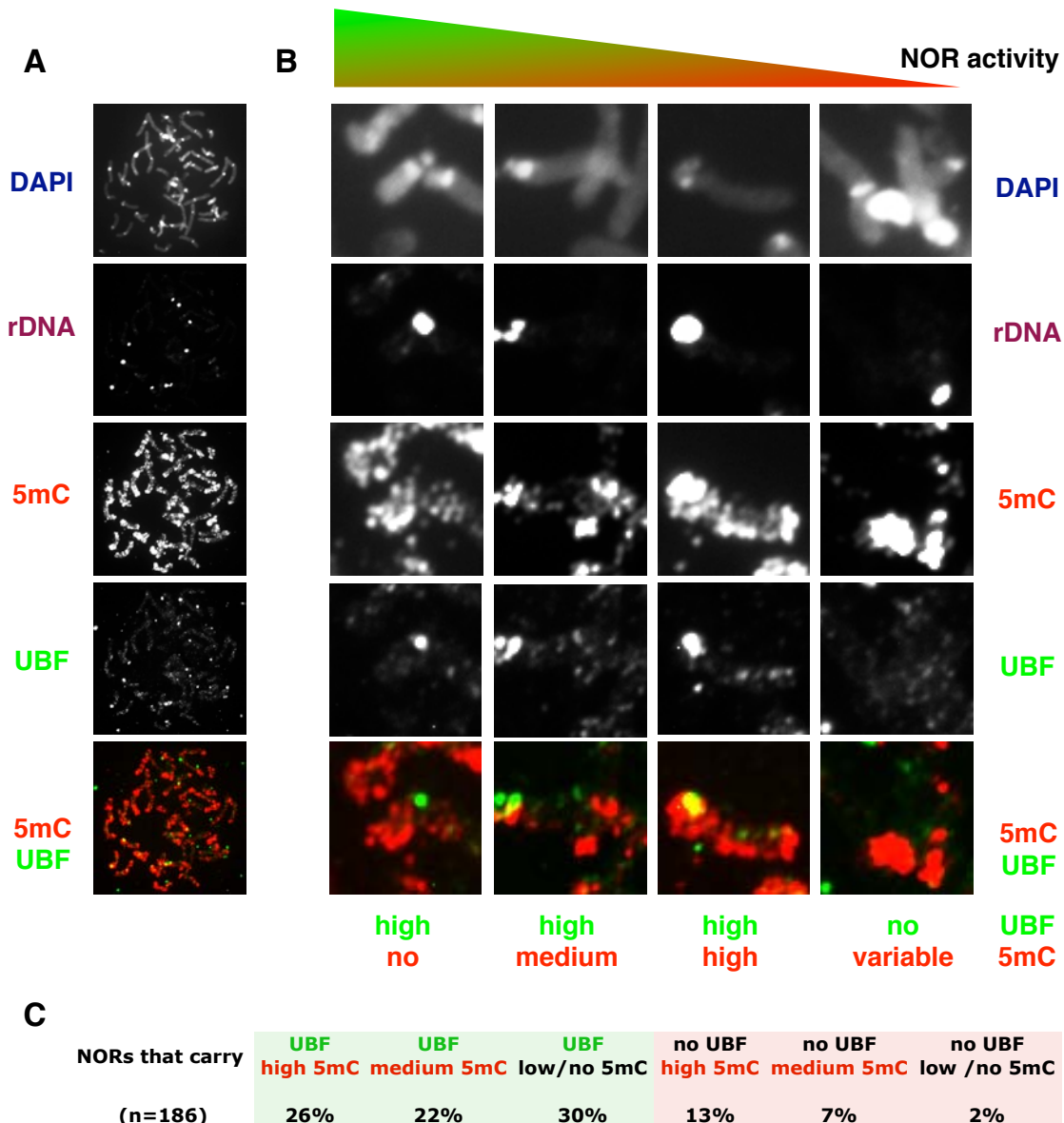
These data suggest gradual shapes of NOR activity and stands in contrast to the classic model of NORs that are either active or inactive. To further analyse these observations, a higher resolution is needed that resolves single rDNA repeats. Therefore, the DNA methylation of ribosomal DNA in different cell lines will be examined by epi-combing.

### 2.3.2 Epigenetic analysis of ribosomal DNA repeats reveal epigenetic clusters and transitions

#### **A bar code system for rDNA allows the discrimination between canonical and non-canonical rDNA repeats**

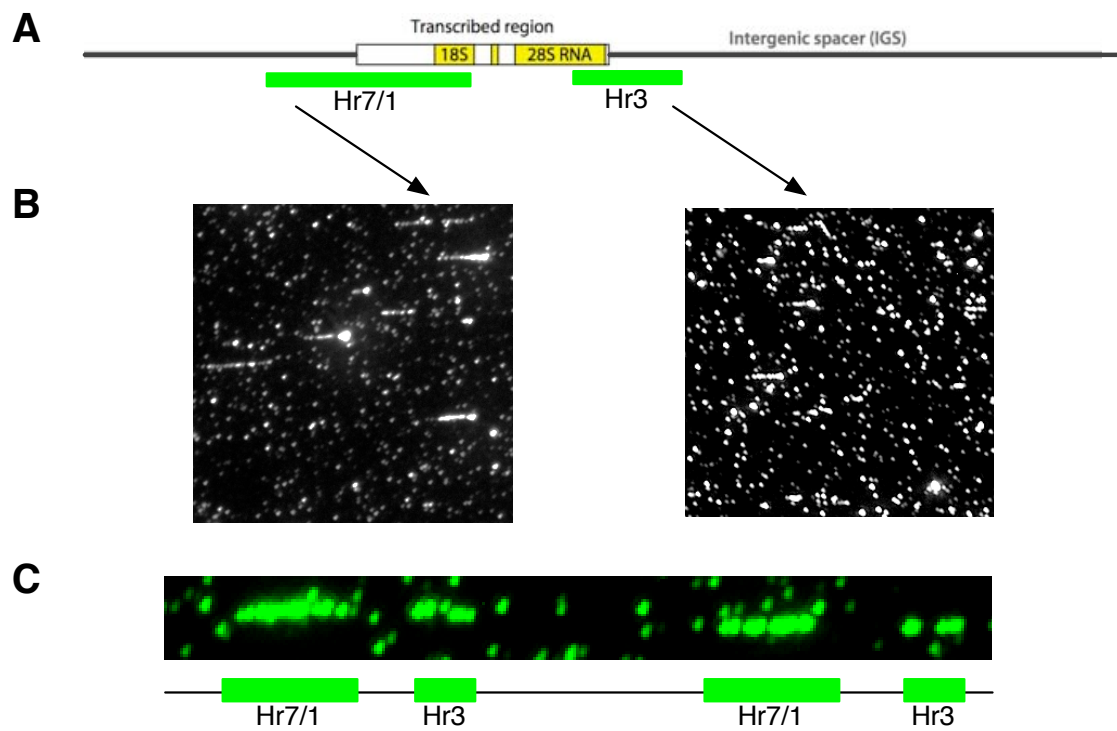
Ribosomal DNA belongs to the class of tandem repeated arrays and is well suited for the analysis by epi-combing because of the size and mosaic methylation pattern of individual repeats. To distinguish between canonical and non-canonical repeats and their orientation, a bar code system will be established that uses a long probe and a short probe. The long probe hybridizes to region 37.9-5.3 kb of the human rDNA (Hr7/1) whereas the short one matches the region 12.4-18.1(Hr3). The numbers show nucleotide positions in kilobase pairs relative to the GenBank U13369 sequence. +1 is the transcriptional start site, the primary transcript terminates at 13.3 kb. The bar code scheme as well as the *in vitro* tests of corresponding hybridization mixtures are shown in Figure 2.23.

To test the hybridization mixture for Hr7/1 as well as Hr3, the corresponding plasmids are digested with HindIII and 1ng/ $\mu$ L is combed and detected by fluorescence hybridization according to 4.2.6.7 and subsequently analyzed under a fluorescence microscope. Short stretches of linear sequences indicate specific signals for Hr7/1 and Hr3, respectively (see 2.23B). A 1:1 mixture of those two hybridization mixtures were added to combed Imr90 genomic DNA and detected as outlined above by fluorescence hybridization. The mixture of long and short repeats result in a bar



**Figure 2.22: Colocalization of UBF and rDNA methylation reveals gradual decreases in NOR activity in male blood lymphocytes**

A metaphase spread of female blood lymphocytes was subjected to metaphase-immunoFISH. First, UBF was immunodetected with layers of A488-labelled antibodies. Afterwards, NORs are visualized by fluorescence hybridization using Cy3-labelled probes against human ribosomal DNA (Hr4, Hr6). Finally, DNA methylation was detected with M- $\alpha$ -5mC and a secondary Cy5-labelled secondary antibody, as well as DNA was stained by DAPI. **A:** Ten NORs, according to a diploid cell can be clearly identified (rDNA lane). Different combinations of UBF and 5mC signals can be detected and examples are shown in B at a higher resolution. For better illustration, UBF (green) and 5mC (red) are merged in the lower lane. **B:** In the first panel, the highest level of NOR activity can be detected where a NOR carries UBF but no DNA methylation. In the second panel, the activity of the NOR is already decreased, since a medium level of DNA methylation partially overlaps with UBF. This activity decreases further to a co-localization of UBF with strong DNA methylation. The most inactive version of NORs can be found in the right panel, where DNA methylation is present, but no UBF signal. **C:** In male blood lymphocytes can be found a gradual decrease in NOR activity from NORs (30%) display high UBF levels and no DNA methylation to intermediate activity of UBF with colocalization of UBF and DNA methylation and inactive NORs that do not display UBF signals.



**Figure 2.23: Locus-specific detection of ribosomal DNA by epi-combing**  
**A:** rDNA unit illustrating the bar code system with a long and a short probe to distinguish canonical and non-canonical repeats. **B:** HindIII-linearized plasmids containing the rDNA regions Hr7/1 and Hr3 are combed and detected by fluorescence hybridization using the StrA488 detection system to test the home made biotinylated probes. **C:** Hybridization mixture of Hr7/1 and Hr3 detects specifically ribosomal DNA and its orientation on genomic DNA, resulting in a bar code as depicted in A.

code system on DNA fibers, in this case two canonical repeats next to each other in a head-to-tail orientation.

Unless it is stated otherwise in the text, all experiments with different cell lines were performed according to described protocols. The cells were harvested and embedded in plugs as explained in 4.2.6.1 and the DNA was isolated as described in 4.2.6.2. The high molecular weight DNA was combed according to 4.2.6.3. After the DNA was fixed, the DNA modification was detected with M- $\alpha$ -5mC (see subsection 4.2.6.6). Ribosomal DNA was detected for each cell line as illustrated in Figure 2.23 and described in subsection 4.2.6.7. Then, the slides were subsequently analyzed with a fluorescence microscope.

### **Epi-combing analysis of rDNA in proliferating IMR90 cells**

Imr90 cells were analyzed by epi-combing (500,000 cells/plug) and ribosomal DNA repeats are exemplarily shown in Figure 2.24 and Figure 2.25. Epigenetic clustering of neighbouring repeats were observed in many cases (Figure 2.24).

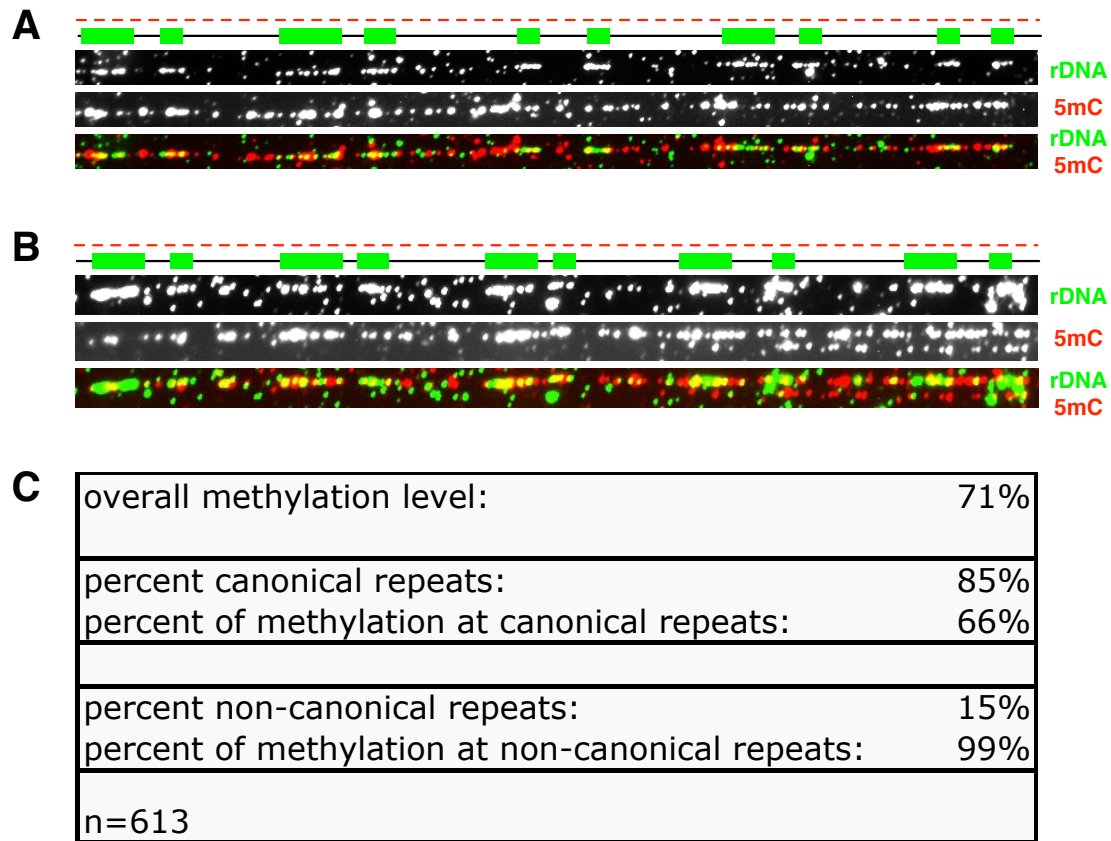
Figure 2.24A shows a combination of canonical and one non-canonical repeat, which are all methylated. In addition, Figure 2.24B illustrates the epigenetic clustering of DNA methylation found at a stretch of five neighbouring canonical rDNA repeats. After analysis of individual repeats ( $n=613$ ), it was calculated that Imr90 possess 85% canonical repeats and 15% non-canonical repeats. Furthermore, the overall methylation level of Imr90 rDNA was determined to be 71%. This value consists of DNA methylation at canonical repeats which is 66% and methylation at non-canonical repeats, which are apparently almost all methylated (99%).

In addition to the epigenetic clustering displayed in Figure 2.24, which is the case for many repeats, also so far unknown transitions between the DNA methylation of neighbouring repeats could be proven (Figure 2.25). DNA fibers presented in 2.25 show that not all neighbouring repeats share DNA methylation, but alterations of epigenetic states can happen, in the case of a series of canonical repeats (2.25 A and B) as well as in the case of canonical repeats that are interspersed by non-canonical rDNA (2.25 B). Interestingly, these epigenetic transitions are not a rare event but occur in 11.5% of all possible transition sites (45 out of 393).

### **Epi-combing analysis of rDNA in senescent IMR90 cells**

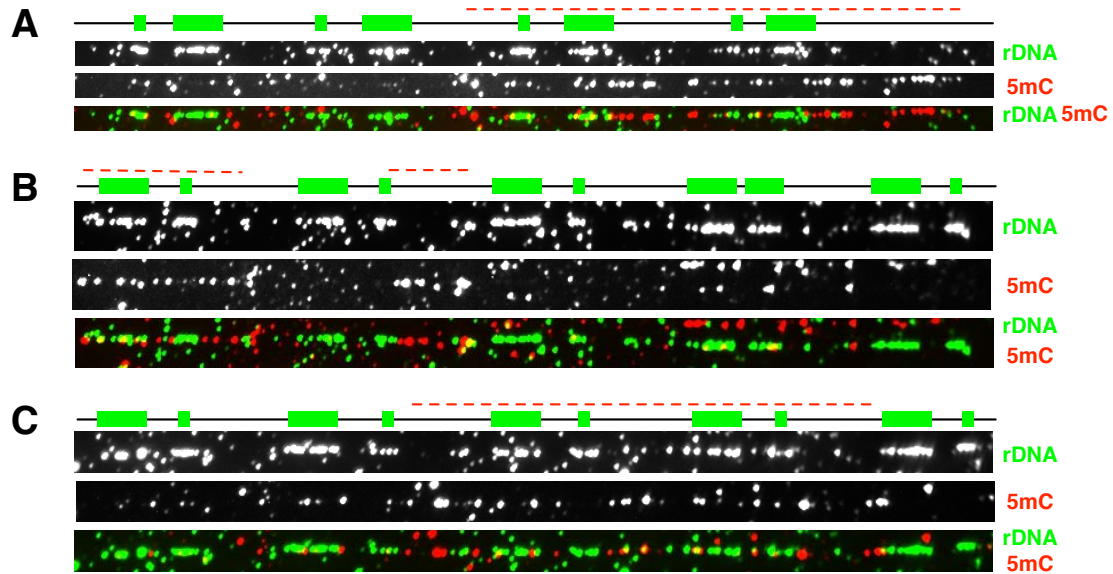
Imr90 cells were cultured to a point where they do not proliferate anymore, called replicative senescence (p27), to analyze possible dynamics in rDNA methylation upon ageing. Senescent cells are characterized by expression of the lysosomal enzyme  $\beta$ -galactosidase ( $\beta$ -gal) [6, 32] as well as the absence of the protein Ki-67 [59]. The expression of the human Ki-67 protein is strictly associated with cell proliferation and is absent from resting cells (G(0)), which makes it an excellent marker for determining the so-called growth fraction of a given cell population.

$\beta$ -gal is tested as described in subsection 4.2.3.3 and Ki-67 is immunodetected with the protocol for immunofluorescence 4.2.5.2. The corresponding microscopic pictures are shown in Figure 2.26A and B, respectively, and both prove that the cells



**Figure 2.24: Epigenetic clustering of ribosomal DNA in Imr90 cells**

Imr90 cells are subjected to epi-combing analysis by detection of 5mC and ribosomal DNA by the bar code system. **A:-B:** The bar code for each DNA fiber as well as the detected DNA methylation is depicted as a scheme above the microscopic pictures for clearer illustration. In both cases, all neighbouring repeats are methylated, resulting in clustering of the epigenetic mark. **A:** The bar code system reveals from left to right two canonical repeats followed by two short signals, indicating a non-canonical repeat. This one is neighboured by another canonical repeat which follows another non-canonical rDNA. **B:** Shown are five canonical repeats arranged in a head-to-tail orientation depicted in the standard model of rDNA. **C:** The statistics of 613 analyzed repeats reveal that 99% of non-canonical repeats in Imr90 cells are silenced by DNA methylation.



**Figure 2.25: Epi-combing analysis illustrates transitions of epigenetic states between neighbouring repeats**

Imr90 cells are subjected to epi-combing analysis by detection of 5mC and ribosomal DNA by the bar code system. **A:-C:** The bar code for each DNA fiber as well as the detected DNA methylation is depicted as a scheme above the microscopic pictures for clearer illustration. Transitions are illustrated as large gaps in the red 5mC signal in the scheme.

**A:** Four canonical repeats are found on a single DNA fiber in a classic head-to-tail orientation. The first two repeats on the left are unmethylated, whereas the two repeats on the right are methylated, resulting in a transition. **B:** A series of four canonical repeats is interspersed by one non-canonical repeat which consists of two long signals. In this series, all but the left one are not methylated. **C:** Five canonical repeats are depicted, where DNA methylation can only be found in the middle of the rDNA repeats.

do not proliferate anymore but are senescent <sup>2</sup>. After confirmation of senescence, the cells were subjected to analysis of rDNA methylation by epi-combing (Figure 2.26C and D). Also in senescent cells, epigenetic clustering and transitions can be found, exemplarily shown by two stretches of four canonical repeats, each. In the case of senescent Imr90 cells, transitions happen more frequently with a percentage of 26.7% of transition sites (54 out of 202). The statistical analysis (n=290) revealed that the ratio of canonical (87%) and non-canonical repeats (13%) stays constant compared to young Imr90. However, the overall methylation level is slightly decreased to 58%. The most remarkable difference is the methylation status of non-canonical repeats, which is reduced to 64% in senescent cells (see Figure 2.26E).

### **Epi-combing analysis of rDNA in proliferating HCT116 cells**

The epi-combing results of DNA methylation at ribosomal DNA loci in the colon cancer cell line HCT116 are shown in Figure 2.27. Epigenetic clustering of DNA methylation (see Figure 2.27A-C) as well as alterations of DNA methylation (Figure 2.27D and E) are exemplarily shown. In this malignant cancer cell line, non-canonical repeats make up 34% of rDNA and these repeats are methylated in almost all cases (percent methylation: 92%). The canonical repeats are methylated to 44%, which sums up to an overall methylation level of ribosomal DNA methylation of 60% (n=243). Interestingly, transitions between epigenetic states of neighbouring repeats happen in 8% of all possible transition sites (16 out of 190 transition sites).

### **Epi-combing analysis of rDNA in proliferating MCF7 cells**

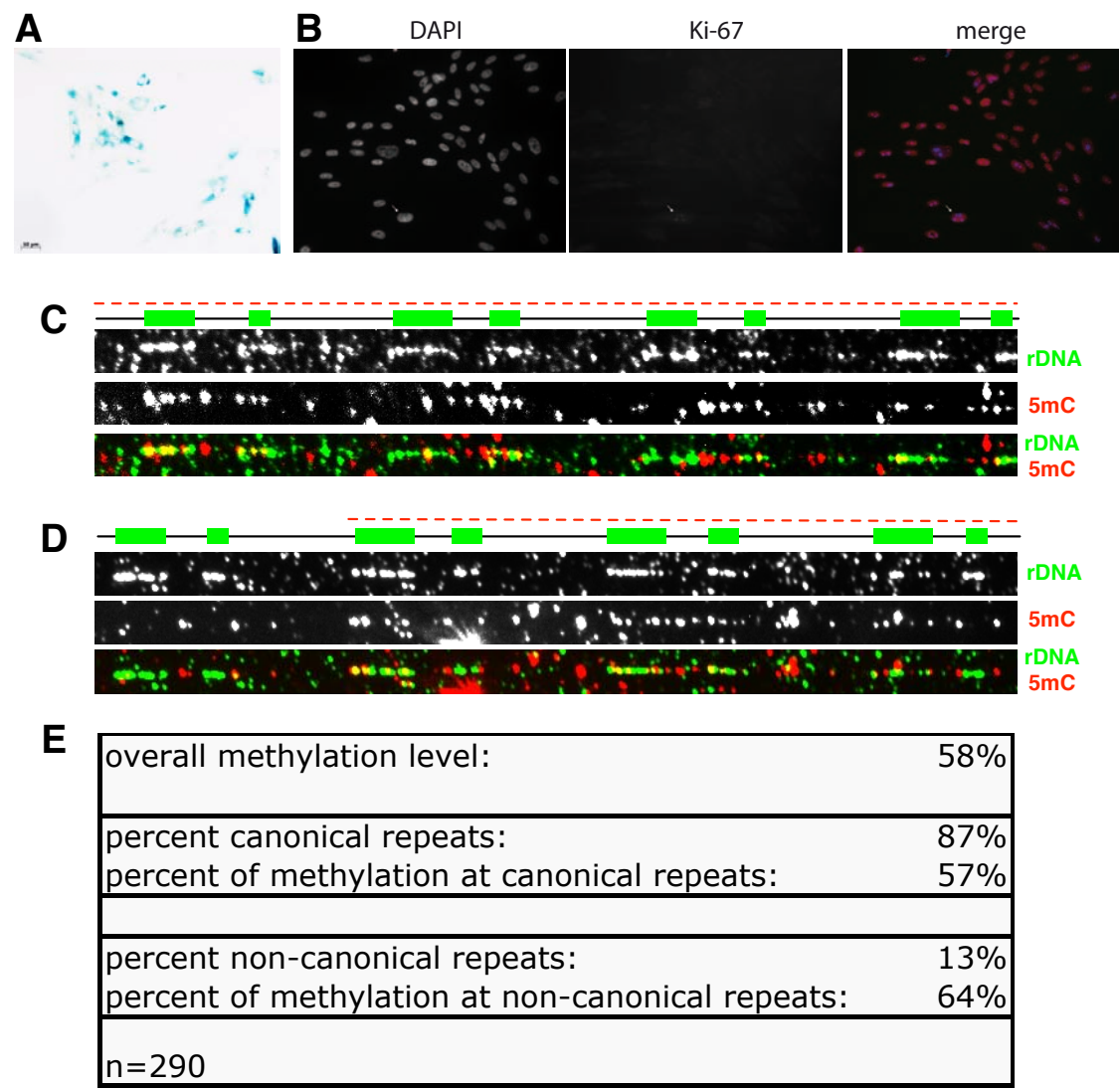
Another cancer cell line is analyzed, the breast cancer cell line MCF-7 and the epi-combing results on rDNA methylation is shown in Figure 2.28.

Epigenetic clustering (Figure 2.27A and D) as well as transitions (Figure 2.27B and C) could be observed. The percentage of non-canonical repeats increased to 36% and all of them are methylated. Interestingly, also 92% of canonical repeats are methylated, which results in an overall high level of rDNA methylation (95%) in MCF7 cells (Figure 2.27E). Though, the low number of analyzed repeats (n=39) should be significantly increased to obtain substantiated data.

---

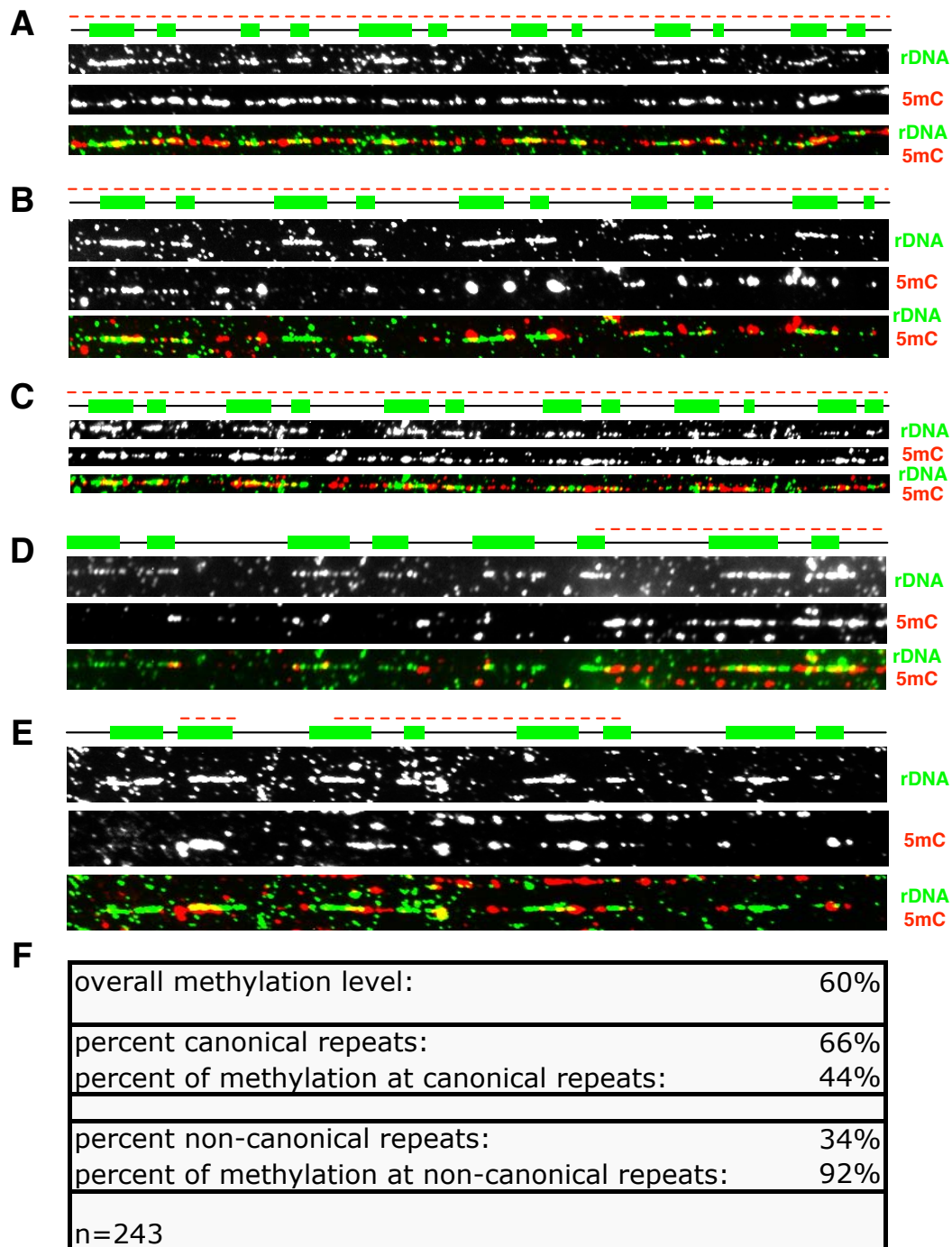
<sup>2</sup>Immunofluorescence of Ki-67 and  $\beta$ -gal staining were performed by Stefan Dillinger



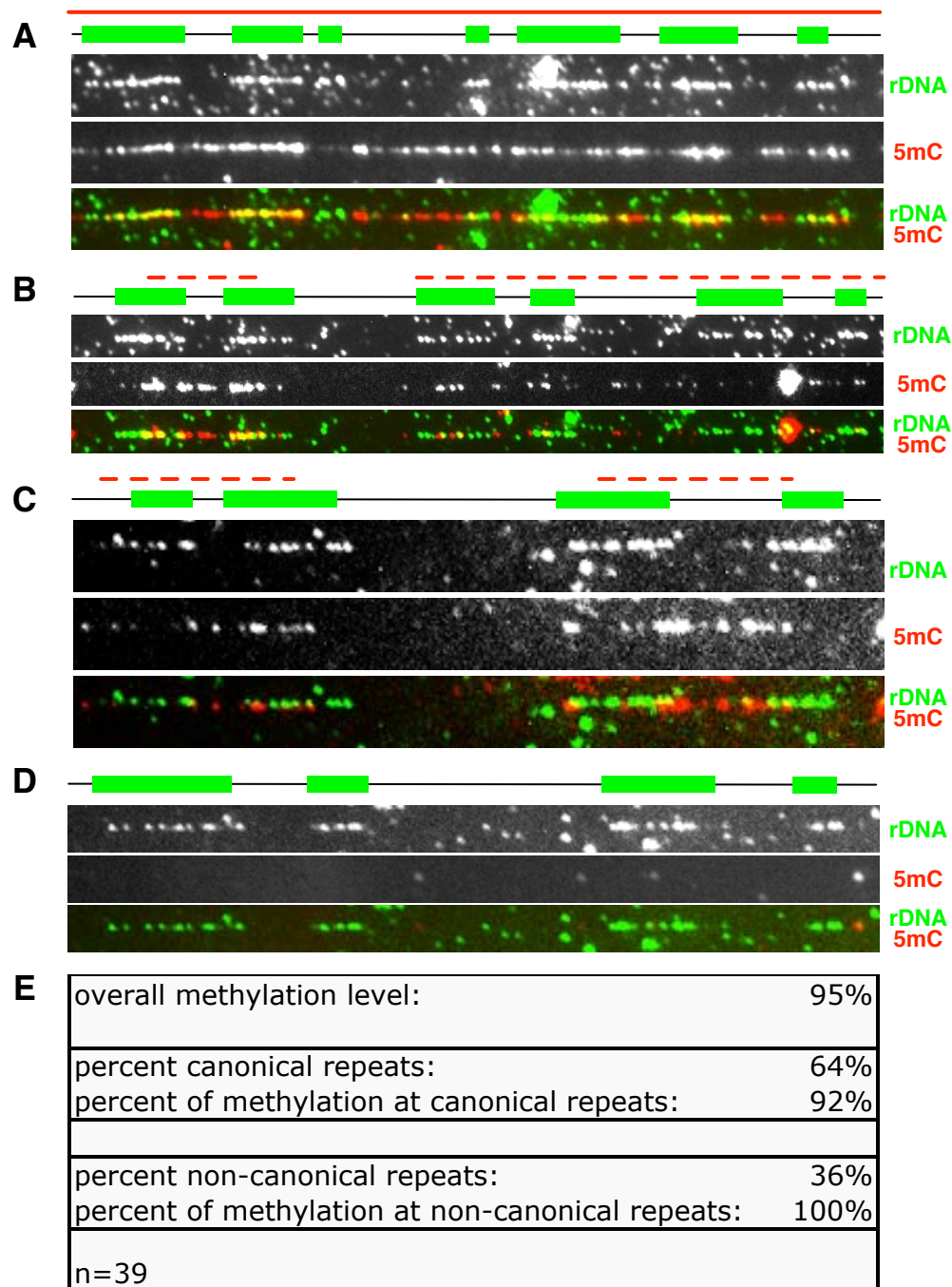


**Figure 2.26: DNA methylation analysis of senescent cells by epi-combing reveals hypomethylation at non-canonical repeats**

**A:** Staining of  $\beta$ -galactosidase of Imr90 (p27) shows that almost all cells are stained positive for the enzyme, thus exhibit replicative senescence. **B:** The lack of the immunostaining of Ki67 is an additional proof for replicative senescence. **C:-D:** Senescent Imr90 cells are subjected to epi-combing analysis by detection of 5mC and ribosomal DNA by the bar code system which is depicted for each DNA fiber as well as the detected DNA methylation as a scheme above the microscopic pictures for clearer illustration. In both cases, a DNA fiber is shown with four canonical repeats in a head to tail orientation. **C:** Epigenetic clustering can also be found in the case of senescence, since the signal of DNA methylation is found along the coding regions. **D:** Four canonical repeats are shown where the left repeat does not exhibit DNA methylation. **E:** Statistics of 290 repeats analyzed revealed the level of non-canonical repeats remained stable, however during senescence these repeats get less methylated (64%).



**Figure 2.27: DNA methylation analysis of HCT116 cells by epi-combing**  
 The rDNA bar code system as well as the detected DNA methylation are depicted as a scheme for clearer illustration. **A:-C:** In all cases, neighbouring repeats display the same epigenetic state, thus DNA methylation is clustered. **A:** A canonical repeat on the left is followed by a non-canonical and four canonical repeats which are methylated. **B:** Five canonical repeats that are all methylated. **C:** Six methylated canonical repeats are illustrated. The stretching factor decreases to the right which might be due to the fact that it is the end or beginning of the DNA fiber. **D:-E:** Transitions in the epigenetic status of neighbouring repeats are observed. **D:** Four canonical repeats where only the right repeat is methylated. **E:** The fiber starts with a non-canonical repeat consisting of two long signals followed by three canonical repeats. Only the right of the two long signals is methylated as well as the two following repeats. The right rDNA repeat is unmethylated. **F:** Statistics of the rDNA repeats analyzed.



**Figure 2.28: DNA methylation analysis of MCF7 cells by epi-combing**  
MCF7 cells are subjected to epi-combing analysis by detection of 5mC and ribosomal DNA by the bar code system which is depicted for each DNA fiber as well as the detected DNA methylation as a scheme above the microscopic pictures for clearer illustration. **A:-D:** In all cases, neighbouring repeats display the same epigenetic state. **A:** Depicted is a canonical repeat followed by a 5'-palindromic structure, which belongs to non-canonical repeats. All repeats are silenced by DNA methylation. **B:** A non-canonical repeat is followed by two canonical ones that are all methylated. **C:** Two methylated repeats are arranged in a head-to-head arrangement **D:** Two canonical repeats that are both unmethylated. **E:** Statistics of the rDNA repeats analyzed. The overall high level of DNA methylation (95%) is also reflected by DNA methylation at non-canonical repeats (100%)

### 2.3.3 Comparison of ribosomal DNA methylation in Werner Syndrome cells and primary fibroblasts

Previous experiments revealed a high level of non-canonical repeats in Werner syndrome patients compared to healthy individuals [20] and it was speculated that this is caused by genomic instability and may contribute to the pathological state. So far, it was not determined whether Werner syndrome patients display aberrant DNA methylation at the rDNA locus and if this change in the levels of non-canonical repeats is also reflected in a change in the DNA methylation status of those repeats. Therefore, DNA of primary fibroblast cells of healthy individuals (AG05283 and AG13077) as well as DNA from Werner syndrome patients (AG12797 and AG06300), that were also used in the aforementioned publication were subjected to epi-combing analysis to reveal new insights into ribosomal DNA methylation. In contrast to other experiments, plugs with isolated DNA (100,000 cells/plug) were obtained from Genomic Vision and the experiment was started by isolating DNA from plugs (subsection 4.2.6.2).

#### rDNA methylation of primary fibroblasts

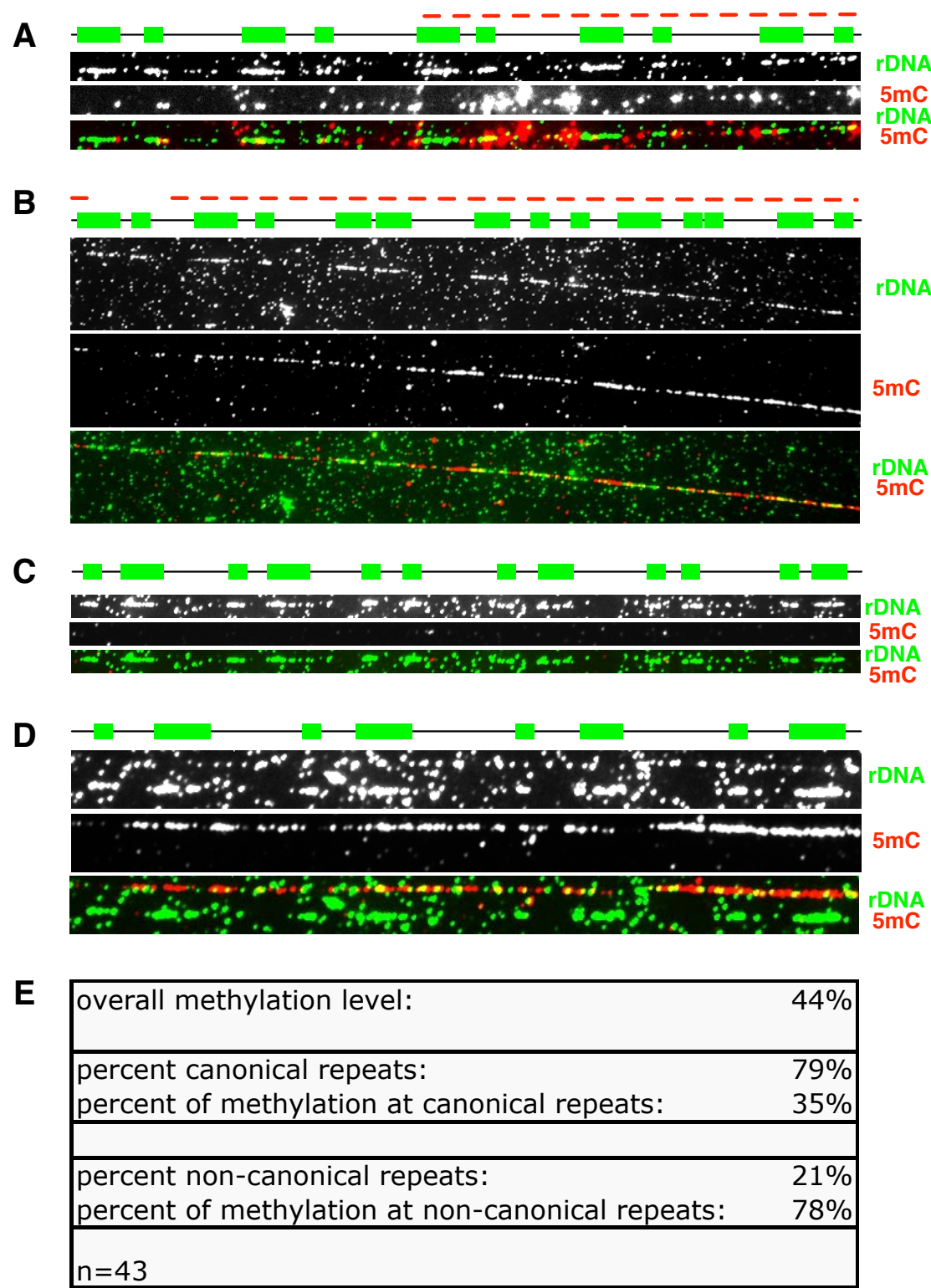
Ribosomal DNA methylation was analyzed by epi-combing in fibroblast cells of a healthy individual. Results obtained are illustrated in Figure 2.29. Also in primary fibroblasts, similar results can be found as in cultured cell lines. Again, not all neighbouring repeats show the same epigenetic status but transitions occur (Figure 2.29A and B). However, the majority of repeats can be found in epigenetic clusters where adjacent repeats show same DNA methylation state (Figure 2.29C and D). The statistical analysis of 43 repeats showed that there is a low percentage of non-canonical repeats (21%) in these cells. These non-canonical repeats are not all methylated, because 22% do not exhibit DNA methylation. The percent of methylation at canonical repeats is in the range of 35%, which leads to a combined methylation level of 44% in AG05283 cells.

The data of the second fibroblast cells are shown in Figure 2.30 with one example of ribosomal DNA patterns revealing transitions (Figure 2.30A) and one stretch of ribosomal DNA showing same DNA methylation status (Figure 2.30B).

Compared to AG05283 cells, this time non-canonical repeats are strongly enriched (59%) and all those non-canonical repeats are methylated. The remaining canonical repeats are methylated to 67%, which leads to an overall methylation level of 86%, which is almost double the DNA methylation as the first primary fibroblast cell line. Noteworthy, the sample volume is small ( $n=43$  and  $n=22$ ) and the data obtained in this presented work need to be extended to get a clearer picture of rDNA methylation in primary fibroblasts.

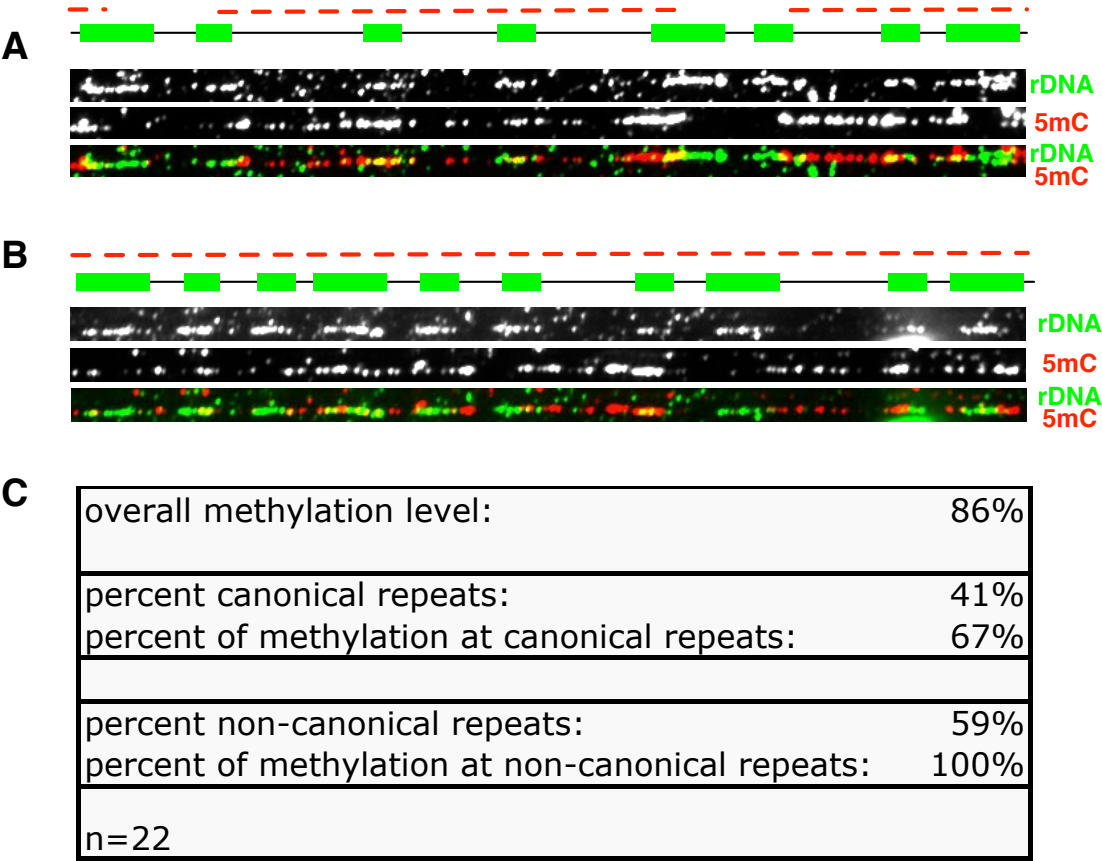
#### rDNA methylation analysis of Werner syndrome cells

Werner syndrome patients display overall genomic instability and therefore, an increase in non-canonical repeats. Analysis of ribosomal DNA methylation should clarify, whether these repeats are methylated. The data of Werner syndrome cells



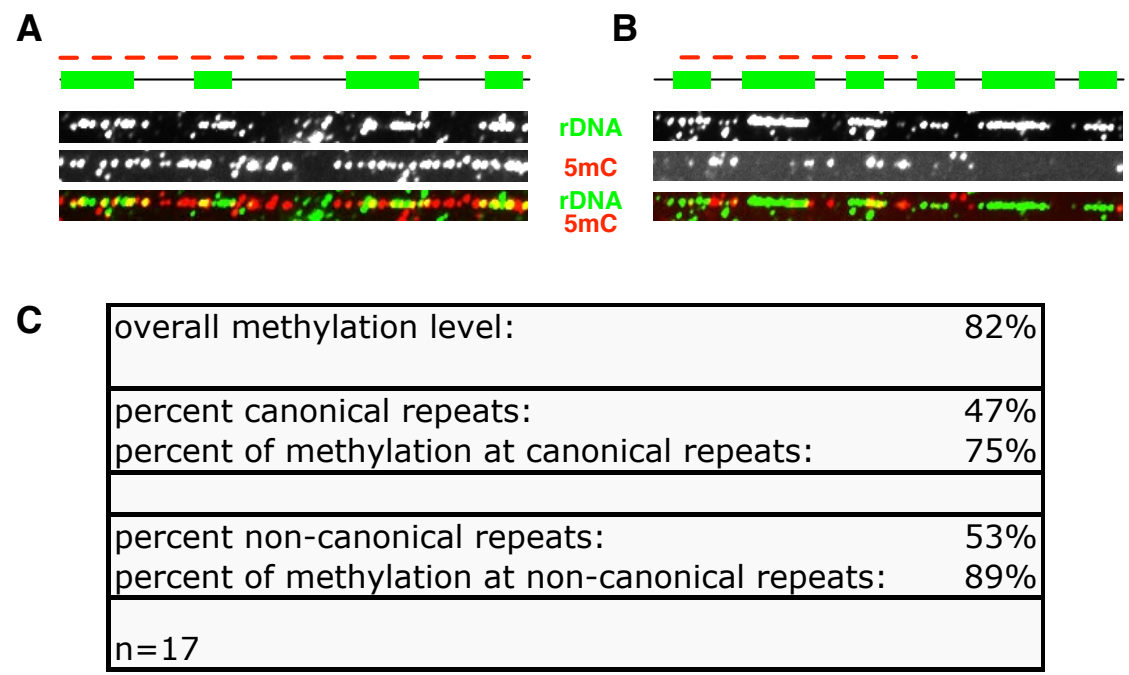
**Figure 2.29: DNA methylation analysis of AG05283 cells by epi-combing**  
Primary fibroblast cells are subjected to epi-combing. The bar code system as well as DNA methylation is depicted as a scheme above the microscopic pictures for clearer illustration. **A:-B:** Repeats display transitions between the epigenetic state. **A:** Depicted is a stretch of five canonical repeats where only the first two repeats on the left are unmethylated. **B:** Two canonical repeats are followed by a non-canonical one and a complex palindromic structure before one canonical repeat. Except the first canonical repeat, all are methylated. **C:** Four canonical repeats are interspersed by two short non-canonical ones. The epigenetic state is clustered since all repeats are unmethylated. **D:** Four canonical repeats that are all unmethylated. **E:** Statistics of the rDNA repeats analyzed.





**Figure 2.30: DNA methylation analysis of AG13077 cells by epi-combing** Primary fibroblast cells are subjected to epi-combing. The bar code system as well as DNA methylation is depicted as a scheme above the microscopic pictures for clearer illustration. **A:** Depicted are two canonical repeats interspersed by a non-canonical repeat. The last repeat is inverted in orientation compared to the others and belongs to the class of non-canonical repeats. Non-canonical repeats are methylated whereas the rDNA repeats in the same orientation are both unmethylated. **B:** A complex 3'-3' palindromic structure is depicted before two canonical repeats. All repeats are marked as methylated. **C:** Statistics of the rDNA repeats analyzed demonstrates an overall high level of DNA methylation.

AG12797 are illustrated in Figure 2.31. Exemplarily shown are two canonical repeats that display the same epigenetic pattern (Figure 2.31A) as well as a complex pattern of non-canonical ribosomal DNA repeats, where only the left one is methylated (Figure 2.31B).

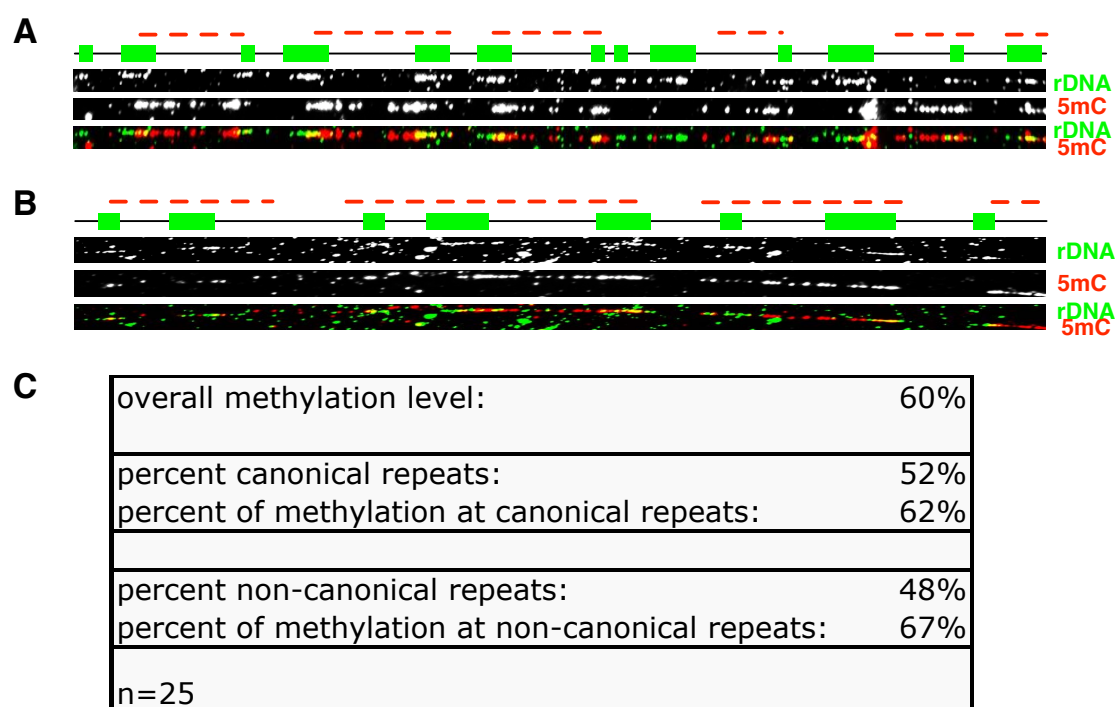


**Figure 2.31: DNA methylation analysis of AG12797 cells by epi-combing** Fibroblasts of a Werner syndrome patient are subjected to epi-combing. The bar code system as well as DNA methylation is depicted as a scheme above the microscopic pictures for clearer illustration.**A:** Depicted are two canonical repeats that are both methylated. **B:** A complex structure is displayed with apparently to canonical repeats separated by a short signal. DNA methylation signals reveal that only the first half of the structure is silenced. **C:** Statistics of the rDNA repeats analyzed.

In the data obtained by epi-combing, 47% of all repeats analyzed are canonical and those are in 25% unmethylated. In contrast, 53% are non-canonical repeats and those are unmethylated in only 11% of the repeats analyzed. Therefore, there is an overall high level of methylation at rDNA repeats of 82%.

The last cell line that was examined with epi-combing is AG06300, which are fibroblasts of a patient with Werner syndrome. Two examples of repetitive rDNA patterns are shown, one for epigenetic transition (Figure 2.32A) and one for clusters of repeats all regulated by DNA methylation (Figure 2.32B).

The overall methylation level obtained from these data is 60%, which is composed of DNA methylation of canonical repeats (62%) and non-canonical rDNA (67%) with a content of non-canonical repeats of 48%. Again, the sample size must be increased before final conclusions can be drawn.



**Figure 2.32: DNA methylation analysis of AG06300 cells by epi-combing**  
 Fibroblasts of a Werner syndrome patient are subjected to epi-combing. The bar code system as well as DNA methylation is depicted as a scheme above the microscopic pictures for clearer illustration. **A:** Two non-canonical repeats in the middle are flanked by two canonical repeats on each side. Transitions of DNA methylation are frequent. **B:** Two canonical repeats on the left are followed by two inverted repeats on the right, forming a 5'-5' palindromic structure in the middle. Since all repeats are methylated, they are epigenetically clustered. **C:** Statistics of the rDNA repeats analyzed.



## 2.4 Association of ribosomal DNA with the nuclear matrix

In addition to DNA methylation, ribosomal DNA may also be regulated by large-scale spatial chromatin organization and the distribution of active and inactive rRNA genes in the nucleus. For example, parts of the ribosomal DNA were found in the nuclear matrix, which might indicate a mechanisms for such a separation. Previous studies on rDNA chromatin regulation revealed the role of the Nucleolar Remodelling Complex (NoRC) in nucleosome positioning, transcriptional repression, epigenetic silencing and replication timing and the large subunit Tip5 is a key regulator for this. In addition, Tip5 harbours a TAM domain and several AT-hooks that are predicted MAR binder motifs. It is therefore hypothesized that Tip5 may function in the recruitment of rDNA to the NM, finally regulating large-scale chromatin reorganization of rDNA as well.

The first requirement for Tip5 mediating ribosomal DNA binding to the Nuclear Matrix is, that Tip5 can be found in NM isolations. This isolation of the NM fraction and the subsequent protein analysis by immunoblotting was performed. The distribution of Tip5 in the fractions of nuclear matrix preparations was monitored. Whole-cell extracts of HEK293 human embryonic kidney cells were fractionated and the resulting cytoplasmic, chromatin, high-salt wash and nuclear matrix fractions were analyzed by immunoblotting. The localization of lamin A/C in the matrix fraction,  $\alpha$ -tubulin in the cytoplasmic fraction and large and small amounts of histone proteins in the chromatin fraction and wash fraction, respectively, served as controls for the nuclear matrix preparations. The results clearly showed that two pools of Tip5 co-exist in the cell. These pools were found in the chromatin and nuclear matrix fractions, where the majority of the protein is located. In contrast, other chromatin remodelling complex subunits, i.e. Brg1, Snf2h and Mi-2 appeared preferentially in the chromatin fraction (out of [166]). Moreover, the distribution of Pol I in the different fractions demonstrated that not all nucleolar transcription factors are concentrated in the nuclear matrix <sup>3</sup>.

Therefore, the localization of Tip5 in the NM does not seem to be a general mechanism for nucleolar proteins or chromatin remodellers, but seems to be specific for the protein Tip5.

### 2.4.1 Tip5 targets rDNA to the Nucleolar Matrix

The fact that Tip5 contains a number of DNA binding domains that potentially bind to MAR sequences, and that the majority of the protein is present in the nuclear matrix fraction suggested that Tip5 could be involved in the nuclear matrix targeting of rDNA. To test this hypothesis, the relative amounts of various rDNA fragments in isolated nuclear matrix fractions of Tip5 transfected and control Hek293 were quantified and compared to the level of the IFN $\beta$  promoter, which is a *bona fide* MAR [13], containing a well-characterized binding site for the AT-hook protein HMGA1 [58].

---

<sup>3</sup>experiment performed by Attila Németh

First, the putative MARs of the human rDNA were determined *in silico* by using a formerly developed web-tool ([69]). Predicted MARs localize to the intergenic spacer (IGS) of rDNA as shown in Figure 2.33A. Real-time qPCR reactions were established to quantify the amount of one selected rDNA IGS sequence (IGS) that is localized between two predicted neighbouring MAR, as well as two additional rDNA regions, which are not predicted MARs (indicated in Figure 2.33A). One of these sites, the rDNA promoter (Prom), is a known binding site of Tip5. Tip5 possesses four AT-hooks and a TAM domain and therefore potentially targets its binding sites to the nuclear matrix. Another sequence was selected from the rDNA coding region (28S) where no Tip5 binding occurs [120]. Thus, our experimental system allows the monitoring of MAR- and Tip5-dependent and -independent associations of rDNA sequences with the nuclear matrix (Figure 2.33).

The immunoblot results show that Tip5 was strongly overexpressed 72 hours post-transfection (Figure 2.33B). In addition, similar amounts of purified nuclear matrix template DNA were analyzed from mock- and Tip5-transfected Hek293 cells in quantitative real-time PCR reactions. Threshold cycle (Ct) differences between mock and transfected cells were determined at each of the three different regions of the human rDNA and normalized to the Ct differences of the IFN $\beta$  promoter (Figure 2.33C). The results of three biological replicates are shown and demonstrate that rDNA is accumulating in the nuclear matrix upon overexpression of Tip5. Irrespective of the predicted MAR regions the whole rDNA unit is sequestered, suggesting a separation mechanism of individual rDNA copies to the nuclear matrix (out of [166]).

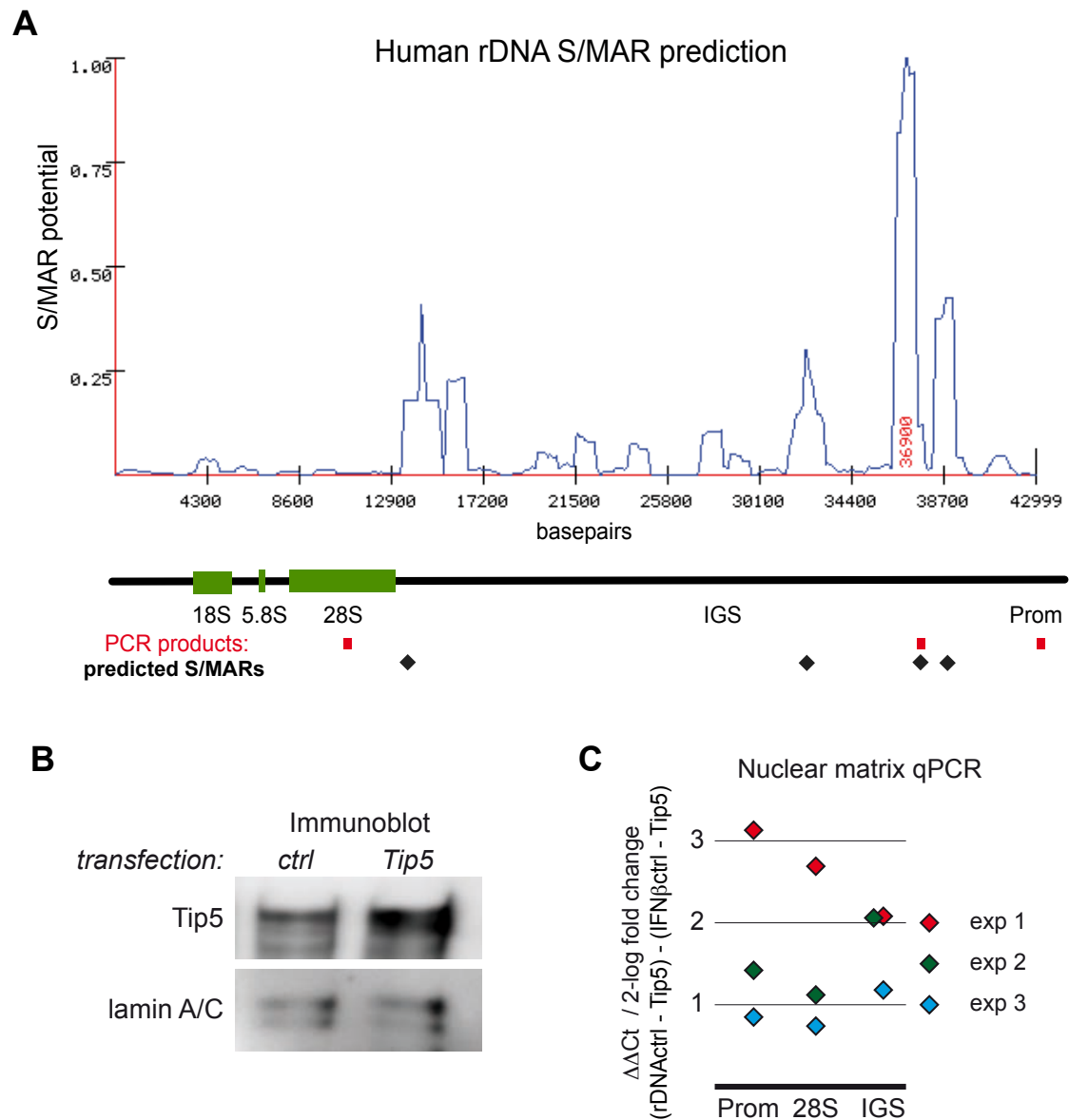
## **2.4.2 AT-hooks of Tip5 display MAR binding with comparable affinities**

### **2.4.2.1 DNA binding features of potential MAR binding domains of Tip5**

Tip5, the large subunit of NoRC, contains a variety of nucleic acid binding domains, e.g. AT-hooks and the TAM domain (2.34A), which were proposed to bind MARs [2]. To begin with the functional characterization of Tip5's potential MAR binding domains, DNA binding assays were performed. The DNA binding properties of the TAM domain has been analyzed in a previous study our working group [136], however, the four AT-hooks remained to be investigated.

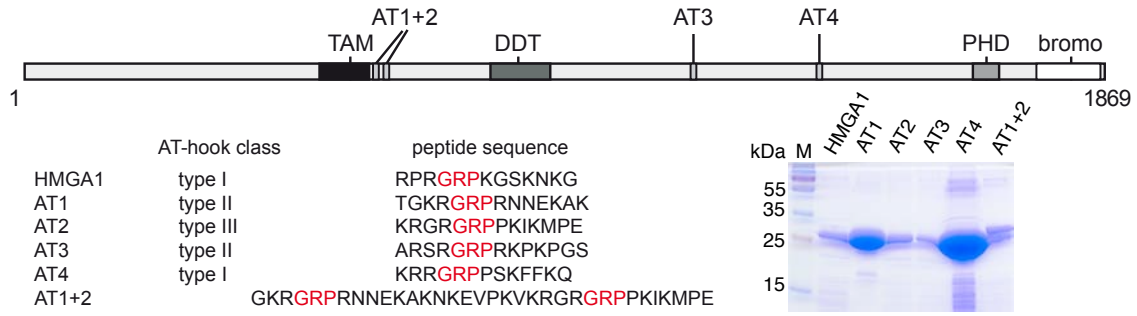
Thus, the four individual AT-hooks and the combination of the first two AT-hooks of Tip5 were expressed and purified as GST-tagged recombinant proteins (see Figure 2.34), and subjected to binding assays. The well characterized second AT-hook of the HMGA1 protein [58] served as a control in the DNA binding assays (abbreviated as HMGA1 below) (purification protocol can be found in 4.2.2.4) and GST alone as negative control.

Two AT-rich sites from the rDNA IGS were selected in addition to the previously characterized HMGA1 binding site of the interferon beta (IFN $\beta$ ) promoter, and the



**Figure 2.33: Overexpression of Tip5 leads to enrichment of rDNA in the NM**

**A:** The scheme of the rDNA repeat is shown at the top, 18S, 5.8S, 28S and IGS mark the coding regions and the intergenic spacer of the human rDNA (GenBank AccNo: U13369), respectively. Brackets and abbreviations indicate the location of rDNA regions analyzed in qPCR reactions (Prom - promoter, 28S - rRNA coding region, IGS - intergenic spacer). Predicted MARs are labelled with hexagons. Red rectangles show amplified regions by qPCR. **B:** HEK293 cells were transfected with full-length, Flag-tagged Tip5 for 72h and cellular extracts were prepared and analyzed by immunoblotting. Lamin A/C served to normalize for differences in loading. **C:** Real-time qPCR analysis of changes in the nuclear matrix association of rDNA after Tip5 overexpression. Similar amounts of purified nuclear matrix template DNA were subjected real-time qPCR from mock- and Tip5-transfected cells. Ct differences between Tip5-overexpressing cells and control cells were determined at the three different human rDNA regions and normalized to the Ct differences of the IFN $\beta$  promoter. The results of three biological replicate experiments are shown in different colours.



**Figure 2.34: GST-purification of the AT-hook peptides of Tip5 and HMGA1**

**A:** Domain structure of Tip5, the large subunit of the NoRC complex. Abbreviations: TAM - Tip5/ARBP/MeCP2 domain, ATx - AT-hook, DDT - domain in different transcription and chromosome remodelling factors, PHD - plant homeodomain zinc finger, bromo - bromodomain. The numbers below the scheme indicate amino acids. **B:** The list of the analyzed AT-hooks is shown below the diagrams. GRP core motifs of the peptides are indicated in red. The classification of AT-hooks according to Aravind and Landsman [2] is shown **C:** Coomassie gel picture of the GST-purification of the AT-hook peptides used for binding analyses.

DNA binding properties of the purified AT-hooks were tested by gel retardation assays.

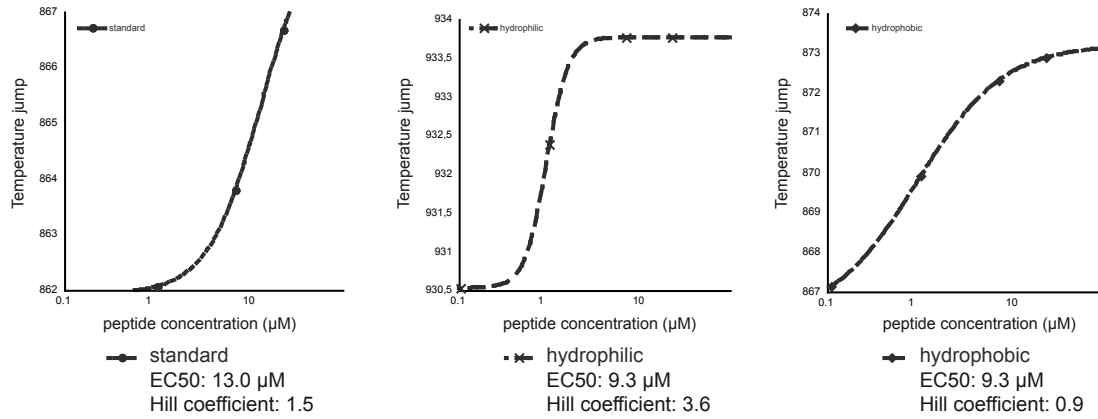
These experiments showed that: i) all AT-hook domains are *bona fide* DNA binding elements, ii) the individual AT-hooks bound the different sequences with similar affinity, iii) more than one AT-hook molecule bound to one DNA molecule in the case of the individual AT-hooks, as indicated by the supershifts, iv) there was only one protein-DNA complex in the case of the larger, double AT-hook AT1+2 protein. This suggests that in the double AT-hook construct both AT-hooks contact the short, 34 base pair DNA fragments, not leaving space for an additional AT-hook to bind this DNA <sup>4</sup>.

To continue with the characterization of Tip5's potential MAR binding domains and to determine the MAR binder with highest affinity, AT-hook-DNA interactions were quantified in microscale thermophoresis experiments. This novel method enables the quantification of DNA binding affinities in solution, however, needed to be established in the lab first. For starting, the second AT-hook of the protein HMGA1 was tested as control with an AT-rich sequence of ribosomal DNA enhancer and the protocol was performed as described in 4.2.8. The concentration of the fluorescently labelled DNA oligonucleotide should be close to the expected  $K_D$  and in the range of 80-1,500 fluorescence counts. Dilutions of the unlabelled peptide should start at a concentration about 20-fold higher as the expected  $K_D$ , being diluted to concentrations of about 0.01-fold of the expected  $K_D$ .

Since there are three different capillary coatings. the best one for AT-hook and MAR binding needs to be determined in the first experiment (see figure 2.35). The binding reaction contains 50 nM fluorescently labelled DNA with varying protein

<sup>4</sup>experiment performed by Katrin Rachow

concentrations as indicated and essentially performed as described in 4.2.8, however, no NP40 was included in the MST buffer.



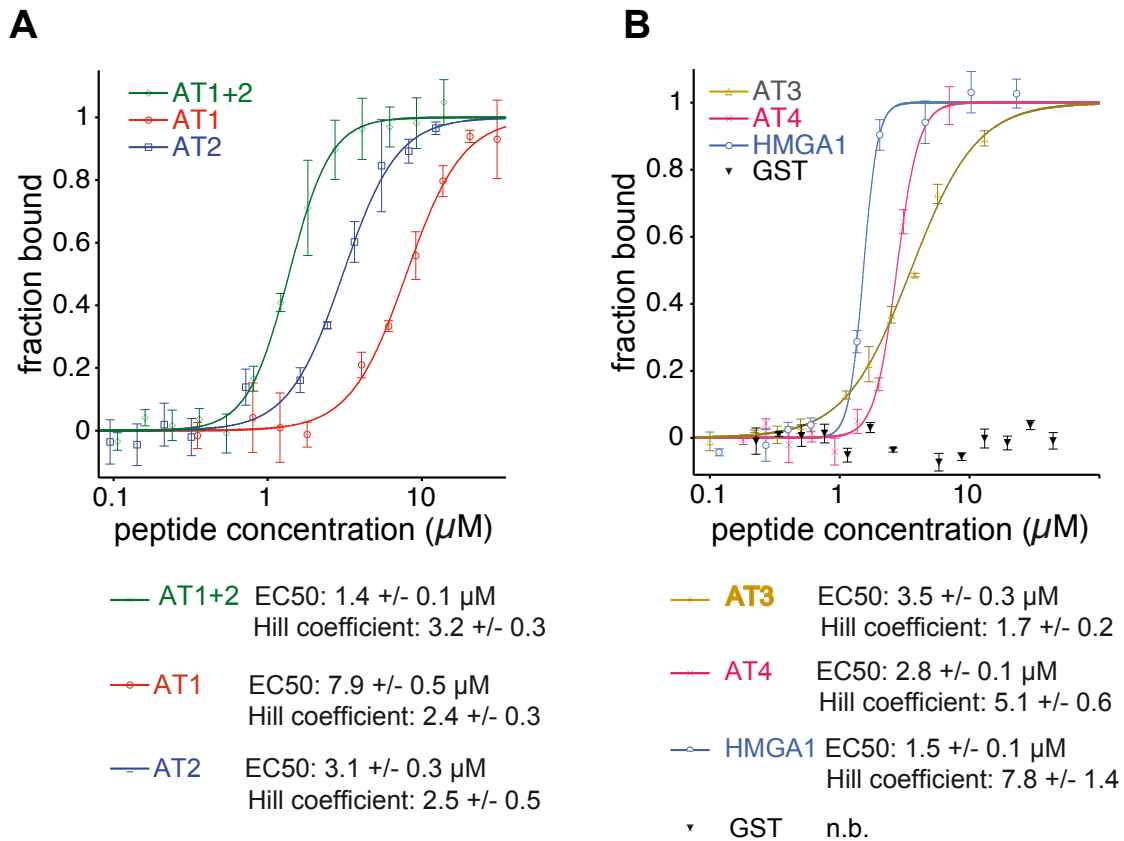
**Figure 2.35: Binding behaviour of HMGA1 to ribosomal DNA in three different capillary types determined by Microscale thermophoresis**

**A:** Binding affinity shown as plot of fraction bound against the peptide concentration in standard capillaries **B:** Same as A, but with hydrophilic capillaries **C:** Same as A, but in hydrophobic capillaries **A-C:** Hydrophobic and hydrophilic capillaries show similar binding affinities, however different binding behaviours. Standard capillaries exhibit weak binding compared to the two other types.

Figure 2.35 shows initial experiments on the different binding behaviours of HMGA1 resulting from different capillary coatings. Standard capillaries show only weak binding affinities in the conditions tested compared to hydrophilic and hydrophobic ones. Hydrophobic and hydrophilic capillaries show same binding affinities (Figure 2.35 C and B, respectively) because half of the present DNA molecules are bound at the same peptide concentration (EC50: 9.3  $\mu\text{M}$ ). The Hill-equation  $1/(1 + (EC50/c)^n)$  was used for fitting since the slope of the binding curve was too steep ( $n > 1$ ) or too smooth ( $n < 1$ ) for using a law of mass action fit. A Hill-coefficient  $n < 1$  or  $n > 1$  can indicate that the binding reaction is more complex than one molecule A in one conformation binds to one molecule B in one conformation. When the Hill coefficients of hydrophilic and hydrophobic capillaries are compared, a difference in the binding behaviour can be observed since hydrophilic capillaries show a greater Hill coefficient indicated by the steep binding curve. Therefore, hydrophobic capillaries present the best binding properties for AT-hook binding to MAR DNA and will be used in further experiments. Furthermore, to prevent unspecific molecule binding to the glass surface, NP40 is included in the buffer (see 4.2.8), which will presumably enhance binding to DNA and thus, lower the EC50 value.

### Tip5 AT-hooks bind to AT-stretches of ribosomal DNA with different affinities

Since the experimental system of microscale thermophoresis with AT-hook binding to MARs is established, the binding affinities of Tip5 AT-hooks to one part of the ribosomal DNA can be analyzed. Measured were thermally induced kinetics of a fluorescently labelled AT-rich site from the rDNA IGS incubated with a serial dilution of the different AT-hooks (see Figure 2.36 as described in 4.2.8. The analysis of the



**Figure 2.36: Binding of AT-hooks on ribosomal DNA measured by Microscale thermophoresis**

The indicated amounts of the different Tip5 AT-hook peptides were analyzed for DNA binding at constant, 50 nM concentration of a 29 bp, Cy5-labelled rDNA IGS oligonucleotide. The amount of bound DNA was plotted against the peptide concentration and the concentrations where half of the oligonucleotide is bound, the EC50 values, were determined. The data points of the plot represent standard deviation values of four measurements. The numerical EC50 values for each peptide are shown below the diagram. **A:** Binding affinities for AT1, AT2 and AT1+2 **B:** Binding affinities for AT3, AT4 as well as HMGA1 as controls.

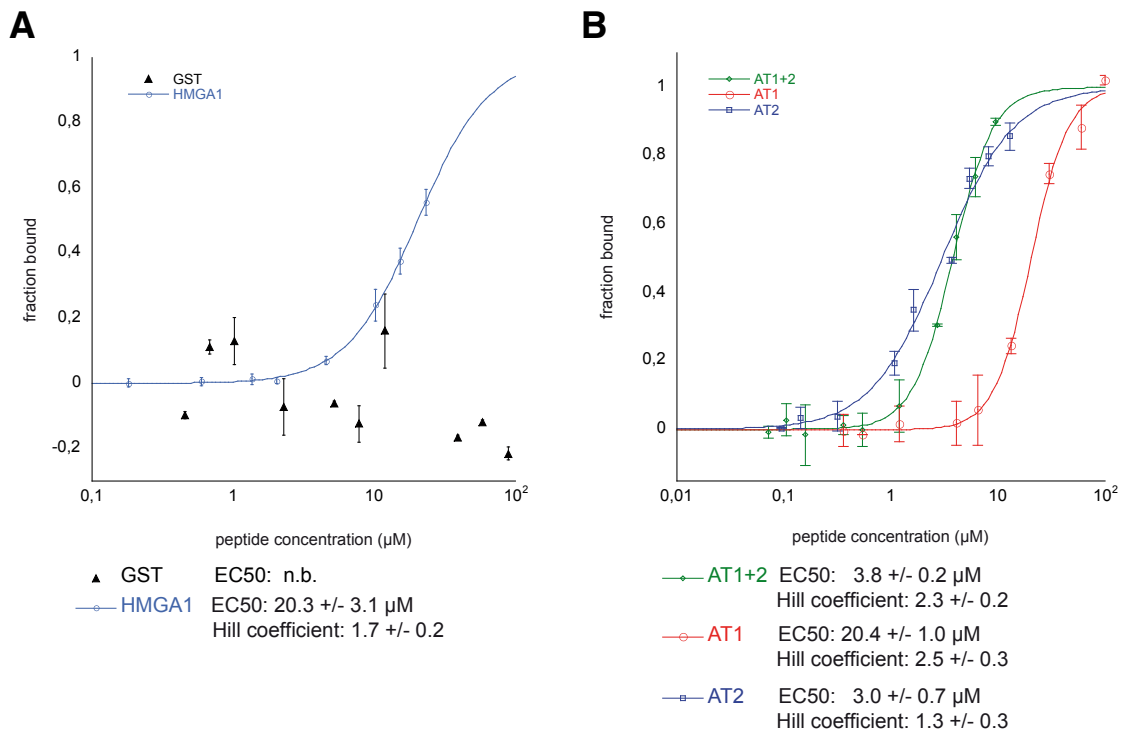
normalized thermophoresis curves provided the equilibrium constant concentration values for each AT-hook, when 50% of the DNA was bound by the protein (EC50). The binding constants exhibit clear differences between the individual AT-hooks, displaying slightly weaker affinities than the HMGA1 control. The EC50 value of the double AT-hook AT1+2 is higher than the EC50 of the individual AT-hooks,

suggesting that both domains contact DNA simultaneously and revealing a binding affinity similar to HMGA1(out of [166]).

### Single-binding site interaction of AT-hooks with the PRDII element of the IFN $\beta$ enhancer shows decreased binding affinities

In previous microscale thermophoresis experiments, the Hill-equation  $1/(1 + (EC_{50}/c)^n)$  was used for fitting since the slope of the binding curve was too steep ( $n > 1$ ) or too smooth ( $n < 1$ ) for fitting with a law of mass action. One possibility is that the used double stranded oligonucleotides are long enough to have more than one binding site for AT-hooks, thus displaying a complicated binding pattern. To eliminate this possibility, a 12 nucleotide short, AT-rich oligonucleotide is ordered that provides one binding site only. This oligonucleotide is a DNA dodecamer comprising a five base pair A, T tract from the PRDII element of the IFN $\beta$  enhancer element (described in [58]) and further named PRDII.

The binding experiments were performed as described and the binding properties are displayed in Figure 2.37. There is an overall decrease in the binding affinities



**Figure 2.37: Binding behaviour of different AT-hook peptides on PRDII measured by Microscale thermophoresis**

**A:** Binding affinity of HMGA1 in contrast to GST, where no binding occurs **B:** Binding properties of AT1, AT2 and AT1+2 showing different binding activities, with AT1 having the lowest affinity to DNA. Error bars indicate standard deviation of four technical replicates. EC50 value and Hill-coefficient is written below the diagrams for each peptide

of tested AT-hooks observed when binding to the short PRDII oligonucleotide is compared to the AT-stretch of ribosomal DNA. In addition, the Hill coefficients

came closer to 1, nevertheless, fitting the data with the law of mass action was still not possible. Therefore, the other AT-hooks were not further analyzed.

In summary, all tested AT-hooks show DNA binding affinity, however to a different extent. The highest affinity is displayed by the double AT-hook AT1+2, followed by similar binding affinities of AT2, AT3 and AT4. In contrast, AT1 shows the lowest DNA binding affinity. This may be the reason why no cooperative effect is seen in the double AT-hook, since the contribution of AT1 is low compared to AT2.

#### 2.4.2.2 Sequestering of rDNA to the nucleolar matrix requires a functional Tip5 molecule

After identifying the double AT-hook domain as the strongest MAR binder, mutation studies were performed. The nuclear matrix association of this protein domain was investigated in transient transfection experiments where a wild type and a mutant version of the double AT-hook domain was fused to GFP resulting in the GFP-AT1+2wt and GFP-AT1+2mut constructs <sup>5</sup>. In the latter one the RGR core motifs of both AT-hooks were mutated to the DGD tripeptide that was previously shown to loose DNA binding activity ([17]). Surprisingly, nuclear matrix analyses of cellular fractions showed that despite the *in vitro* MAR binding activity, the double AT-hook domain is not sufficient to mediate association with the nuclear matrix since the mutant peptide is not found in the Nuclear Matrix fraction <sup>6</sup>.

To test whether the other putative MAR binding domain, the TAM domain is required to mediate association with the nuclear matrix, the GFP-AT1+2wt and GFP-AT1+2mut constructs were extended with this domain resulting in the GFP-TAM-AT1+2wt and GFP-TAM-AT1+2mut constructs. Interestingly, nuclear matrix analyses of cellular fractions revealed that the TAM domain is necessary and sufficient to mediate the association of the GFP-TAM-AT proteins with the nuclear matrix <sup>7</sup>.

Transfections with different GFP-Tip5 peptides lead to overexpression of these peptides as shown by immunoblotting experiments (Figure 2.38A). Nevertheless, the enrichment of rDNA in the Nuclear Matrix as after overexpression of Tip5 could not be repeated in the different experiments (Figure 2.38B compared to Figure 2.33) <sup>8</sup>.

As a summary, even though the nucleolar matrix targeting domain could be assigned to the TAM domain, the sequestering of the rDNA to the nuclear matrix fraction requires a functional Tip5 molecule. DNA quantification after expressing similar levels of the double AT and TAM-AT hook constructs did not exhibit an increase of rDNA in the matrix (see Figure 2.38), as shown for the Tip5 protein, illustrating that the TAM domain is required for nucleolar matrix targeting, however additional DNA binding domains in Tip5 are required for rDNA binding (out of [166]).

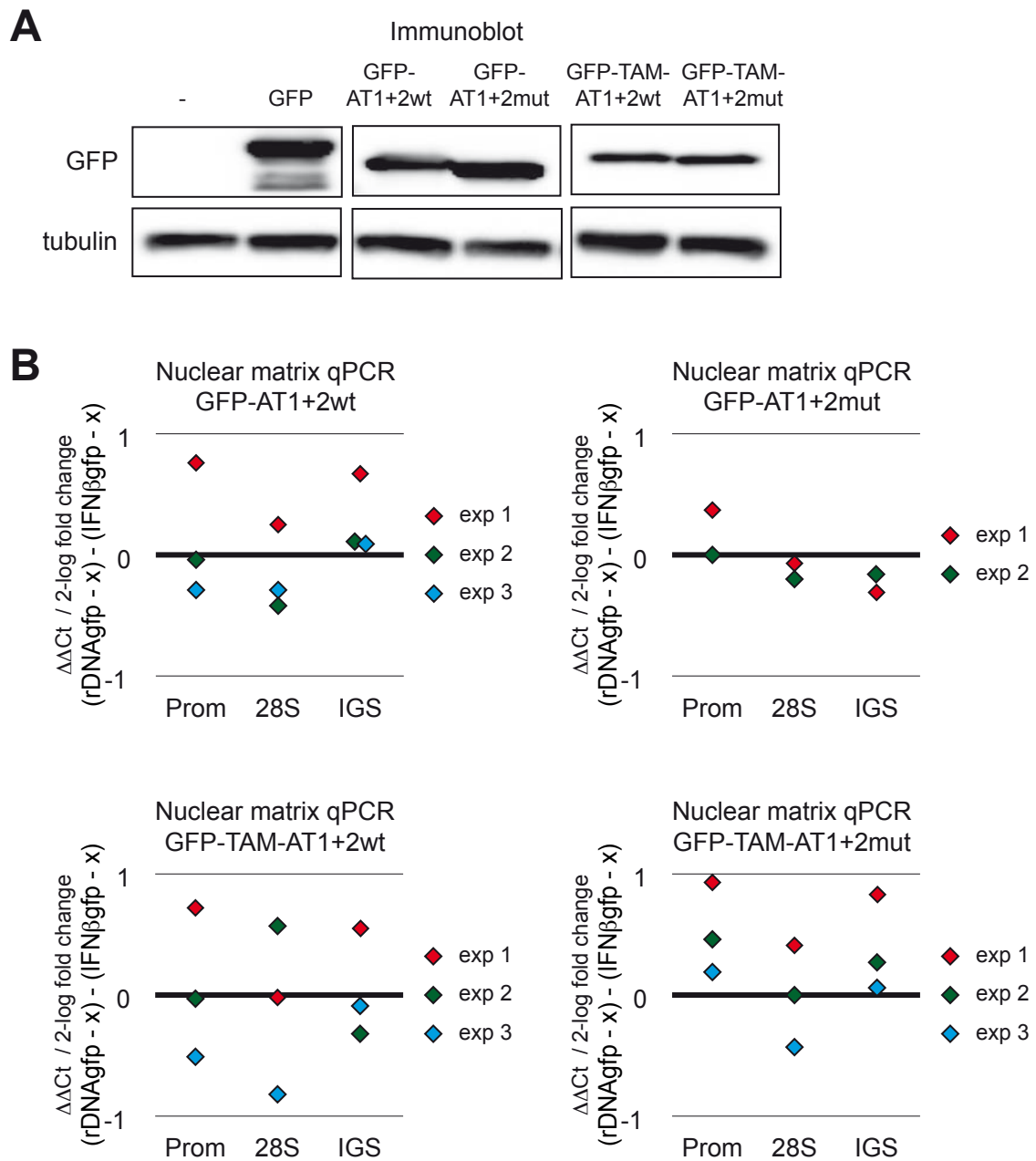
<sup>5</sup>Created and tested by Michael Filarsky

<sup>6</sup>Experiments performed by Michael Filarsky

<sup>7</sup>Experiments performed by Michael Filarsky

<sup>8</sup>Nuclear Matrix isolation and immunoblotting performed by Michael Filarsky





**Figure 2.38: Overexpression of different GFP-Tip5 peptides**

**A:** Immunoblot detection of the overexpression of different GFP-Tip5 proteins. HEK293 cells were transfected with GFP or the indicated GFP-Tip5 proteins for 72h and cellular extracts were prepared and analyzed by immunoblotting. Tubulin served to normalize for differences in loading. **B:** Real-time qPCR analysis of changes in the nuclear matrix association of rDNA induced by overexpression of different GFP-Tip5 peptides. For details about the experimental set up see Fig. 1B and the text. The results of three biological replicates are shown for GFP-AT1+2wt, GFP-TAM-AT1+2wt and GFP-TAM-AT1+2mut, and two biological replicates for GFP-AT1+2mut. Coloured squares indicate individual experiments.

## 2.5 Extended AT-hooks as a novel DNA binding motif

In addition to the four AT-hooks of Tip5 (AT1, AT2, AT3 and AT4) identified by SMART (<http://smart.embl-heidelberg.de/>), another GRP tripeptide motif, the core motif of the AT-hook, was found in the amino acid sequence of the protein. The core motif of AT-hooks is flanked by basic amino acid patches, however in this case, they are located throughout a larger distance to the GRP motif, thus it was termed extended AT-hook (eAT-hook) of Tip5.

Remarkably, a search for (K/R)>3-N6-20-GRP-N6-20-(K/R)>3 sequence motifs in the human and mouse proteomes delivered 81 and 60 different extended AT-hook peptides, respectively (Figure 2.39A). A small proportion of them (19 and 17 peptides in human and mouse, respectively) were identified also as canonical AT-hook, however, most of them could not be recognized by the standard AT-hook search algorithm.

To test the DNA binding ability, a full (eAT) version of this AT-hook as well as a short (eATdNdC) and a one-side-extended (eATdN) version were constructed, expressed and purified to test the impact of the basic amino acid patches in the extended regions on DNA binding. Figure 2.39B shows the sequence of the eAT-hook constructs with the GRP core motif in red and the basic amino acid patches in a larger distance. After the peptides were purified, they were subjected to Microscale thermophoresis to test the DNA binding capability of the putative DNA binding motif. As interaction partner, the AT-stretch of ribosomal DNA IGS is used to compare the binding affinity with the binding of the other canonical AT-hooks of Tip5.

Microscale thermophoresis assays prove that only the full length version of the extended AT-hook is able to bind DNA, but not the truncated versions. Remarkably, the affinity of DNA binding is comparable to the binding properties of the canonical AT-hook AT1 of Tip5. In addition, the large number of such sequence motifs, the conservation between mouse and human and the similar molecular function of the proteins carrying them suggests that they possess the function of canonical AT-hook motifs.

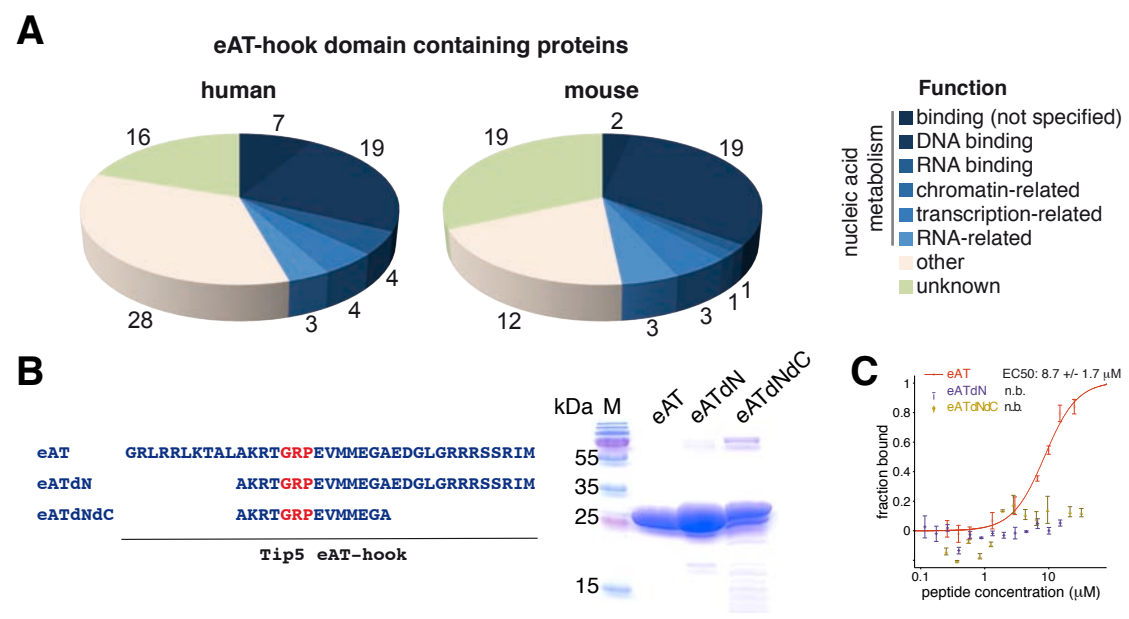


Figure 2.39: Identification of a novel DNA binding motif

**A:** Conservation of eAT hooks in mouse and human proteins by GO search. Pie diagrams of molecular functions of human and mouse proteins that hold extended AT-hooks. The number of proteins containing extended AT-hooks is indicated for each category. Blue colours indicated protein functions in nucleic acid metabolism. **B:** Sequence of different eAT hook constructs of Tip5 and their GST-purification shown in a SDS-PAGE and Coomassie staining. The GRP core motif is illustrated in red. **C:** Microscale thermophoresis experiment with different eAT-hook constructs as described in 4.2.8.

## 3 Discussion

### 3.1 Single molecule epigenomics by Dynamic molecular combing

One part of this work was establishing variants of the Dynamic molecular combing method, which are called psoralen-combing, chromatin-combing and epi-combing.

#### 3.1.1 Psoralen-combing for Sequence-specific Single Molecule Chromatin Analysis

Psoralen-crosslinked DNA has been investigated with a variety of analytical methods including electron microscopy [23, 22], restriction enzyme digestion followed by Southern blot analysis [132], *in vitro* and *in vivo* footprinting [108, 164] also with single nucleotide resolution [[68], nucleosome positioning by primer extension [152, 151] and chromatin endogenous cleavage (ChEC) [44, 47, 88, 155]. Additionally, psoralen photo-crosslinking of DNA has been used recently for the genome-wide analysis of DNA helical tension, a feature of chromosomes that cannot be characterized by other experimental methods [8].

#### Proof-of-principle experiments

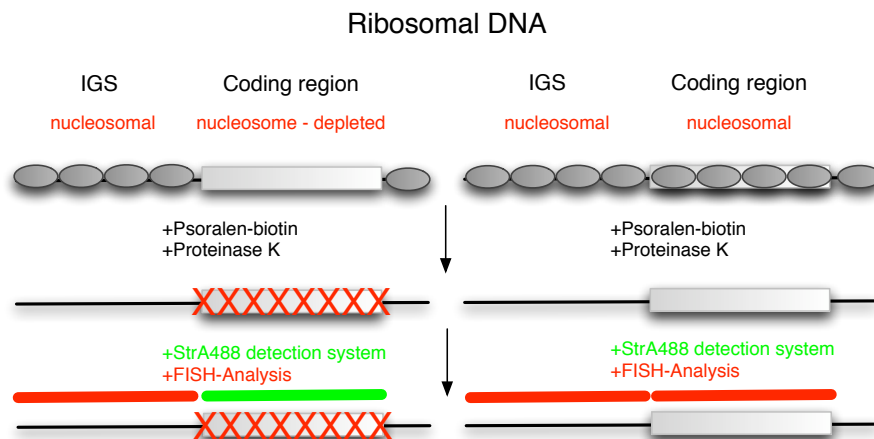
The psoralen-combing assay applies the dynamic molecular combing method on psoralen crosslinked DNA. The technique allows sequence-specific, genome-wide single molecule analysis of chromatin structures based on their psoralen accessibility, at the megabase scale with kilobase resolution. The combination of psoralen photo-crosslinking with subsequent dynamic molecular combing analysis of stretched DNA fibers is presented. The proof-of-principle experiments have been performed on  $\lambda$  DNA, which was subjected first to psoralen treatment and photo-crosslinking with a UVA, blue light lamp at 315-400 nm. The DNA has been stretched on silanized coverslips and visualized with fluorescent hybridization either with or without reversal of the psoralen crosslinking at 254 nm UVB light. The reversible psoralen-dependent masking of the signals demonstrates that the method can be potentially applied for sequence-specific, single molecule characterization of chromatin structures. It was further demonstrated that the use of biotinylated psoralen enables the direct visualization of photo-crosslinked DNA (out of [169]). Finally, the detection of psoralen crosslinked DNA was proven for the case of native chromatin of HCT116 cells. However, detection of specific DNA loci by fluorescence hybridization is still missing and will be the scope of future experiments.

In contrast to electron microscopy, which was the only single molecule method to investigate psoralen-treated chromatin so far, psoralen-combing may work in a sequence-specific fashion. In addition, this approach provides the possibility to monitor more genomic loci in parallel and perform genome-wide analyses of open, psoralen-accessible chromatin structures at kilobase resolution. Thus, psoralen-combing extends the continuously growing list of single molecule methods to analyze the epigenome [168].

### Potential of rDNA analysis by psoralen-combing

The use of psoralen crosslinking in the analysis of the ribosomal DNA, improved extraordinarily our understanding about the chromatin structure of this particular genomic region. Ribosomal DNA exist in multiple copies, however, the repeats are present in two distinct chromatin structures since not all repeats are active. The inactive rDNA is characterized by nucleosomal DNA whereas active rDNA is devoid of nucleosomes or possess dynamic nucleosome structure (reviewed in [85]). The psoralen photocrosslinking of rDNA with subsequent restriction enzyme digestion and Southern blot analysis is a unique tool to characterize nucleolar chromatin as it exposes the two coexisting rDNA chromatin structures. This is based on the different crosslinking of the DNA from active and inactive gene copies, which results in an altered migration behaviour in agarose gel electrophoresis [28, 132] (out of [169]).

Psoralen-combing may be a novel tool to distinguish those two classes of rDNA repeats at the single molecule level (see figure 3.1).



**Figure 3.1: Discrimination of open and closed rDNA chromatin by psoralen-combing**

Native genomic DNA will be subjected to treatment with psoralen, before the DNA fibers will be combed. Active, nucleosome-depleted repeats will be crosslinked with psoralen-biotin and detected by StrA488. Due to the psoralen:DNA adducts, further detection by fluorescence hybridization is abrogated. Nucleosomal DNA, on the contrary, does not have intercalated psoralen-biotin which makes detection of the rDNA repeat possible by fluorescence hybridization.

Psoralen-treated DNA is further analyzed with the dynamic molecular combing

method that provides special information about rDNA clusters. The high complexity of the arrangement pattern of clustered repetitive DNA, the human rDNA, was uncovered for the first time by using dynamic molecular combing [90], and this feature cannot be assessed by other analytical methods. Therefore, this assay is well suited for the investigation of the arrangement and rearrangement of single copy and repetitive chromosomal regions.

### 3.1.2 Epi-combing

DNA modifications represent an integral part of the epigenome and they have a pivotal role in regulation of genome function. Despite the wide variety of analytical techniques that have been developed to detect DNA modifications, their investigation at the single genome level is only beginning to emerge. In contrast to population-averaged analyses, single molecule approaches potentially allow the mapping of epigenetic linkage between distantly located genomic regions, the locus-specific determination of the DNA modifications of repetitive DNA elements as well as the analysis of allelic epigenomic differences.

#### **Epi-combing complements the analysis of the epigenome on a single molecule level**

In this work, epi-combing is described as a novel sequence-specific single molecule method to analyse DNA modifications by combining Dynamic molecular combing with DNA modification immunodetection. High molecular weight DNA in solution is aligned and stretched by a receding air-water meniscus and fixed on a glass surface. Tens to hundreds of genomes can be immobilized on one glass slide allowing analysis of thousands of DNA molecules in a single experiment. Specific antibodies detect 5mC as well as 5hmC and the DNA locus of interest is visualized using fluorescently labelled hybridization probes. Analysis is done manually with a conventional fluorescence microscope or automated with a high-resolution fluorescence scanner.

Proof-of-principles were performed on native and *in vitro* modified  $\lambda$  DNA and these experiments demonstrate that the detection of DNA modifications on single DNA fibers is specific, sensitive as well as selective. In addition, experiments show the discrimination between different DNA modifications (5mC and 5hmC), which cannot be distinguished by bisulfite treatment, a method commonly used to investigate DNA methylation.

Furthermore, it can be applied for the investigation of DNA methylation patterns at large, highly repetitive regions of genomic DNA, as well as for epigenetic linkage analysis of kilobase-sized, modified DNA bearing chromosomal regions. Remarkably, a barcode fluorescence labelling may enable multiplex, parallel analysis of different genomic regions in a single experiment.

A technique similar to the epi-combing has been already invented for the epigenomic analysis of bacterial genomes and it was successfully tested on human genomic DNA, too. This approach uses optical mapping to analyze DNA methylation patterns [1]. High molecular weight DNA fibers are arrayed on a positively charged glass surface

using a microfluidic device. Then, the molecules are digested with restriction enzymes and visualized by a fluorescent dye. Image acquiring and processing is fully automated and converts the DNA signals to a molecular barcode system. These barcodes are based on the distances between different restriction enzyme cleavage events and thus, represent a unique identification for each individual molecule. Detection of DNA methylation patterns is possible by using one methylation-sensitive restriction enzyme and comparing the resulting pattern with a theoretical *in silico* digestion pattern. Missing cutting events are caused by DNA methylation at the respective site. This approach was further developed for the human genome by combination of two enzymes, one being methylation insensitive and one being sensitive to methylation. The first one was used as a barcode tag to find the location of the genome, whereas the latter one to identify the methylated recognition sites. In summary, the method offers genome-scale DNA modification profiling on unmodified and unamplified DNA. Although the technique is feasible for repeat regions, the analysis is restricted to sites where a restriction enzyme has a recognition sequence (out of [168]). In addition, possible sequence variations at the restriction site further complicate the analysis of the signals.

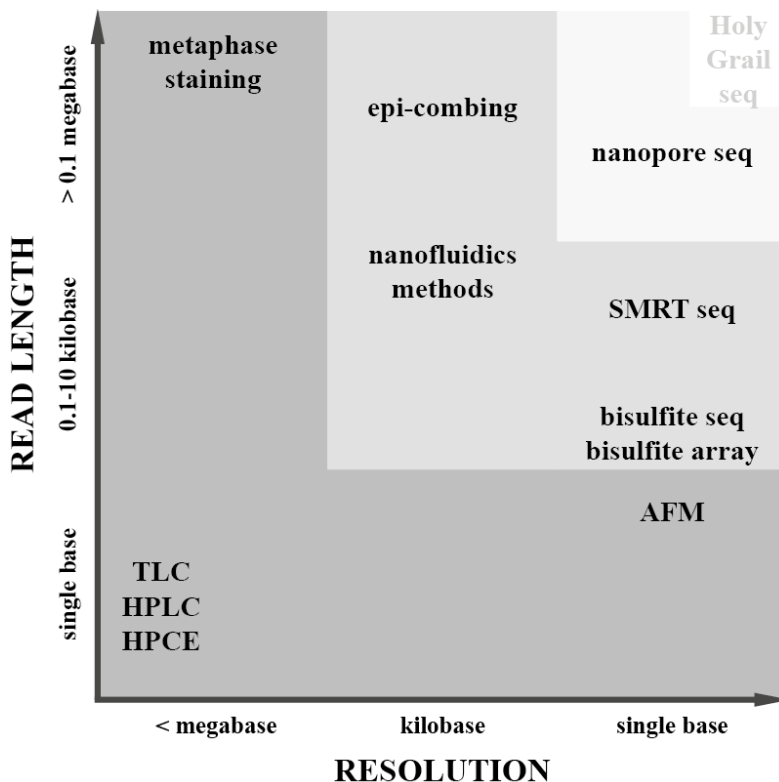
In contrast, epi-combing does not require restriction digestion and in theory, sequence-specific detection of DNA modifications is feasible at any part of the genome. Epi-combing belongs the group of recently emerging single molecule techniques, which assess DNA modifications using different visualization methods of DNA polymers without sequencing and microarray hybridization, also among them atomic force microscopy and nanofluidics approaches. In contrast to other methods, epi-combing was successfully tested on human genomic DNA. Thus, the features of epi-combing with a kilobase resolution and a megabase read length complements the so far existing single molecule DNA methylation analysis technologies, as can be seen in figure 3.2.

In summary, epi-combing is a new analysis tool, which helps to answer so far unresolved questions concerning DNA modifications at repeats, individual DNA strands, and individual alleles and to provide data for epigenetic linkage mapping and completes the list of methods that are able to capture various DNA modifications of individual, native DNA molecules and provide sequence-specific information at the genome scale.

### **Potential biomedical applications of single molecule epigenome analyses**

In recent years, elucidating the genetic and non-genetic determinants of human diseases was a major challenge in biomedical research. In the future, sequencing of whole genomes will shed light onto the complex interplay of genetic variants leading to diseases. However, not all common human diseases have genetic reasons but studies show that epigenetic factors are involved in a large proportion of them. Normal cellular function relies on correct epigenome function and many reports show the causality between epigenetic factors and diseases [37, 107, 111, 154, 37]. To understand complex human diseases, epigenome-wide studies may deliver essential information.

Changes in the DNA methylation patterns are recognized as a common molecular



**Figure 3.2: Comparison of read length and resolution between different single molecule analysis methods**

The ideal DNA modification analysis method exposes the epigenome in one sequence read at single base-pair resolution (Holy-Grail Sequencing). Methods such as thin layer chromatography (TLC), high-performance liquid chromatography (HPLC) and high-performance capillary electrophoresis (HPCE) work on hydrolyzed DNA and therefore reveal only the overall genomic level of modified bases. Metaphase chromosome spreads can be specifically stained with antibodies against 5-methylcytosine or 5-hydroxymethylcytosine, however they show DNA modification patterns only at low resolution. Nanofluidics and epi-combing techniques provide information about DNA modifications at kilobase resolution on 0.01-1 megabase sized DNA molecules. Atomic force microscopy (AFM), bisulfite sequencing and bisulfite array methods, single molecule real-time (SMRT) sequencing and nanopore sequencing operate at single base resolution and deliver information about the modifications of DNA molecules in the size range of several ten, hundred, thousand or several ten thousand base pairs, respectively. (out of [168])



hallmark of human neoplasia [5]. Hypermethylation of CpG islands in the promoter region of tumour suppressor genes is the most frequent mechanism of gene inactivation in cancers. The single molecule analysis of the epigenetic linkage of such regions could thus promote better tumour classification, which in turn may advance therapeutic intervention. For example, the promoter regions of the p14ARF and p16INK4a tumour suppressor genes are separated by approximately 20 kilobases. Both promoters are located in a partially methylated domain in normal diploid fibroblast cells [74], which means that population-based epigenetic analyses could determine the percentage of the methylation states of the two promoters. However, they could not tell whether in a single cell and on the same allele both promoters are identically modified or not, i.e. inactive or active. Though, this is the information, which will determine the cell's fate. An essentially similar situation with more partially methylated domains can be observed at the HOX gene cluster [74], where the promoter methylation states were also correlated with cancer [5]. In contrast to promoter-specific hypermethylation, global hypomethylation was also observed in cancer cells leading to increased gene expression of oncogenes. Therefore, it would be feasible to use a combination of already established as well as novel DNA modification markers, to identify patterns of them that aid more precise specification of different cancer subtypes as well as different stages of mutagenesis and better prediction of treatment effectiveness and patient survival (out of [168]). Epi-combing may contribute to this new field of personalized biomedicine by analyzing the epigenome of DNA methylation at tandem repeat arrays, which will shed light onto the epigenetic regulation of repetitive DNA.

## **3.2 DNA methylation patterns of tandem repeat arrays**

### **3.2.1 DNA methylation analysis of satellite-2 repeats**

Satellite-2 DNA is found in pericentromeric regions and its DNA methylation was found to be crucial for chromosome segregation [41, 109, 110, 159]. In the presented work, experiments were performed that analyzed the DNA methylation status of satellite-2 DNA by metaphase-immunoFISH and epi-combing. Metaphase-immunoFISH experiments were performed on male blood lymphocytes and revealed that strong signals of satellite-2 DNA on chromosomes 1,9, 15 and 16. In most cases, the signals overlapped with heavy DNA methylation, however some fluorescence hybridization signals do not overlap or do only partially overlap with strong DNA methylation.

Analysis by epi-combing yields a higher resolution and was done with the colon cancer cell line HCT116. The data demonstrate that stretches of satellite-2 DNA are heavily methylated with a great variety in length ranging from a few kilobases up to 200 kilobases. The average length of signals was calculated as 76 kb, however the real value may differ from the observed one. For example, it is known that satellite-2 blocks can be several megabases long and a reason why these long regions are not detected is that they break during the combing process. In addition, due to the resolution limit of epi-combing and the manual scanning process of signals,

satellite-2 repeats shorter than 1 kb may be missed and not taken into account for the calculation of the average length. However, it is proven that satellite-2 DNA is present in long stretches of DNA which seem to 100% methylated.

Hence, no DNA hypomethylation was observed in the colon cancer cell line as was reported from other cancerous cells such as breast cancer [62, 116, 141, 154]. There are two possibilities for this discrepancy: Either hypomethylation at satellite-2 repeats is not a common feature of all cancer cell lines (no data of HCT116 are available), or unmethylated repeats cannot be resolved with epi-combing. Satellite-2 DNA consists of repeats of a 23 - or 26mer of basepairs and the resolution limit of epi-combing is in the kilobase range. This means that single unmethylated repeats in a longer stretch of methylated DNA will be missed with epi-combing.

To solve this problem, a mixture of different methods such as bisulfite treatment-based 5mC analysis should complement the data obtained by epi-combing. Interestingly, a loss in methylation at satellite-2 arrays was observed in human fibroblasts that were cultured to replicative senescence [139]. Further analysis of satellite-2 DNA should therefore compare these results with senescent Imr90 that were cultured to replicative senescence and used for rDNA methylation analysis in the presented work.

Furthermore, epi-combing could reveal the DNA methylation patterns of satellite-2 DNA in ICF syndrome patients, where demethylation in these sequences is caused by mutations in the DNA methyltransferase Dnmt3b [53, 52, 54, 159].

### 3.2.2 DNA methylation analysis of ribosomal DNA

#### 3.2.2.1 DNA methylation clustering and transition of ribosomal DNA repeats

Ribosomal DNA is found on the short arms of the five acrocentric chromosomes and is clustered there in "Nucleolar organizer regions" (NOR). Previous experiments lead to the conclusion that NORs are either completely active, or completely inactive, and that this activity status is kept and propagated throughout the cell cycle (reviewed in [85]). In the presented work, NOR activity states were determined by metaphase-immunoFISH experiments. The data confirm that UBF and Pol I are marker for active NORs and that they colocalize in 86% of all NORs. However, the combination of UBF and DNA methylation detection in metaphase-immunoFISH experiments reveals a more complicated situation of activity states than active and inactive. The frequent colocalization of UBF with different DNA methylation levels on NORs indicates gradual shades of NOR activities. The most active NOR is composed of UBF and no (or low) DNA methylation levels. Then medium DNA methylation levels are found with UBF which represents the second most active NOR state. A further decrease in NOR activity is characterized by UBF staining and colocalization of heavy DNA methylation, which occurs in at least about one third of NORs in the tested male blood lymphocytes. Finally, inactive NORs do not carry UBF but have variable levels of DNA methylation.

These observations suggest that there are different states of NOR activity in a cell and the colocalization of DNA methylation and UBF indicates that not all repeats on a given NOR display the same epigenetic status. However, the resolution of

metaphase-immunoFISH is low and therefore, epi-combing is used to answer the question about epigenetic clustering at ribosomal DNA.

Different cell lines were subjected to the analysis such as the cancer cell lines HCT116 and MCF7 and various fibroblasts from different sources such as skin fibroblasts from healthy individuals (AG05283 and AG13077) and from patients with Werner syndrome (AG12797 and AG06300) as well as lung fibroblasts (Imr90) that were in addition cultured to replicative senescence.

In all tested cells, epigenetic clustering of adjacent repeats was observed as predicted from observations of entire NORs that are either active or inactive. In addition to the epigenetic clustering, transitions between epigenetic states of neighbouring repeats were also proven. This surprising result explains the observation of UBF and DNA methylation colocalizing at individual NORs.

Interestingly, a mosaic methylation pattern of ribosomal DNA was proposed, where the coding region is either completely methylated or unmethylated according to the activity status of the repeat, and an IGS that is always methylated to nearly 100% of all CpG sites [18]. In epi-combing experiments, however, the DNA methylation of the IGS is highly variable from being totally unmethylated (see figure 2.29C) to strongly methylated all over the repeat (see figure 2.28A). This discrepancy and the high level of DNA methylation of the IGS region although it is depleted of CpG sites suggests either sequence variation of individual IGS or non-CpG methylation or could also be caused by non-homogenous detection of 5mC during epi-combing. Furthermore, the total percentage of DNA methylation at ribosomal DNA was calculated according to epi-combing data obtained in this work. The value ranged from 44% in AG05283 up to 95% in MCF7. The high rDNA methylation level in the breast cancer cell line MCF7 is supported by data that demonstrate hypermethylation at ribosomal RNA genes in breast cancer [160], however, more repeats need to be analyzed to reveal the true value of rDNA methylation.

At this point, it is not possible to compare the methylation level data of the tested WS cell lines (AG05283 and AG13077) as well as tested primary fibroblasts (AG12797 and AG06300) with previous results, since no published data were available. Therefore, to verify the findings, a significant increase in the number of analyzed repeats is necessary, too. Regarding the fact that about 400 copies of rRNA genes exist in a diploid cell, at least 1000 repeats, which corresponds to 2.5 rDNA genomes, should be analyzed to reach a significant data volume.

In the case of HCT116, epi-combing revealed an overall methylation level at ribosomal DNA of 53%, which is in good agreement with the published value of 60% [35]. Further analysis with available Dnmt knock out cell lines of HCT116 could reveal insights into the epigenetic regulation of rDNA clusters by DNA methyltransferases. For Imr90 cells, a total percentage of 71% at the rDNA locus was observed using rDNA methylation analysis by epi-combing. Interestingly, MethylC-Seq was performed for Imr90 fetal lung fibroblasts in another study [75] and a bioinformatic analysis of the available data focussing on the rDNA locus could verify the results obtained in the presented work.

The rDNA methylation level of fibroblasts either cultured to replicative senescence (Imr90: 58%) or being pathologically senescent in Werner syndrome cells (AG12797: 82% and AG06300: 60%) was determined in a previous study [78]. In senescent WS and old fibroblasts, rDNA species with unusually slow electrophoretic mobility were detected after restriction endonuclease digestion using the enzyme EcoRI followed

by Southern blotting and hybridization to the 28S region of the rDNA locus [78]. This analysis revealed that some rDNA species of senescent cells and Werner syndrome cells were longer compared to predicted EcoRI fragments. The detection of longer fragments in the Southern blot analysis of genomic DNA from near senescent cells could be due to inhibition of restriction endonuclease (EcoRI) activity by DNA modifications (such as 5mC) at or near the EcoRI recognition sequence.

This observation is accompanied by significant increases in cytosine methylation of rDNA in *in vitro* senescence of fibroblasts. Generally, genomic DNA was digested with either HpaII or MspI and subjected to Southern blot analysis where the coding region of the rDNA locus was detected. For both MspI and HpaII digestions, the intensity of the lowest band was quantitated and a methylation index was defined as the intensity of this band in the MspI-digested sample divided by the intensity of the same band in the HpaII-digested sample. It was shown that replicative senescence results in increased methylation at the ribosomal DNA compared to low passage numbers and that this increased methylation compared to early passage numbers is maximal in WS fibroblasts [78], indicating accelerated methylation in WS patients. In this presented work, the rDNA methylation level of replicative senescent Imr90 fibroblasts slightly decreased (58%) compared to the overall methylation of low passage number cells. In contrast, the percentage of transitions increased from 11.5% in low passage number cells to 26.7% in senescent cells. It can be speculated that there is a link between decrease in DNA methylation and an increase in transitions. However, due to the small sample volume, more data of rDNA methylation and transitions in senescent fibroblasts from Imr90 cells as well as other fibroblast cell lines need to be collected to draw final conclusions.

### 3.2.2.2 Epigenetic status of non-canonical rDNA repeats

Non-canonical repeats of ribosomal DNA were first observed by Dynamic molecular combing and this is the only method to determine the arrangement of individual rDNA units within a repeat cluster[20]. So far, nothing is known about the epigenetic state of non-canonical repeats, since they are impossible to analyze with conventional DNA methylation methods due to the limitation in read length or merging of individual repeats in ensemble methods. Therefore, epi-combing is so far the only DNA methylation analysis technology that can resolve this problem. By combining conventional Dynamic molecular combing with immunodetection, the DNA methylation status of non-canonical repeats was revealed in different cell types.

### Non-canonical rDNA repeats of cancer cells and Imr90 fibroblasts

In cancer cell lines HCT116 and MCF7, 24% and 36% of all repeats analyzed were non-canonical. These repeats were in 65% methylated in the case of HCT116, and all of the repeats were methylated in MCF7. These data cannot be compared with any previous results, since no data exist on the DNA methylation status of non-canonical repeats. If those data are specific for colon and breast cancer can only be confirmed with the acquisition of more data in these cells, since the number of analyzed repeats so far is limited and hence, false positive or negative data have a great impact on such a small sample size.

Imr90 cells displayed a low level of non-canonical repeat units, and this did not change upon senescence (15% and 13%, respectively). However, previous data state a level of about 25% in Imr90 cells [20]. This variation can be either explained by the different sample size used in the experiments (613 repeats in this thesis compared to 6357 repeats) or by the fact that former experiments used a two-colour detection system to identify different kind of repeats. Due to the additional level of DNA modification detection in epi-combing, a bar code system was invented for the discrimination of different repeats. This bar code has the drawback that complex patterns of ribosomal DNA repeats cannot be resolved as well as fusion events of different rDNA parts. These signals that could not be clearly identified were eliminated in the statistical analysis. These false negative repeats reduce the value of non-canonical repeats, since they do not appear in the analysis.

Interestingly, all non-canonical repeats were methylated in not senescent cells (99%), whereas this value decreases to 64% upon senescence. In addition, the overall methylation level is also reduced to 58% compared to 71%. It is tempting to speculate whether the specific demethylation of non-canonical rDNA repeats is a hallmark of ageing. However, the sample size should be amplified to more than 1000 repeats in each sample to observe significant changes.

### **Non-canonical rDNA of primary fibroblasts compared to fibroblasts of Werner syndrome patients**

Epi-combing data on non-canonical rDNA of primary fibroblasts of healthy individuals and patients suffering from Werner syndrome were obtained, too. The data of healthy individuals revealed levels of 21% (AG05283) and 59% (AG13077) which are methylated to 78% and 100%, respectively. In contrast, the levels of non-canonical rDNA in cells of Werner syndrome patients were calculated as 53% (AG12797) with a methylation level of 89% as well as a non-canonical rDNA level of 48% (AG06300) with a corresponding percentage of methylation of 67%.

In former experiments, the total percentage of non-canonical units were determined around 27% in AG05283 and AG13077 [20]. While the values are comparable in the case of AG05283, the level of AG13077 determined in this work is much higher. To reveal the real number of non-canonical repeats, the sample size must be increased with additional experiments. Although the sample size is limited, it can already be stated that the majority of non-canonical repeats seems to be silenced. However, obtaining more data will test this observation.

The non-canonical percentage was about 60% in AG12797 as well as AG06300. The results obtained by epi-combing are between 53% (AG12797) and 48% (AG06300). These absolute numbers may increase to 60% when more repeats are analyzed, however due to the limitation of the bar code system, the value may remain below the formerly observed 60%.

In both patients not all non-canonical repeats are methylated, with the higher one (89%: AG12797) belonging to a 36 year old donor and the lower level (67%:AG06300) coming from a donor of 37 years of age. Interestingly, the lower value is very similar to the value observed in senescent fibroblasts and it is possible that DNA methylation gets lost with the onset of the disease, just as observed in the cell culture system of replicative senescence. However, far more epi-combing data are needed to confirm this speculation as well as the exact clinical features of

the different donors.

In summary, analysis of more repeats from each cell line and thereby amplifying the sample size will corroborate the initial findings on the DNA methylation of ribosomal DNA with this novel method and will confirm the novel insights about the epigenetic regulation of ribosomal DNA.

### **3.3 Large-scale organization of ribosomal DNA chromatin is regulated by Tip5**

#### **3.3.1 Targeting of rDNA to the nuclear matrix**

The chromatin remodelling complex NoRC is a key factor required for repression of the gene by repositioning the promoter bound nucleosome and initiates heterochromatin formation by its interaction with HDACs and Dnmts (reviewed in [85]). In this work, it is shown that NoRC mediates also the level of higher order rDNA chromatin organization by recruiting rDNA to the matrix. The results imply that in addition to its well-defined role in regulating local chromatin structures at the rDNA promoter, NoRC is further involved in large-scale chromatin domain organization of the rDNA locus.

#### **Association of ribosomal DNA to the nuclear matrix**

The association of mammalian rDNA with the nuclear matrix was already shown earlier. Genome-scale biochemical [101] and cell biology experiments [30] demonstrated the specific enrichment of rDNA in the nuclear matrix. However, neither the transcriptional activity of the nuclear matrix-associated rDNA, nor the sequences within the rDNA repeat unit, which mediate the association with the nuclear matrix, were identified. Regarding these questions, seemingly contradictory models were proposed: Keppel suggested that the entire rDNA repeat unit is associated with the nuclear matrix [66], whereas others found that the coding sequence itself [125] or non-transcribed regions flanking the 47S rRNA coding sequence are predominantly enriched in the nuclear or nucleolar matrix [131, 14, 134]. With regard to the transcriptional activity of nuclear matrix associated DNA, it was suggested on one side that active rDNA is associated with the nuclear matrix [66, 131], and on the other side that the nuclear matrix contains transcriptionally inactive rDNA [14], which could also represent sequences that are being replicated [76] (summarised in [166]).

#### **Definition of the nuclear matrix**

These discrepancies can be explained largely by differently used terminology and differences in the experimental procedures. The nuclear matrix [7], nuclear scaffold [92] and nuclear skeleton [61] are operational definitions, which are based on biochemical fractionation approaches. The experimental procedures include different endonuclease digestions followed by high-salt or low-salt extractions, or the fractionation is carried out at physiological salt concentration. Remarkably, the

concentration of DNase I and the incubation time of the endonuclease digestion vary frequently between the protocols of different laboratories, which may affect the observed association of the rDNA with the nuclear matrix as shown in an initial study [101]. The addition of nucleolus isolation steps to the nuclear matrix isolation procedure in particular studies [14, 134] further complicates the comparability of the published data about the nuclear matrix association of rDNA as summarised in [166].

### **Tip5 regulates nucleolar topology**

In this work, the nuclear matrix is isolated by applying extensive DNase I digestion and high-salt extractions as described in former publications [7, 55, 114] and named the last insoluble fraction as nuclear matrix, according to the nomenclature of the initial publication [7]. It is important to note that this nuclear matrix does not represent an identifiable sub-nuclear structure [51, 103]. However, the DNA content of the nuclear matrix represents a fraction of the genome, which is resistant to extensive DNase I digestion, and specific sequences that are enriched in this fraction possess gene regulatory functions [100].

The presented data on Tip5-dependent nuclear matrix targeting of rDNA indicate that besides its other functions, Tip5 also regulates the DNase I accessibility of rDNA in the nucleus and in turn nucleolar topology. Upon overexpression of Tip5 and subsequent quantification of the rDNA content in the nuclear matrix fraction, it is demonstrated that not only the MAR, but also the *bona fide* binding site at the promoter and a 28S rRNA coding region, where no Tip5 binding occurs, were enriched in the nuclear matrix fraction. This suggests that Tip5-mediated silencing of rRNA genes, but not only its direct DNA binding augments the association of rDNA with the nuclear matrix.

### **3.3.2 DNA binding features of Tip5 and its functional consequences**

#### **Binding properties of Tip5's AT-hooks**

Tip5 contains the TAM domain and four minor groove binder AT-hooks, which are supposed to bind MARs and mediate nuclear matrix association [2]. To identify Tip5's protein domain that could mediate association with the nuclear matrix, the DNA binding features of the AT-hooks were investigated by microscale thermophoresis experiments. It was already shown that the TAM domain binds considerably less efficiently to DNA than the AT-hooks [136]. Similar DNA binding affinities were detected for three AT-hooks, while one of them bound less efficiently to the DNA fragments tested. In summary, the comparison of experimentally observed DNA binding activities of the AT-hooks showed the following order:  $AT1 < AT2 \sim AT3 \sim AT4 < HMGA1$  in contrast to the expected  $AT1 \sim AT3 < AT2 < AT4 \sim HMGA1$ , which is based on the classification of [2].

### Nuclear matrix binding activity of AT-hooks and the TAM domain

Quantification of the DNA binding efficiencies also revealed that the combination of the first two AT-hooks bound most efficiently to putative matrix attachment regions. Indeed, the role of the AT-hook domain of another chromatin remodelling protein, Brm/Snf2 $\alpha$ , in nuclear matrix binding was demonstrated in a previous study [17]. Thus, this double AT-hook domain along with its mutant was tested for nuclear matrix binding activity. However, the Tip5 AT-hook domains alone did not yield nuclear matrix binding and therefore, this domain as well as its mutant were extended with the TAM domain and tested again for nuclear matrix binding activity<sup>1</sup>. The results revealed that the TAM domain is a nuclear matrix targeting domain, which is in agreement with its proposed role [2]. In addition, both the TAM domain and the double AT-hook domain of Tip5 were identified as nucleolar targeting sequences.

Finally, the targeting of rDNA to the nuclear matrix by these Tip5 domains was investigated, where no significant changes in the matrix association of rDNA upon overexpression of the different proteins could be detected. This result indicates that additional parts of Tip5 are required for the specific enrichment of rDNA in the nuclear matrix. The overexpression of these domains could result in genome-wide MAR binding, which prevents detectable rDNA-specific targeting effects.

In contrast, overexpression of the full-length Tip5 clearly showed such an effect. In summary, our findings suggest a dual role for Tip5's double AT-hook and TAM domain, targeting the nucleolus and anchoring to the nuclear matrix.

### Functional consequence of AT-hook binding to DNA

Since DNA binding probably defines the substrate specificity of the NoRC complex, the role of the AT-hooks might also be essential in genomic targeting and/or anchoring of Tip5. Until now only a few genomic loci containing repetitive DNA have been described for NoRC activity: targeting is achieved through the sequence-specific DNA binding protein TTF-I at the rDNA promoter [94, 137], and the binding is possibly maintained by the NoRC-associated non-coding pRNA [81]. Further, the association of Tip5 with satellite DNA has been reported [120]. However, there is no current genome-wide binding analysis of NoRC available to demonstrate or exclude the existence of additional genomic binding sites due to the lack of high-quality antibodies against the Tip5 protein.

Notably, AT-hooks can be modified by post-translational modifications. The effects of arginine methylation by PRMT6 and serine phosphorylation by Cdc2 on the DNA binding of HMG proteins have already been demonstrated [126, 124, 170]. Indeed, the acetylation of the second AT-hook of Tip5 has been shown to have an important function in the regulation of rRNA synthesis [165]. Such posttranslational modifications of chromatin modifying complexes could represent an additional level of regulation in the shaping of the epigenetic landscape of the genome.

---

<sup>1</sup>Experiment performed by Michael Filarsky



### Extended AT-hook as a novel DNA binding motif

Interestingly, a non-canonical, extended AT-hook (eAT) of Tip5 was identified and shown that its DNA binding affinity is comparable to the canonical AT-hook AT1. Remarkably, a search for (K/R)>3-N6-20-GRP-N6-20-(K/R)>3 sequence motifs in the human and mouse proteomes revealed 81 and 60 different extended AT-hook peptides, respectively. Reducing these numbers by previously identified canonical AT-hooks, still a substantial proportion of novel eAT hooks were identified which approximately double the number of known AT-hook proteins. The high number of this sequence motif, the conservation between mouse and human and the similar molecular functions of the identified proteins suggests that they indeed possess the function of canonical AT-hook motifs.

However, more experiments are needed to determine whether this is a novel binding motif or whether DNA binding activity is merely caused by electrostatic interaction of the basic amino acid patches flanking the GRP motif. To test this, a peptide that carries a point mutation that changes the GRP core motif to GDP would be useful. If DNA binding is abrogated in microscale thermophoresis experiments, it will indicate that the extended AT-hook is a novel motif.

In addition, the DNA binding properties of further eAT hooks of other proteins should be investigated. Proteins that are involved in nucleic acid metabolism but nevertheless lack any other identified DNA binding domain except eAT-hooks will be an ideal system to test for DNA binding.

### 3.3.3 A biological role for NoRC binding to DNA

The following role for Tip5 in the organization of nucle(ol)ar transcription and architecture is known: Tip5 in the NoRC complex is recruited to the rDNA promoter through TTF-I and then alters chromatin structure (53). Through subsequent interactions with regulatory RNA(s), HDACs and Dnmt proteins, the silenced state of rDNA is established (reviewed in [85]). In addition, the data presented on Tip5-dependent recruitment of the rDNA to the nuclear matrix demonstrates that Tip5 also regulates the localisation of the subpopulation of rRNA genes in the nucleus.

The results suggest that the entire rDNA repeat can be associated with the nuclear matrix upon Tip5 overexpression. Thus, a further model is proposed in which Tip5 plays a key role in recruiting the rDNA to the nuclear matrix and NoRC-mediated heterochromatin formation and chromatin compaction leads to limited DNase I accessibility and the accumulation of large rDNA chromatin domains in the nuclear matrix. Taken together, the results provide insights into the activity-dependent large-scale organization of nucleolar rDNA chromatin and reveal a novel function of Tip5 in this process.

The organization of the nucleolus-associated genome has been analyzed recently [95]; however, little is known about the role of chromatin-dependent regulator proteins in the organization of the higher-ordered arrangement of chromatin domains in this subspace of the nucleus. The importance of the epigenetic regulator protein Dnmt1 has been demonstrated in the maintenance of the nucleolar structure [35].

The role of CTCF in tethering chromatin insulator sites to the perinucleolar chromatin was also revealed [163]. Future studies will verify the possible involvement of Tip5 in the organization of nucleolus-associated chromatin domains, as well as the role of its posttranslational modifications in the higher-ordered chromatin regulation of rRNA genes, which is a novel property of this multifunctional chromatin regulator protein.

## 4 Materials and Methods

### 4.1 Materials

#### 4.1.1 Technical devices

Instrument	Supplier
15 W Blacklight blue lamps (315-400 nm)	Sankyo-Denki
-80°C freezer	Sanyo
37°C incubator	Heraeus Instruments
37°C plate incubator	Memmert
agarose gel chamber	Werkstatt Universität Regensburg
Autoclave	Verioklav
Autoclave LTA 25B2	Zirbus
Axiovert 200M	Zeiss
Centrifuge 5415R	Eppendorf
Centrifuge Centrikon T-324	Kontron Instruments
Centrifuge 3-16 K	Sigma
Centrifuge 1-14 (non-cooled)	Sigma
Centrifuge 4-15 (ZK)	Sigma
Chilling jar	Roth
Coplin staining trough	Roth
Diamond grinding pen	Roth
Dumont #5 Inox medical forceps	Roth
FluorescenceImageReader LAS3000	Fuji
GelMax UV imaging system	Intas
Laminar flow hood	antair BSK
Ice machine	Ziegren
Incubat	Melag
Inkubator 1000, Titramax 1000	Heidolph
Incubator SafeCell UV	Sanyo
Cooled incubator	LMS
Magnetic stirrer MR Hei-Mix L	Heidolph
Magnetic stirrer MR 3001	Heidolph

<b>Instrument</b>	<b>Supplier</b>
Microscope IX50	Olympus
Microwave	Sharp
Microwave	MDA
Millipore machine	ELGA
Monolith NT.115	Nanotemper technologies
Molecular combing device	Genomic vision
NanoDrop UV Spectrometer ND-1000	PeqLab
Overhead shaker	Biosan
PCR machine Primus 96 advanced	PeqLab
pH electrode	Knick
pH meter	Knick
Photometer	Amersham Biosciences
Pipetman pipets	Gilson
Pipethelper "Pipetboy comfort"	Integra Biosciences
Power supply - EPS301	GE Healthcare
Protein gel chambers	Invitrogen
Rotilabo wash bottles	Roth
Rotor-Gene RG3000 system	CORBETT Research
RotorShake genie	Scientific Instruments
SafeImager	Invitrogen
Sonifier 250	Branson
Stratalinker <sup>®</sup> UV Crosslinker 2400	Stratagene
table top centrifuge	Eppendorf
Thermomixer comfort	Eppendorf
Trans-Blot <sup>®</sup> SD Semi-dry transfer cell	BioRad
Unigeldryer 3545D	Biosan
Unimax 1010 Shaker	Heidolph

### 4.1.2 Chemicals and Reagents

Unless otherwise stated, all common chemicals and consumables were purchased from Abcam, GE Healthcare, Genomic Vision, Merck, Invitrogen, Fermentas, New England Biolabs, Promega, Roche, Roth, Santa Cruz Biotechnologies, Serva, Bio-rad, Stratagene/Agilent and Sigma Aldrich.

<b>Material</b>	<b>Supplier</b>
4,5',8 - Trimethylpsoralen	Sigma
5-hm-dCTP	BIOLINE
Acrylamide	Biorad

Material	Supplier
Agarose (ME, LE GP and low melting)	Biozym
Ammoniumpersulfate (APS)	Fluka
Ammonium acetate	Merck
Bacto™ Agar	BD Biosciences
Bacto™ Peptone	BD Biosciences
Bacto™ Tryptone	BD Biosciences
Bacto™ Yeast extract	BD Biosciences
$\beta$ -Mercaptoethanol	Sigma
biotin-XX, SE	Invitrogen
Bradford reagent	BioRad
Bromphenolblue	Serva
Boric acid	Merck
BSA (albumin fraction V)	NEB
Complete Protease Inhibitor Cocktail Tablets	Roche
Coomassie Brilliant Blue G-250	Biorad
DAPI, dilactate	Sigma Aldrich
Digoxigenin-SE	Invitrogen
Dimethylsulfoxide	Merck
dNTPs	Invitrogen
dNTP for nick translation	Boehringer Mannheim
DMEM+GlutaMAX™	GIBCO
EGTA	Sigma
Ethidium bromide	Sigma
EtOH (tech., p.a.)	Merck
Fetal Calf Serum (FCS)	GIBCO
Formamide	Sigma
FuGENE®	Roche
Glycerol 87%	Merck
Glycogen	Roche (901-393)
HEPES	Roth
Isopropanol p.a.	Merck
$\lambda$ phage DNA	Invitrogen
Lithium 3,5-diiodosalicylate	Sigma
N,N-Bisacrylamide	Biorad
NP40	Sigma
Magnesium chloride	Merck
McCoy's 5A (modified) Medium	GIBCO
Methanol p.a.	Merck

Material	Supplier
Milk powder	Sucofin
Orange-G	Sigma
PMSF	Sigma
Potassium chloride	Merck
Protino Glutathione Agarose 4B	Macherey-Nagel
Sodium dodecyl sulfate	Roth
SYBR Safe	Invitrogen
N,N,N',N'-Tetramethylethylenediamine (TEMED)	Roth
Titriplex	Sigma
Tris	Invitrogen
Triton-X 100	Sigma
TE-buffer	Invitrogen
Trypsin-EDTA	Invitrogen
Tween <sup>®</sup> 20	Roth
S-(5'-Adenosyl)-L-methionine chloride	Sigma
Urea	Merck
YOYO <sup>®</sup> -1 iodide	Invitrogen

### 4.1.3 Consumables

Consumables	Supplier
Cell culture dishes	Sarstedt
Cell culture plates	Sarstedt
Cell culture flasks	Sarstedt
Combing surfaces	GenomicVision
Culture tubes (snap lid)	Sarstedt
Glass coverslips	Roth
Fixogum rubber cement	Marabu
Immobilon-P Membrane	Millipore
Instant glue (Cyanoacrilate glue)	UHU
Microscope slides SuperFrost	Roth
Parafilm "M" Laboratory Film	Pechiney
Pasteur pipets glass 230mm	Hirschmann
Petri dishes	Greiner
Pipet tips	Sarstedt
Plastic cuvettes 10x4x45mm	Sarstedt

Consumables	Supplier
Reaction tubes 1.5ml/2ml	Roth
Sterile filter 0.2 $\mu$ m	Millipore
Stuffed pipet tips	Biosphere
Syringes Omnifix	Braun Biotech
Test tube with screw cap 15ml	Greiner
Test tube with screw cap 50ml	Greiner
Whatman paper	Macherey-Nagel

#### 4.1.4 Software and online tools

Software	Source
Axiovision	Zeiss
NEB double digest finder	New England Biolabs (www.neb.com)
Kaleidograph 4.1	Synergy software
LabLife	LabLife software
VectorNTI	Invitrogen (www.invitrogen.com)
Netprimer	Premierbiosoft (www.premierbiosoft.com)
MacVector	MacVector, Inc (www.macvector.com)
Oligoengine	Oligoengine (www.oligoengine.com)
UCSC genome browser	<a href="http://genome.ucsc.edu/ENCODE">http://genome.ucsc.edu/ENCODE</a>

#### 4.1.5 Buffers, Solutions and Media

Stock solutions and buffers were prepared according to standard protocols [89],[115]. Protease inhibitors (either Leupeptin 0.5 $\mu$ g/mL, Pepsatin 1 $\mu$ g/mL, Aprotinin 1 $\mu$ g/mL and PMSF 0.5mM or Complete Protease Inhibitor Cocktail Tablets by Roche) as well as antibiotics were freshly added before use.

Name	Composition
Annealing buffer (10x)	200 mM Tris-HCl pH 7.4, 20 mM MgCl <sub>2</sub> , 500 mM NaCl
Coomassie staining solution	45% (v/v) water, 45% (v/v) methanol, 10% (v/v) acetic acid

Name	Composition
CSK buffer	10mM Pipes pH 6.8, 100mM NaCl, 300mM sucrose, 3mM MgCl <sub>2</sub> , 1mM EGTA, 1mM TCEP, 0.05% (v/v) Triton X-100
CSK I buffer	100mM NaCl, 300 mM Saccharose, 3mM MgCl <sub>2</sub> , 1 mM EGTA pH 8.0, 10mM Pipes pH 6.8, 0.5% Triton X-100, 1mM TCEP, 1x Protease Inhibitor cocktail, 1M NH <sub>4</sub> Ac
CSK II NaCl buffer high salt	100mM NaCl, 300 mM Saccharose, 3mM MgCl <sub>2</sub> , 1 mM EGTA pH 8.0, 10mM Pipes pH 6.8, 0.5% TritonX-100, 1mM TCEP, 1x Protease Inhibitor cocktail, 2 M NaCl
CSK II LIS buffer low salt	3.75 mM Tris HCl, pH 7.5, 0.05 mM spermine, 0.125 mM spermidine, 0.5 mM EDTA, 5 mM MgCl <sub>2</sub> , 20 mM KCl
DNA lysis buffer	50mM Tris pH 8.0, 100mM EDTA, 100mM NaCl, 1% SDS
DNA sample buffer (10x)	50% glycerol, 50mM Tris/HCl pH 7.6, 10mM EDTA, 0.05% Orange G
GST purification lysis buffer	1xPBS, 1mM TCEP, 0,5% Triton X-100, Protease inhibitor Mix
GST purification wash buffer	1xPBS, 1mM TCEP, 0,5% Triton X-100, Protease inhibitor Mix
GST purification elution buffer	20mM Tris, pH 8.0, 20mM Glutathione, Protease inhibitor Mix
MST buffer (5x)	250 mM Hepes pH 7.4, 25 mM MgCl <sub>2</sub> , 500 mM NaCl and 0.25% (v/v) NP40
Nucleolus isolation Buffer A	10mM HEPES-KOH pH 7.9, 1.5mM MgCl <sub>2</sub> 10mM KCl, 0.5mM TCEP
Nucleolus isolation S1	0.25M sucrose, 10mM MgCl <sub>2</sub>
Nucleolus isolation S2	0.35M sucrose, 0.5mM MgCl <sub>2</sub>
Nucleolus isolation S3	0.88M sucrose, 0.5mM MgCl <sub>2</sub>
Laemmli buffer	300mM Tris/HCl pH 6.8, 10% SDS, 50% glycerol, 0.05% Bromphenolblue, 5% Mercaptoethanol
LB medium	1% (w/v) BactoTM Tryptone, 0.5% BactoTM Yeast Extract, 1% (w/v) NaCl, add 2% agar for agar plates
Combing buffer	MES 0.5M, pH 5.5 with NaOH



Name	Composition
PBS	137mM NaCl, 2.7mM KCl, 1mM Na <sub>2</sub> HPO <sub>4</sub> , 2mM KH <sub>2</sub> PO <sub>4</sub> pH 7.4 with HCl
PBS/T buffer	PBS supplemented with 0.1% Tween <sup>®</sup> 20
PBS/T milk buffer	same as PBS/T but with 5% (w/v) milk powder
SDS running buffer	192mM glycine, 25mM Tris, 0.1% (w/v) SDS
SDS PAGE stacking buffer (4x)	0.5 M Tris-HCl, 0.4% SDS, to pH 6.8 with HCl
SDS PAGE separating buffer (4x)	1.5M Tris-HCl, 0.4% SDS, adjust to pH 8.8 with HCl
20xSSC	3.0 M NaCl and 0.3 M sodium citrate, at pH 7.0 with HCl
TE buffer	10mM Tris-HCl pH 7.6, 1mM EDTA
TBE buffer	90mM Tris-HCl, 90mM borate, 2mM EDTA, pH 8.0
Towbin Buffer	0.025 M Tris, 0.192 M Glycine, 20% Methanol
1% Triton-X buffer	500mM Tris-HCl pH 7.4, 150mM NaCl, 5mM EDTA, 1% Triton X-100
8M urea buffer	10mM Tris-HCl pH 8.0, 100mM NaH <sub>2</sub> PO <sub>4</sub> , adjusted to pH 8.0

### Additiva for bacteria media

Reagent	Final concentration ( $\mu\text{g/mL}$ )
BluoGal	100
IPTG (for agar plates)	40
Tetracycline	10
Kanamycin	50
Ampicillin	100

### 4.1.6 Enzymes

Enzymes	Supplier
$\beta$ -agarase	New England Biolabs
DNase I	Roche
DNase I (grade II, from bovine pancreas)	Roche

Enzymes	Supplier
DNA Polymerase I (Kornberg Polymerase)	Fermentas
Herculase polymerase	STRATAGENE
MNase	Roche, Invitrogen
M.SssI	New England Biolabs
Proteinase K	Sigma
Restriction endonucleases	New England Biolabs
RNase A	Roche, Invitrogen
Taq DNA polymerase	Längst lab
HotStarTaq	Qiagen

### 4.1.7 Kits

If not stated otherwise, all kits were used according to the manufacturers' manuals.

Kit	Supplier
BM Chemilumineszenz Blotting Substrate (POD)	Roche
Expand 20kb <sup>PLUS</sup> PCR System, dNTPack	Roche
Expand Long Template PCR System	Roche
iScript Select cDNA Synthesis Kit	Bio-Rad
NucleoSpin RNA II	Macherey Nagel
plasmid purification kit	Qiagen
PCR purification kit	Qiagen
Pure Link HiPure Plasmid Maxiprep Kit	Invitrogen
QIAEXGel Extraction kit	Qiagen
Senescence Detection Kit	BioVision
Super signal WEST Dura WB kit	Pierce

### 4.1.8 Standard DNA and protein weight markers

Weight marker	Supplier
2-log ladder	NEB
100 bp ladder	NEB
Gene ruler 1kb plus DNA ladder	Fermentas
Gene ruler (ultra low range)	Fermentas
Pre-stained protein marker page ruler	Fermentas

### 4.1.9 Antibodies

Stock solutions of antibodies were prepared according to the instructions given by the supplier.

#### 4.1.9.1 Primary antibodies

Antibodies	Supplier	Description	Dilution
M- $\alpha$ -33D3	Diagenode	mouse, monoclonal	EC(1:50) IF(1:25)
R- $\alpha$ -5hmC	Diagenode	rat monoclonal	EC(1:50)
Rb- $\alpha$ -5hmC	Active Motif	rabbit, polyclonal	EC(1:25)
Rb- $\alpha$ -Actin	Sigma	rabbit, polyclonal	WB (1:500)
Rb- $\alpha$ -B23	Santa Cruz sc-56622	rabbit, polyclonal	IF (1:100)
Rb- $\alpha$ -H3	Abcam	rabbit, polyclonal	CC (1:200)
Rb- $\alpha$ -Ki67	Santa Cruz sc-15402	rabbit, polyclonal	IF (1:100)
Rb- $\alpha$ -LaminAC	Santa Cruz sc-20681	rabbit, polyclonal	WB(1: 500)
MBD2-Fc	Rehli laboratory, University of Regensburg	catalytic domain of MBD2 fused to Fc fragment	EC(1:100)
M- $\alpha$ -RPA194	Santa Cruz	mouse, polyclonal	IF(1:50)

Antibodies	Supplier	Description	Dilution
M- $\alpha$ -UBF	Santa sc-13125	Cruz mouse, polyclonal	WB(1:500) IF(1:25)
Rb- $\alpha$ -UBF	Santa sc-9131	Cruz rabbit, polyclonal	WB(1:500) IF(1:25)
Rb- $\alpha$ -mTIP5-N1-18 <sup>1</sup>	Grummt laboratory, DKFZ	rabbit, polyclonal	WB(1:250)
M- $\alpha$ -Tubulin	Abcam	mouse, monoclonal	WB (1:10,000)

**Abbreviations:**  $\alpha$  (anti), G (goat), Rb (rabbit), M (mouse), WB (Western Blot), EC (Epi-combing), IF (Immunofluorescence), CC (Chromatin-combing).

#### 4.1.9.2 Secondary and tertiary antibodies

Commercial secondary antibodies (anti-rat, anti-rabbit both horseradish peroxidase conjugated) were used at a 1:3000 dilution in Western Blot experiments.

For secondary antibodies used in Immunofluorescence experiments and dilutions, see below:

Antibodies	Supplier	Description	Dilution
G- $\alpha$ -M-Cy3	Jackson	Goat polyclonal	Combing (1:50) IF (1:25)
G- $\alpha$ -M-Cy5	Jackson	Goat polyclonal	IF (1:25)
G- $\alpha$ -Rb-Cy3	Jackson	Goat polyclonal	IF (1:25)
D- $\alpha$ -H-Cy5	kind gift from AG Cremer	Goat	IF (1:25)
D- $\alpha$ -G-Cy3	Rockland	Donkey polyclonal	Combing (1:50)
D- $\alpha$ -Rb-DL488	Jackson	Donkey polyclonal	IF (1:50)
Streptavidin-Alexa488	Invitrogen		Combing (1:50) IF (1:50)

<sup>1</sup>antibody described in [136]

Antibodies	Supplier	Description	Dilution
Rb- $\alpha$ -streptavidin-biotin	Rockland	rabbit	WB (1:500) IF (1:25)
G- $\alpha$ -M-Alexa350	Invitrogen	Goat polyclonal	Combing (1:25)
M- $\alpha$ -G-AMCA	Jackson	Mouse polyclonal	Combing (1:25)
S- $\alpha$ -dig	Roche	Sheep Fab fragments	Combing (1:50)
D- $\alpha$ -S-DL549	Jackson	Donkey	Combing (1:50)
S- $\alpha$ -dig-rhodamine	Roche	Sheep Fab fragments	Combing (1:50)

**Abbreviations:**  $\alpha$  (anti), D (Donkey), G (goat), H (Human), Rb (rabbit), M (mouse), EC (Epi-Combing), IF (Immunofluorescence)

#### 4.1.10 Bacterial strains and mammalian cell types

##### Bacterial strains

<b>XL1 Blue</b>	
general	F-episome, general DNA plasmid propagation, blue/white screening
resistance	tetracycline
genotype	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F'proAB lacIqZDM15 Tn10 (Tetr)]
<b>BL21</b>	
general	Expression Strain: deficient in lon and ompT proteases
resistance	chloramphenicol, tetracycline
genotype	F-ompT gal hsdSB (rB-mB-) dcm lon lambdaDE3

##### Mammalian cell types

<b>Hek293</b>	
cell type	human embryonic kidney

maintenance	37°C, 5% CO <sub>2</sub> , DMEM-GlutaMAX, 10% FCS
supplier	ATCC <sup>2</sup>
<b>HeLa</b>	
cell type	human cervix carcinoma
maintenance	37°C, 5% CO <sub>2</sub> , DMEM-GlutaMAX, 10% FCS
supplier	ATCC <sup>3</sup>
<b>HCT116</b>	
cell type	human colon carcinoma
maintenance	37°C, 5% CO <sub>2</sub> , McCoy's 5A (Modified), 10% FCS
supplier	ATCC <sup>4</sup>
<b>Imr90</b>	
cell type	human fetal lung fibroblast
maintenance	37°C, 5% CO <sub>2</sub> , DMEM-GlutaMAX, 10% FCS
supplier	ATCC <sup>5</sup>
<b>U2OS</b>	
cell type	human osteosarcoma <sup>6</sup>
maintenance	37°C, 5% CO <sub>2</sub> , DMEM-GlutaMAX, 10% FCS
supplier	ATCC <sup>7</sup>
<b>MCF-7</b>	
cell type	human breast carcinoma
maintenance	37°C, 5% CO <sub>2</sub> , DMEM-GlutaMAX, 10% FCS
supplier	ATCC <sup>8</sup>

#### 4.1.11 Plasmids

Plasmid	Source
HrDNA 1/7	Attila Németh
HrDNA 3	Attila Németh

<sup>2</sup>Hek293 cells were obtained from ATCC second hand at low passage number

<sup>3</sup>HeLa cells were obtained from ATCC second hand at low passage number

<sup>4</sup>HCT116 cells were obtained second hand at low passage number

<sup>5</sup>Imr90 cells were obtained from ATCC second hand at low passage number

<sup>6</sup>described in [64]

<sup>7</sup>U2OS cells were obtained from ATCC second hand at low passage number

<sup>8</sup>MCF-7 cells were obtained from ATCC second hand at low passage number

Plasmid	Source
HrDNA 4	Attila Németh
HrDNA 6	Attila Németh

#### 4.1.12 Oligonucleotides

Name	Sequence
HsRRN3g_24210F	5'- TTT GAG CGG AAA CCT GAA AG -3'
HsRRN3g_24407R	5'- CTT AAG GGG TTT GTA GGG GC -3'
HsRRN3_1236F	5'- TTG CAG GAC CCA AGT AAT CC -3'
HsRRN3_1435R	5'- AAT GGT CCA TGG AGA GCA AC -3'
IFNb102F	5'- GGA GAA GTG AAA GTG GGA AAT TCC TCT GAA TAG A -3'
IFNb318R	5'-T CTA TTC AGA GGA ATT TCC CAC TTT CAC TTC TCC -3'
PrdII - F	5'- Cy3- GGG AAA TTC CTC-3'
PrdII R	5'-GAG GAA TTT CCC -3'
CBG25624F	5'- GGT CTT CCT CTG AAA AAC CCA TAA CC-3'
CBG25784R	5'-CCT TAG GCT TCT CTT GGC TGT GAC -3'
pIRES-SeqFor_KZ	5'- TAG CGC TAC CGG ACT CAG AT -3'
pIRES-SeqRev_KZ	5'- GAG GAA CTG CTT CCT TCA CG -3'
Hr36189F	5'- TCG CCG ACT CTC TCT TGA CTT G-3'
Hr36399R	5'- TGG AGC ACA GTG ACA CAA CTA TGG-3'
Hr9661F	5'- CGA ATG ATT AGA GGT CTT GGG GC -3'
Hr9860R	5'- TGG GGT CTG ATG AGC GTC GG -3'
Hr42857F	5'- ATG GTG GCG TTT TTG GGG AC -3'
Hr42964R	5'- CGA AAG ATA TAC CTC CCC CG -3'
en3-MST	5'-Cy3- GGG AAA TTC CTC -3'
en3-MST- complementary	5'- GAG GAA TTT CCC -3'
L35496F	5'- CAA AGC CTT CTG CTT TGA ATG -3'
L48011R	5'- ACA GTG ACA GAC TGC GTG TTG -3'

## 4.2 Methods

Preparation and transformation of chemically competent bacteria with DNA, amplification of plasmid DNA in *E. coli* bacteria, purification, concentration determination, restriction enzyme digestion, ligation of DNA fragments, analysis of DNA on agarose and polyacrylamide gels, as well as amplification of the DNA by the polymerase chain reaction (PCR) was performed according to the standard protocols [115], [89]. Bacteria were cultured in Luria Bertani (LB) medium and selective antibiotics were added corresponding to the plasmid encoded resistance.

### 4.2.1 Working with nucleic acids

#### 4.2.1.1 Isolation of nucleic acids

##### Isolation of plasmid DNA from bacteria

Small scale DNA preparations were performed with 2 mL of an overnight bacteria culture using the Qiagen kit together with the QiaCube to isolate plasmid DNA. For large scale DNA preparations, 100 mL of an overnight culture was used and the corresponding Invitrogen Midi prep kit was used according to the manufacturer's instructions.

##### Isolation of genomic DNA from human cells

After aspirating media from the cells, pre-warmed DNA lysis buffer (37°C) was added to the cells and incubated for about 30 sec (300 µL per 10<sup>6</sup> cells). The viscous suspension was transferred into a sterile cup and 10 µL of a 10 mg/mL RNaseA solution was added per millilitre lysate. After incubation for a minimum of 2 hours at 37°C, 25 µL of a 10 mg/mL proteinaseK solution per millilitre lysate was added and incubated over night at 50°C. The next day, the DNA was purified with ammonium acetate/ethanol precipitation.

##### Isolation of RNA from human cells and reverse transcription

RNA was isolated using NucleoSpin RNA II by Macherey Nagel according to the manufacturer's instructions except the DNase I incubation step was prolonged to 40 min. cDNA was synthesized using iScript Select cDNA Synthesis Kit by Bio-Rad with random primers to transcribe also ribosomal RNA.

#### 4.2.1.2 DNA purification

DNA precipitation is carried out to remove impurities such as proteins, salts and sugars. This can be done by two methods, ethanol ammonium acetate precipitation or by gel extraction. The first one serves the purpose of purification of DNA from impurities whereas the gel extraction is used when a DNA fragment needs to be specifically separated from other DNA as well as impurities.



### **DNA purification by precipitation**

All following steps can be carried out at room temperature. 0.5 Volume of 7.5 M ammonium acetate solution and 2 Volume of absolute ethanol were added to the DNA solution and the mixture was then spun down in a tabletop centrifuge for 20 min at 16,100 rcf. The supernatant was carefully removed and the pellet washed with 1 mL 70% ethanol and centrifuged for 15 min. The pellet was air-dried for about 10 min and resuspended in the required volume of TE buffer.

### **DNA purification by gel extraction**

The DNA was mixed with 10xDNA loading dye and loaded on an agarose gel. Electrophoresis were run at 100 V until the fragments were sufficiently separated. An agarose block containing the desired DNA fragment was then sliced out of the gel and the DNA was extracted using the QIAquick Gel Extraction kit according to the manufacturer's protocol.

#### **4.2.1.3 Checking the quality of nucleic acids**

##### **Assessing the concentration of nucleic acids**

The DNA and RNA concentration is determined via absorbance measurement at 260 nm using NanoDrop ND1000. The purity of the DNA and RNA sample is checked by comparison of the absorbance at 260 nm relative to 280 nm which arises from absorbance of aromatic amino acids (280 nm). A ratio between 1.8 and 2.0 indicates absence of proteins and therefore, high-quality DNA, a ratio of about 2.0 is generally accepted as "pure" for RNA.

##### **Agarose gel electrophoresis**

DNA was analysed by electrophoresis on 1% agarose gels unless otherwise stated. 0.5 g agarose was mixed with 50 mL 1xTBE buffer and heated until the agarose was completely dissolved. Before the mixture was poured in a gel chamber, SYBR Safe was added in a 1:10,000 dilution as a staining reagent. The gel solidified and 1xTBE buffer was used as running buffer. Then, samples were mixed with 10xDNA loading dye and pipetted into gel slots. The samples were subjected to electrophoresis at 100 V for the appropriate time to separate different DNA fragment lengths. To visualise DNA, the Invitrogen SafeImager system was used.

#### **4.2.1.4 Polymerase chain reaction**

Polymerase chain reaction (PCR) served three purposes during the PhD work. First, it was used to amplify the desired sequence to be cloned into a vector. Second, analytical colony PCR identified positive bacterial clones carrying the insert when screening for transformants. Third, quantitative PCR was used to assess the amplified DNA concentration in real time.

### Primer design

Primers were designed using Netprimer or MacVector (Primer3). The complementary part of the primers was designed to have a melting temperature of  $60^{\circ}\text{C}$  if possible. Furthermore, they were designed to show a minimum of secondary structure and primer dimerisation and to specifically bind to the template. Unless otherwise stated, unmodified primers were obtained from Eurofins MWG Operon whereas fluorescently labelled oligonucleotides were purchased from Metabion.

### Preparative PCR

In general, Taq polymerase was used to amplify the sequence that will be cloned into a vector. Only when the insert was longer than 5 kb, the enzyme Herculanase polymerase was used instead of Taq polymerase.

Pipetting scheme (1 reaction):

5 $\mu\text{l}$	10xPCR buffer
1 $\mu\text{l}$	10mM dNTPs
2.5 $\mu\text{l}$	forward primer (10 $\mu\text{M}$ )
2.5 $\mu\text{l}$	reverse primer (10 $\mu\text{M}$ )
20 ng	DNA template
1 $\mu\text{l}$	Taq Polymerase
ad 50 $\mu\text{l}$	water

The PCR protocol was the following:

Initial denaturation	$95^{\circ}\text{C}$	300sec	1x
Denaturation	$95^{\circ}\text{C}$	30sec	
Annealing	primer specific	30sec	35x
Extension	$72^{\circ}\text{C}$	60sec	
Final extension	$72^{\circ}\text{C}$	420sec	1x
Storage	$4^{\circ}\text{C}$		

Note that the annealing temperature was primer specific but ranged between  $55^{\circ}\text{C}$  and  $58^{\circ}\text{C}$  and that all steps were exerted on ice.

A special PCR set up was used to prepare hydroxymethylated  $\lambda$  DNA. To amplify a 12.5 kb long DNA fragment, either the Expand 20kb <sup>PLUS</sup> PCR System, dNTPack or the Expand Long Template PCR System (Roche) was used, both with same efficiency. To obtain hydroxymethylated DNA, a PCR reaction on  $\lambda$  DNA was performed using primer pairs L35496F and L48011R (hybridization temperature:  $52^{\circ}\text{C}$ ) as well as hydroxymethylated dCTP and an elongation time of 10 min that was increased for 20 sec each cycle at  $68^{\circ}\text{C}$ , a 12.5 kb long, hydroxymethylated

template is synthesized after 20 cycles which can be used in subsequent epi-combing experiments.

### Colony PCR

A pipet tip was dipped into a single bacteria colony that was to be tested for the correct insert. The cells were resuspended in 25  $\mu\text{L}$  water in a 0.2 mL PCR tube and restreaked on a new agar plate containing the respective antibiotics. For cell lysis, the sample was boiled at  $100^\circ\text{C}$  for 10 min in a PCR cycler. Then, 25  $\mu\text{L}$  colony PCR master mix was added and the Thermocycler program was started. After the reaction, the sample was analysed on an agarose gel for the presence of the amplicon.

Pipetting scheme for colony PCR master mix (1 reaction):

5.0 $\mu\text{l}$	10xPCR buffer
1.0 $\mu\text{l}$	Taq polymerase
0.5 $\mu\text{l}$	forward primer (10 $\mu\text{M}$ )
0.5 $\mu\text{l}$	reverse primer (10 $\mu\text{M}$ )
1.0 $\mu\text{l}$	dNTPs (25 $\mu\text{M}$ )
17.0 $\mu\text{l}$	water

PCR program:

Initial denaturation	$95^\circ\text{C}$	300sec	1x
Denaturation	$95^\circ\text{C}$	30sec	
Annealing	primer specific	30sec	35x
Extension	$72^\circ\text{C}$	60sec	
Final extension	$72^\circ\text{C}$	420sec	1x
Storage	$4^\circ\text{C}$		

The annealing temperature was primer specific but ranged between  $55^\circ\text{C}$  and  $58^\circ\text{C}$ .

### Real-time quantitative PCR

Quantitative PCR was used to measure the amount of a specific DNA fragment with high accuracy. Each reaction was performed as a technical triplicate. The quantity of DNA was detected after each PCR cycle by measuring the fluorescence of SYBR-Green (Roche). This is a dye that fluoresces when intercalated into double-stranded DNA. Thus, the intensity of the signal is proportional to the amount of double-stranded DNA present in a sample.

During a PCR reaction, the original amount of DNA  $Z(0)$  is amplified exponentially. This occurs until the polymerase fails or one of the reagents is consumed. The total

amount of DNA  $Z(n)$  can be calculated as follows:

$$Z(n) = Z(0) \cdot E^n \quad (4.1)$$

Where  $Z(0)$  is the original amount of DNA,  $n$  the number of cycles and  $E$  the efficiency of the reaction ( $1 < E < 2$ ). The amplification efficiency can be derived from the slope in a semi-logarithmic plot of fluorescence against cycle number when a standard curve for a primer pair is measured. When  $E$  is known for a primer pair, the relative amounts of specific fragments can be determined. The measured fluorescence is plotted against the number of cycles on a logarithmic scale. A threshold for detection of DNA-based fluorescence is set above background and the number of cycles at which the fluorescence exceeds the threshold is called the threshold cycle ( $C_t$ ). This  $C_t$  value represents the intersection between an amplification curve and a specific threshold line and is used to compare the concentrations of different DNA fragments in the same preparations.

qPCR reactions were performed in 0.2 mL PCR tubes with a reaction volume of 20  $\mu$ L. The reaction contained 4  $\mu$ L of DNA sample (concentration 10 ng/ $\mu$ L or less) and 16  $\mu$ L master mix. This master mix contained 4 pmol of forward and reverse primer each, 0.25  $\mu$ L of a 1:400,000 SYBR-Green stock solution, 0.4 U HotStarTaq and premix. This premix consists of  $MgCl_2$  (final concentration: 2,5mM), (final concentration 0.2mM), 10xPCR buffer (final concentration 1x). qPCR was performed using the Rotor-Gene RG3000 system from CORBETT Research, SYBR-Green was excited at 480 nm, fluorescence was recorded at 510 nm.

The qPCR cyclyer program was:

Initial denaturation	95°C	15min	1x
Denaturation	95°C	10 sec	
Annealing	55°C	20 sec	35x
Extension	72°C	40sec	

Standard curves were measured in serial 1:4 dilutions of a 10 ng (DNA)/ $\mu$ L stock solution till a dilution of 1:1,024. After evaluating the standard curves, the efficiencies of every standard curve were higher than 96%.

#### 4.2.1.5 Restriction digest of DNA

Restriction endonucleases specifically cleave double-stranded DNA at recognition sites. They create either blunt ends or single stranded overhangs which can be used for subsequent ligation reactions (preparative digest) or for the analysis of the right insert during cloning (analytical digest). Additionally, analytical digestions can be applied while testing plasmids for the desired insert. Restriction enzymes were used according to the manufacturer's recommendations concerning buffer, addition of BSA and temperature ([www.neb.com](http://www.neb.com)). These restriction enzymes are stored in a buffer containing 50% (v/v) glycerol which might lead to non-specific enzyme activity during the digestion or to inhibition of digestion. Therefore, all digestions were carried out at a maximum glycerol concentration of 5% (v/v).

For the analytical digest 0.1  $\mu\text{g}$  to 1  $\mu\text{g}$  DNA was incubated with 5 Units of the respective restriction endonuclease in a total volume of 20  $\mu\text{L}$ . The preparative restriction digest was done with 15  $\mu\text{g}$  DNA using 60 Units restriction endonuclease in a total volume of 60  $\mu\text{L}$ . To check the completion of the digest, DNA was electrophoretically separated using TBE agarose gels supplemented with SYBR Safe (Invitrogen).

#### 4.2.1.6 Ligation of DNA fragments

Unless otherwise stated, a molar ratio of 1:5 between vector and insert is used. The amount of insert DNA was calculated as follows:

$$m(insert) = \frac{5 \cdot m(vector) \cdot bp(insert)}{bp(vector)} \quad (4.2)$$

The pipetting scheme for a ligation reaction was:

---

1ul	10mM ATP
1ul	10xligase buffer
1ul	T4 DNA ligase
100ng	100 ng vector
5xmolar	excess insert
ad 10ul	ad 10 $\mu\text{l}$ water

---

#### 4.2.1.7 DNA Sequencing

For sequencing, 300 ng of purified plasmid was mixed with 1  $\mu\text{L}$  of 10  $\mu\text{M}$  sequencing primer and filled to a final volume of 8  $\mu\text{L}$  with water. Sequencing was performed by GeneArt, Regensburg.

### 4.2.2 Protein Methods

#### 4.2.2.1 Denaturing polyacrylamide gel electrophoresis

Proteins were separated according to their molecular weight by discontinuous SDS-PAGE.

#### Preparation of polyacrylamide gels

Invitrogen MiniCell cassettes were used for SDS gels of different percentages. The solutions were prepared according to a standard pipetting scheme (0.2 mL/1% gel) for one gel (Sep: Separating gel, Stack: Stacking gel).

Gel type	Sep x%	Stack 5%
water	ad 6 mL	0.87 mL
30% A-BA <sup>9</sup> mix	x (0.2 mL/1%)	0.25 mL
SDS PAGE separating buffer 4x	1.50 mL	/
SDS PAGE stacking buffer 4x	/	0.38 mL
20% APS	0.03 mL	0.0075 mL
TEMED	0.006 mL	0.0015 mL

Note that APS and TEMED were finally added to start polymerisation. The solution for the separating gel was subsequently poured into the gel chamber and overlaid with isopropanol. After polymerisation, the alcohol was decanted, the stacking solution prepared and immediately cast onto the separating gel. Then, combs were inserted and the gel polymerised. Finally, gels were either immediately used or wrapped in wet tissue paper and plastic foil for storage at 4°C up to 4 weeks.

### Sample preparation and electrophoresis

Protein samples were mixed with the appropriate amount of 5x Laemmli buffer and boiled for 10 min at 95°C. The samples were spun down and pipetted into the slots of the gel. Electrophoresis was carried out at constant current of 35 mA for the required time at room temperature. The gel chamber was then disassembled and gels were either stained or Western Blot analysis was performed.

### Coomassie Staining of gels

Gels were fixed and stained by incubation in Coomassie staining solution at room temperature for 10 min on a belly shaker. The staining solution was poured away and water was added for destaining. The gels were boiled in fresh water until the protein bands got visible.

#### 4.2.2.2 Protein Quantification

The Bradford Assay is used for protein quantification of cell lysates. The Biorad Protein Assay reagent was diluted 1:5 with water and 1 µL of the cell lysate was added to 1000 µL dye reagent in cuvettes. After 5 min incubation, the absorbance was measured at 595 nm ( $A_{595}$ ) and the protein concentration was calculated as follows:

$$conc(mg/mL) = A_{595} \cdot 23 \quad (4.3)$$

<sup>9</sup>30% Rotiphorese mix acrylamide and N-N'-methylenebisacrylamide

#### 4.2.2.3 Semi-dry Western Blot and protein detection

In Western Blots, proteins are transferred by electrophoresis, capillary action or diffusion to a carrier membrane where they get immobilised and detected. As part of this project, the transfer is carried out by electroblotting and the proteins are visualised by immunodetection.

##### Western Blot Semidry method

Negatively charged SDS-protein complexes are transferred from an SDS-gel onto a PVDF membrane by blotting. An SDS-PAGE was done as previously described and after this, the gel was cut to a size of 8x5 cm. Six blotting papers and the PVDF membrane were also cut to the same size. The blotting papers were soaked in Towbin Buffer for 5 min whereas the membrane was rinsed in methanol:

The blot was assembled in a BioRad semidry blotting apparatus in the following order from anode to cathode: three blotting papers, PVDF membrane, SDS-gel, three blotting papers. The blot was assembled without remaining air-bubbles between the different layers. Electrophoresis was carried out at room temperature at 100 mA ( $2.5 \text{ mA/cm}^2$ ) at constant voltage for 60 min.

##### Detection of membrane-bound proteins by Ponceau S staining

Whether the transfer was successful was confirmed by Ponceau S staining. This method labels proteins reversibly and does not interfere with subsequent immunodetection. For staining, the membrane was incubated with Ponceau S solution for 5 min at room temperature on a belly shaker. When protein bands were visible, the membrane was destained by washing several times with water.

##### Immunodetecting the protein of interest

The immobilised proteins were visualised by immunodetection. First, a primary antibody is directed against the protein of interest. Then, a secondary, enzyme-coupled antibody recognises the Fc part of the primary antibody. The coupled enzyme, a horse-radish peroxidase makes the protein visible by catalysing a light-reaction which is detected.

The membrane was blocked in 1xPBS/T containing 5% milk powder for one hour at room temperature. Then, the primary antibody diluted in 1xPBS/T, 5% dry milk was incubated on the membrane overnight at  $4^\circ\text{C}$ . The membrane was washed with 1xPBS/T three times for 10 min with mild shaking at room temperature. Afterwards, the membrane was incubated with the secondary antibody diluted in 1xPBS/T, 5% dry milk for 45 min and washed again three times for 10 min at room temperature.

After washing, the membrane was developed using a Chemiluminescence kit. For low expressed genes, the sensitive Super Signal WEST Dura WB kit (Pierce), for other proteins the Roche Chemiluminescence Substrate (POD) was used. Both kits were applied at room temperature according to the instructions given. The

chemiluminescence was detected at different exposure times with the Fuji LAS3000 FluorescenceImageReader.

#### 4.2.2.4 Purification of GST-tagged proteins

##### **IPTG induction of cells**

2 mL of LB medium/Ampicillin was inoculated with the corresponding clone and incubated over night at 37°C at mild shaking. The next day, 0.5 mL were taken from the culture and added to 50 mL LB medium/Ampicillin in a 250 ml flask. After 2-3 hours, the OD<sub>600</sub> value is checked whether it is in the range of 0.5 (exponentially growing cells). Then, the cells are induced with IPTG and the OD<sub>600</sub> value is checked several times. Successful protein expression is indicated by an inhibition of cell division, resulting in a flattened growth curve. After typically 5-6 hours, cells are harvested by centrifuging at 1,000 rcf and 4°C.

##### **Cell lysis**

After removing the supernatant, cells are resuspended in 1 mL GST purification lysis buffer. To further lyse the cells, they are frozen in liquid nitrogen twice and slowly thawed at room temperature. Afterwards, cells are sonified in an ice-bath for 10 times 10 sec each with 30 sec break between (settings: 50% output, hold). Cell lysate is centrifuged for 1 hour at 16,000 rcf and 4°C, then the supernatant is added to the GST-agarose beads.

##### **Protein binding to the GST-agarose beads**

150 µL of beads slurry is removed for each purification and washed three times with 1 mL 1xPBS each. Afterwards, the supernatant is added to 100 µL beads and incubated for 1 hour at room temperature with an overhead shaker.

##### **Washing the beads and subsequent protein elution**

Beads are washed 4 times with 1 mL GST purification washing buffer each for 5 min at room temperature using an overhead shaker. Afterwards, the protein is eluted by adding 150 µL GST purification elution buffer and incubation for 1 h at room temperature with an overhead shaker. This step is repeated for completely extracting the protein from the beads. 15 µL glycerol is added to the protein fractions to store them at -20°C.



### 4.2.3 Mammalian cell culture

#### 4.2.3.1 Maintenance

##### Maintenance

All work with mammalian tissue culture was done according to the standard methods with standard precautions. All work was performed under a sterile hood with laminar flow, solutions were purchased sterile or sterilised by autoclaving. Working space, gloves and all devices were cleaned with 70% ethanol before use.

Mammalian cells were propagated in their recommended medium. Media were stored at 4°C but pre-warmed to 37°C right before use. Cells were split at an estimated confluency of 80% and the medium was replaced by fresh one every three days. For splitting, medium was aspirated and cells were carefully washed with 1xPBS. The required amount of Trypsin/EDTA solution was added to the cells and cells were incubated at 37°C until they detached (3-4 min). Then, fresh medium was added and cells were split depending on their growth properties and the diameter of the plate. Cells were incubated in a humidified incubator reserved for mammalian cell culture at 37°C and 5% CO<sub>2</sub>.

##### Harvesting

If not harvested using Trypsin/EDTA, cells were harvested by scraping off with a cell scraper. After removing medium, cells were washed with PBS and scraped off using a cell scraper. The cell suspension was centrifuged for 5 min at 500 rcf and then washed with PBS once. After centrifuging, the supernatant was removed and the cell pellet was immediately used or frozen in liquid nitrogen which can then be stored at -80°C.

##### Cryo-conservation

Mammalian cell cultures can be cryopreserved and stored for long time at -80°C or in liquid nitrogen. For cryopreservation, the cells were harvested by Trypsin/EDTA and washed with 1xPBS several three times. The cell pellet was carefully resuspended in a volume of FCS containing 5% DMSO to reach a final cell density of  $5 \cdot 10^6$  cells/mL. The suspension was quickly split in aliquots of 1 mL in sterile cryotubes which were pre-cooled to -20°C. The tubes stored for 24 hours at -20°C before replacing the tubes for long time storage in liquid nitrogen or the -80°C freezer.

When cryo-cultures were needed, they were removed from the -80°C freezer, thawed and poured into a 10 cm plate containing 13 mL of the required medium. The cells were incubated at 37°C and the medium was changed the next day to remove remaining DMSO.

#### 4.2.3.2 Transfection of mammalian cells

Cells were transfected with FuGENE<sup>®</sup> HD Transfection Reagent by Roche according to the manufacturer's instructions.

#### 4.2.3.3 Detection of senescent cells

A known characteristic of senescent cells is the SA- $\beta$ -Galactosidase ( $\beta$ -Gal) activity, which can be histochemically detected. The enzyme  $\beta$ -Gal is present at high concentration only in senescent cells and is not found in presenescent, quiescent or immortal cells. Therefore, a senescence detection kit of Biovision was used according to the manufacturer's constructions.

In addition, a lack of Ki-67, a marker for proliferating cells, indicates senescent cell status when monitored by immunofluorescence (see 4.2.5.2).

#### 4.2.4 Nucleolus isolation

Nucleoli were isolated according to a protocol modified from [19]. To control the purification, different quality controls sheet were used (see figure 4.1) to check the overall quality of the nucleolus preparation. For one isolation, about 100 million cells (10x14cm plates) are used.

##### Nucleolus Isolation without cross-linking

Medium is removed and cells are washed with 1xPBS. Then, cells are trypsinized off the dishes with 3 mL Trypsin-EDTA and 4 min incubation.

Cells are spun down for 4 min at 400 rcf and 4°C and washed three times with 1xPBS.

After final wash, cells are resuspended in 5 mL Buffer A (10 mM Hepes-KOH pH 7.9, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.5 mM TCEP) which causes the cells to swell. Cells are then broken open with a pre-chilled dounce homogenizer with a tight pestle (10 strokes).

Dounced cells are centrifuged at 240 rcf for 5min at 4°C to pellet nuclei and other fragments. This nuclear pellet was resuspended in 3 mL 0.25 mM sucrose, 10 mM MgCl<sub>2</sub> and layered over 3 mL 0.35 mM sucrose, 0.5 mM MgCl<sub>2</sub> followed by centrifugation for 5 min at 4°C and 1400 rcf. This results in a cleaner nuclear pellet.

This pellet is resuspended in 3 mL 0.35 mM sucrose, 0.5 mM MgCl<sub>2</sub> and sonicated in an ice-bath for 6-8x10 sec with 30 sec cooling each using a microtip sonicator. (The release of nucleoli is checked under a light microscope and the sonication step is adjusted according to the result).

The sonicated sample is layered over 3 mL 0.88 mM sucrose, 0.5 mM MgCl<sub>2</sub> and centrifuged for 10 min at 2,800 rcf and 4°C. The pellet contains the nucleoli whereas the supernatant can be retained as a nucleoplasmic fraction.

<b>Nucleolus preparation QC sheet    Sample Nr:</b>		
<b>species*</b>		
<b>cell line*</b>		
<b>protocol*</b>	<b>w/o HCHO</b>	
<b>microscopy*</b>	<b>pict of cells</b>	
	<b>pict of nucleoli*</b>	
	<b>% nucleoli/total particles</b>	
<b>proteins (WB)*</b>	<b>CP - alpha-tubulin*</b>	
	<b>N/No - H2B*</b>	
	<b>N/No - H3*</b>	
	<b>No - UBF*</b>	
	<b>No - RPA194*</b>	
<b>DNA*</b>	<b>concentration (ng/μl)*</b>	
	<b>OD<sub>260/280</sub>*</b>	
<b>qPCR No/WCE</b>	<b>1% agarose- SYBR Safe</b>	
	<b>5'ETS*</b>	
	<b>28S*</b>	
	<b>IFNb Prom*</b>	
<b>comments</b>		
<b>*obligatory controls (with attached electronic documentation)</b>		

**Figure 4.1: Nucleolus isolation quality control sheet**

The scheme illustrates the parameters that are essential for documentation of nucleolus isolations and its quality.

Nucleoli were washed by resuspension in 0.5 mL 0.35 mM sucrose, 0.5 mM MgCl<sub>2</sub> and centrifuged at 2,000 rcf and 4°C.

The nucleoli are either further processed or resuspended in 0.5 mL 0.35 mM sucrose, 0.5 mM MgCl<sub>2</sub> and frozen in liquid nitrogen to store them at -80°C.

## 4.2.5 Microscopy methods

### 4.2.5.1 Fixation of adherently growing cells

Protocol for the fixation of cultured cells was adapted from [133] and performed as follows:

Cells are grown on a sterile coverslip until the desired confluency of cells is achieved.

Medium is aspirated and the coverslip is rinsed 2-3 times with pre-warmed 1xPBS at 37°C.

Cells are fixed in 2% paraformaldehyde in PBS (freshly made, pH 7.0) for 10 min at room temperature. During the last minute a few drops of 0.5% Triton X-100 in 1x PBS are added where fixation is performed.

Coverslip is washed in 1x PBS with 0.01% Triton X-100 for three times 3 min at room temperature.

Cells are permeabilised in 0.5% Triton X-100 for 5 min at room temperature.

The coverslip is incubated in 20% glycerol in 1x PBS for at least 60 min at room temperature (can be left for overnight).

Cells are frozen in liquid nitrogen (15-30 sec), thawed (gradually, at room temperature) and soaked in 20% glycerol. These steps are repeated 4-5 times.

Cells are washed in PBS with 0.01% Triton X-100 for three times 10 min each.

The sample is incubated for 5 min in 0.1M HCl and rinsed in 2xSSC.

Note:

Some cell types require longer incubation (10 min) with 0.1M HCl, some need (after 5 min of HCl) an additional incubation in pepsin. Digestion with pepsin is difficult due to possible over-digestion which results in losing cells or disturbing nuclear morphology.

Pepsin digestion:

Incubation in 0.002% pepsin in 0.01N HCl at 37°C ( time varies depending on cell type, e.g. HeLa 5min).

Cells are washed twice in 1xPBS for 5min each.

Cells are postfixed in 1% PFA / 1xPBS for 10 min and rinsed in 1xPBS three times for 5min each.

The last step is incubation in 50% Formamide/2xSSC at least 30 min at room temperature, or up to several weeks stored at 4°C).

Fixed cells are kept in 50% FA/2xSSC for weeks and up to 3 months (Longer storage leads to bad nuclear morphology).

#### **4.2.5.2 Immunofluorescence**

After cells are fixed, the cells can subsequently be subjected to immunofluorescence.

The coverslip is washed three times for 3 min each in 1xPBS/0.1% Tween 20® at mild shaking and room temperature.

The primary antibody is diluted in BlockAid™ Blocking Solution by Life Technologies and added to the cells.

The sample is covered with a coverslip at the same size and incubated in a humidified chamber at 37°C for 45 min.

After incubation, the sample is washed three times for 3 min each in 1xPBS/0.1% Tween 20® at mild shaking and room temperature.

The incubation and the following washing step is repeated with the fluorescently coupled secondary antibody.

During the last washing step with 1xPBS/0.1% Tween 20®, the sample is washed with 1xPBS where DAPI is added at a 1:10,000 dilution for 10 min at room temperature and protected from light.

After 10 min, the coverslip is washed with 1xPBS for 3 min and rinsed three times with water.

When the coverslip has dried, it can be mounted with Vectashield® and analyzed under the fluorescence microscope.

#### **4.2.5.3 Fluorescence *in situ* hybridization - FISH**

##### **DNA probe preparation for hybridization**

The labelling of the DNA probe can be achieved both by random prime labelling and nick translation, the latter one being used here based on a protocol developed in the laboratory of Thomas Cremer [31].

Nick translation in 50 µL:

---

---

H <sub>2</sub> O	50 - (18+ x ) $\mu$ L
$\lambda$ DNA ( 2 $\mu$ g)	x $\mu$ L
10 x NT reaction buffer	5 $\mu$ L
0.1 M 2-Mercaptoethanol	5 $\mu$ L
1 mM dNTP mixture	5 $\mu$ L
1 mM labeled dUTP	1 $\mu$ L
DNA polymerase I	1 $\mu$ L
DNaseI (1:100 dilution in H <sub>2</sub> O)	1 $\mu$ L

---

Labeled dUTPs were either Biotin-16-dUTP or Cy3-dUTP as well as digoxigenin-dUTP depending on the kind of experiment. The reaction is incubated at 16°C for 60 min.

After 60 min, 5  $\mu$ L of the nick translation reaction mix are checked on a 1.5% agarose gel with 1x SYBR Safe DNA stain and visualized on a UV detector. The reaction mix is stored at 4°C during analysis time. Fragments should ideally range between 200 and 1000 bp in size.

If fragments are too large, the reaction mix is placed again at 16°C and further incubated until the right size of the fragments is reached.

100  $\mu$ L nick translation stop mixture is added (storage at -20°C is possible at this step).

The following reaction to precipitate DNA is set up:

---

H <sub>2</sub> O	96 $\mu$ L
Nick translation reaction mix	100 $\mu$ L
salmon sperm DNA	4 $\mu$ L
7.5 M ammonium-acetate	100 $\mu$ L
Absolute ethanol	400 $\mu$ L

---

The sample is vortexed briefly, subsequently centrifuged at full speed (>10,000 rcf) for 20 min at room temperature and the supernatant is removed carefully.

The pellet is washed by adding 500  $\mu$ L 70 % ethanol, followed by centrifugation at full speed for 10 min at room temperature.

The supernatant is removed carefully and the pellet is dried for a few minutes.

The pellet is dissolved in 20  $\mu$ L formamide at 1 h and 50°C with vigorous shaking.

After 1 hour, 20  $\mu$ L of prewarmed hybridization mix is added and shaken for another 1 hour at 50°C.

The DNA probe is denatured by boiling at 95°C for 5 min and the probe is stored at -20°C up to several months.

### Metaphase-FISH

Human male and female metaphase Slides were purchased from Applied Genetics Laboratories Inc. (Catalog#:HMM, Lot#: 111111, 101111 and 072511). They are prepared from PHA-stimulated blood lymphocytes cultured for 72 hours, synchronized and cultured for additional 5-7 hours before continuing with standard cytogenetic slide preparation methods [84].

Slides are stored at  $-20^{\circ}\text{C}$  and ready for subsequent fluorescence in situ hybridization. After thawing, the spots with cells are marked using a diamond glass cutter.

After rinsing the slide in 2xSSC and washing for 5 min in 50% Formamide/2xSSC, hybridization probe is added and covered with a coverslip. Attention is paid that no air bubbles remain before it is sealed with rubber cement.

When the rubber cement is dry, the DNA is denatured for 3 min at  $75^{\circ}\text{C}$  and hybridized over night at  $37^{\circ}\text{C}$  in a humidified chamber.

The next day, DNA is detected as described in 4.2.5.3 or the probe is directly labeled with a fluorescence dye.

### 3D-FISH

After fixation of cells (see 4.2.5.1), a probe could be mounted directly on a coverslip and sealed with rubber cement.

Then, both - probe and chromosomes - are denatured on a hot block simultaneously, 2-3 min at  $75^{\circ}\text{C}$ . The probe is either directly labeled or detected via fluorescently labeled antibodies (detection as described in 4.2.5.3) and analysed under a fluorescence confocal microscope.

### Immuno DNA-FISH

Immuno DNA-FISH is the combination of Immunofluorescence with FISH DNA detection. In the presented work, Immuno DNA-FISH was performed on commercially purchased lymphocyte metaphase spreads. The slides are treated with 1M HCl for 40 min to make the epitope for the antibody against 5mC accessible. After neutralization with 1xTBE for 5min, the DNA mark is detected by M $\alpha$ -33D3 (1:50 dilution in BlockAid) and one layer secondary antibody G $\alpha$ M-Cy3 or -Cy5, each for 45min at  $37^{\circ}\text{C}$  as described in 4.2.5.2. Then, the immunofluorescence protocol was combined with the FISH-protocol (see 4.2.5.3).

As hybridization probe for satellite-2 DNA, commercially available biotinylated LNA probes for satellite-2 DNA (Exiqon) was used and detected the next day with one layer of StrA488.

In the case of rDNA, the IGS region was hybridized using a mixture of two probes Hsr4 and Hsr6, both labelled directly with Cy3.

### 4.2.6 Dynamic molecular combing variants

Dynamic molecular combing steps such as DNA plug preparation, DNA isolation and the combing process essentially were performed according to [29].

#### 4.2.6.1 DNA plug preparation of adherent cells

The medium is removed, cells are washed with 1xPBS and 1 mL of Trypsin/EDTA is added to the plate. When cells are detached from the bottom of the flask, add 4 mL of medium (for a 10 cm plate).

The cell suspension is put in a 15 mL centrifugal tube and mixed by repetitive (ca. 5 times) suction and expelling of the suspension using a 5 mL pipette.

10  $\mu\text{L}$  of the suspension are applied to a counting chamber while the rest of the cell suspension is spun down at 1,000 rpm for 5 min at 4°C.

According to the cell number and the cells per plug, the number of plugs to be prepared from the cells can be calculated.

#### 4.2.6.2 DNA isolation for combing

One can start this process immediately after DNA plug preparation. In that case, step 1 can be skipped. If plugs were stored in EDTA, plugs are washed 3 times in TE 10.1 for 30 minutes each with slow rotation at room temperature.

TE 40.2 + YOYO-1 solution is prepared by adding 0.3  $\mu\text{L}$  of YOYO-1 stock solution to 100  $\mu\text{L}$  of TE 40.2 for each plug. This should be done just before using it.

One plug and 100  $\mu\text{L}$  of TE/YOYO-1 mixture is put in one round-bottom 2mL tube where it is protected from light and kept for 1 hour at room temperature

TE40.2/YOYO-1 mixture is removed carefully and 1 mL of MES buffer (0.5M, pH5.5) is added to the tube.

The tube is heated at 68°C for 20 minutes without shaking. Handle the solution with care, because shaking leads to unwanted stretching artefacts.

The tube is incubated at 42°C for 10 minutes in thermomixer without shaking. This step can be longer. It is crucial to reach 42°C in the solution, because higher temperature will inactivate the  $\beta$ -agarase enzyme.

1.5  $\mu\text{L}$  of  $\beta$ -agarase solution is taken by micropipette and kept in ambient air for a few seconds to warm up (addition of cold  $\beta$ -agarase solution to melted agarose solution might generate microaggregates of agarose which cannot be digested by  $\beta$ -agarase).

$\beta$ -agarase solution is added to the tube (Do not mix the solution, let  $\beta$ -agarase diffuse spontaneously.) and the solution is kept at 42°C overnight in the incubator.

The digested solution (1 mL) is poured into a reservoir which already has 500 $\mu\text{L}$  of MES buffer. To increase level of solution, appropriate volume of MES buffer (0.5-1.0



mL) can be added to the reservoir. Do not rush the additional MES solution. Extra MES is added very careful using a syringe putted to the bottom of the reservoir.

The prepared DNA can be stored up to several months at 4°C. In case of longer storage time, it is better to store it in a plastic tube. If the solution is used frequently, it can be stored in the teflon reservoir. In general: avoid repeated transfer and any other manipulations such as freezing, which could break the high molecular weight genomic DNA.

#### **4.2.6.3 The combing process**

The filled Teflon reservoirs of the Molecular Combing system get attached with the device as well as GV silanized coverslips wearing powder-free gloves.

After turning on the combing machine, the DNA is stretched on silanized coverslips by using the standard settings of the Molecular Combing System, i.e. 5 min incubation time and 300  $\mu\text{m/s}$  speed for the removal of the surface from the DNA solution.

Optional: The slides may be cut after this step to save material.

The DNA is fixed on the glass surface by incubating at 65°C for 4 hours.

Then, coverslips are adhered with a tiny drop of cyanoacrylate glue on a slide and covered with YOYO-1 diluted in Vectashield® 1: 2000 by using another 22x22 mm coverslip.

Then, the DNA quality is checked under the microscope. If the quality is sufficient, more slides are prepared for analysis. Teflon reservoirs are sealed with parafilm after use and stored at 4°C.

The microscopy slide can now be stored at -20°C for a few weeks (about 3 weeks) or further processed with fluorescence *in situ* hybridization to detect the DNA.

#### **4.2.6.4 *In vitro* psoralen-combing**

Psoralen is dissolved in ethanol to a final concentration of 0.2 mg/mL and stirred for approximately 30 min prior to use.

#### **Psoralen treatment of $\lambda$ DNA**

$\lambda$  DNA is dissolved to a concentration of 50 ng/L in 1xPBS to a total volume of 200  $\mu\text{L}$  (1.6nM) and the DNA solution is filled into a plate with 200  $\text{mm}^2$  surface (e.g. 24-well plate). In the case of irradiation of a low melting agarose plug, the plug is placed on a glass microscope coverslip and the respective amount of psoralen is applied directly on the plug.

The UVA-lamp is switched on about 10 min before use to reach constantly 8mW/cm<sup>2</sup> at 2-3 cm distance from the lamp.

Aluminium foil is wrapped around the metal block and placed in an ice bath to cool down. The foil will reflect UV light during radiation.

The 24-well plate with the DNA solution is placed on the metal block in the ice bath and the cover of the plate is removed.

10  $\mu\text{L}$  psoralen mixture is added to the solution or 10  $\mu\text{L}$  ethanol to the control (1/20 of total volume), respectively, the plate is shaken to mix the solutions and incubated for 5min.

The plate on the metal block in the ice bath is put under the UV lamp and irradiated for 5 min.

The plate is removed and 10  $\mu\text{L}$  psoralen or ethanol are added again to the samples, respectively, and the irradiation for 5min is repeated without prior incubation. This step is repeated twice for a total of 4 x 5 min UV irradiation.

The crosslink of the DNA treated with psoralen is checked compared to the control on a 0.5% agarose gel in 1xTBE. The samples are ready for the process of Dynamic Molecular Combing or can be stored at  $-20^{\circ}\text{C}$ .

For combing analysis, each  $\lambda$  DNA sample is diluted with 0.5 M MES buffer pH 5.5 to obtain a concentration of 16 pM in 2 mL final volume and then, the slides were further processed according to the standard combing protocol with subsequent fluorescence *in situ* hybridization.

### **Psoralen treatment of DNA in agarose plugs**

In the case of nucleosome-free DNA, the plug was washed with 1xPBS for 30min on a rotating wheel at RT to remove 0.5M EDTA. Afterwards, it was placed on a glass coverslip and residual liquid was removed. Then, 1  $\mu\text{g}$  PB diluted in ethanol was added (= 5 $\mu\text{L}$  of a 0.2 mg/mL solution, since the plug has a volume of about 100  $\mu\text{L}$ ) and the plug was incubated for 10 min at RT. Afterwards, the plug was irradiated with UVA light as described (4.2.6.4) for different irradiation times as indicated. Before each irradiation step, fresh PB (1  $\mu\text{g}$  each) was added without further incubation. The plug was washed after irradiation for 3x30min in 10.1 TE buffer and the DNA was isolated according to 4.2.6.2 and combed as described in 4.2.6.3. Afterwards, PB:DNA adducts are detected by three layers of antibodies (StrA488 - Rb $\alpha$ -Str-biotin - StrA488) (see 4.2.6.7 starting with DNA detection) and analyzed under a fluorescence microscope.

### **Psoralen treatment of cells lysed with 1% sarcosyl**

HCT116 cells were harvested according to 4.2.6.1 and embedded in low melting agarose plugs (100,000 cells/plug). Then, the cells were lysed as described for 30 min at RT with 0.5M EDTA containing 1% sarcosyl and no Proteinase K. Afterwards, plugs were washed for 10min in 1xPBS at RT on a rotating wheel and placed on a glass coverslip and treated with psoralen as described in 4.2.6.4.

### **Psoralen treatment of cells lysed with NP40**

HCT116 cells are embedded in a low melting agarose plug according to 4.2.6.1, cells are lysed in the plug for 30 min at RT with a buffer containing 50mM Tris and 0.5% NP40. Plugs are washed for 10 min at RT on a rotating wheel. Then, the DNA in the plugs is irradiated for 4x5min as described, washed for 3x30min with 10.1 TE and subsequently, the DNA is isolated according to 4.2.6.2. After combing and detection of psoralen-biotin:DNA products using three layers of antibodies (StrA488 - Rb $\alpha$ -Str-biotin - StrA488) (see 4.2.6.7 starting with DNA detection), the slide is analyzed with a fluorescence microscope.

#### **4.2.6.5 Chromatin-combing**

For chromatin-combing, cells were fixed by addition of formaldehyde (37% liquid stock solution) directly to the culture medium to a final concentration of 1%, followed by a 10 min incubation at RT. The crosslink is stopped by addition of 1 M glycine to a final concentration of 125 mM directly to the culture medium and incubation for 5 min at RT on a rocking platform. After aspirating the medium, cells are washed twice with 10 ml ice-cold 1xPBS and transferred into a 15 ml tube by scraping off, followed by centrifugation for 5 min at 700 rcf at 4°C. After centrifugation, the supernatant is removed and cells are resuspended in the appropriate volume of 1xPBS for further analysis.

Cells are then ready to be embedded in low melting 38 2.1 Single molecule epigenomics by dynamic molecular combing agarose plugs according to 4.2.6.1 and lysed by incubation of 0.5M EDTA supplemented with 1% sarcosyl for 30 min at RT. Plugs are washed with 10.1 TE for 3x30 min at RT on a rotating wheel and finally, plugs are added to 2 mL of 0.5M MES pH 5.5. The sample is then heated in a microwave at a pre-warmed microwave at 600W for 30 sec, which dissolves the low melting agarose plug. The mixture is then poured in a teflon reservoir and combed onto a silanized glass surface (described in 4.2.6.3). The surface is dried for 1h at 40 °C and subsequently, proteins are detected by immunofluorescence (see 4.2.5.2).

#### **4.2.6.6 Epi-combing**

Plugs are prepared and DNA is isolated as described before. After the slides are combed as explained earlier in the section, they are subjected to first, detection of DNA methylation and second, detection of the locus of interest.

Note that all steps except antibody incubation are performed with mild shaking of the slides in a Coplin Jar unless otherwise stated.

For the detection of DNA methylation with an 5mC-antibody, the cells are first incubated for one hour in 1M HCl. This step depurinates DNA and makes the DNA methylation accessible for the antibody.

After incubation, the slide is rinsed three times with water and neutralized for 5 min with 1xTBE.

After rinsing three times in water, the slide is washed for 3 min in 1xPBS.

Note: If the DNA modification is to be detected with an MBD2-Fc-Fragment, the previous steps are skipped and the procedure is started with washing in 1xPBS.

After that, the slides are washed 3 min each with three successive ethanol steps (70%, 90% and 100% ethanol) to dehydrate the DNA.

When the slide is dry, the antibody with the respective dilution in BlockAid is pipeted on the slide and covered with a cover slip of the same size. This is followed by a 20 min incubation at 37°C in a humidified chamber using a wet tissue.

After removing the cover slip with forceps, the slide is washed three times for 3 min each with 2xSSC/1% Tween 20<sup>®</sup>. This washing step is done after each round of antibody incubation.

The respective secondary, fluorescently labelled antibody is applied in a 1:50 dilution in BlockAid onto the slide, covered and incubated for 20 min for the second round.

To increase the signal intensity, a third layer of fluorescently coupled antibody may be of advantage. After the final wash in 2xSSC/1% Tween 20, the slide is rinsed in 1xPBS. Then, the sample is incubated for 10 min with 4% PFA/1xPBS to fix the antibody. The incubation is followed by two steps of washing in 1xPBS for 3 each.

The slides are rinsed in 2xSSC and ready for fluorescence *in situ* hybridization.

#### 4.2.6.7 Fluorescence *in situ* hybridization for all combing variants

After rinsing the slides with 2xSSC, the slides are washed with 50% FA/2xSSC for 5 min at room temperature.

The excess liquid on the slide is removed, 10 µL of the respective probe is applied on the slide and covered with a coverslip without remaining air bubbles. The edges of the coverslip are sealed with rubber cement. (for probe preparation see 4.2.5.3)

When the rubber cement is dry (meaning it got clear), the slide is denatured at 95°C for 3 min. Note that this step is only 1.5 min when the 12.5 kb long PCR fragment of hydroxymethylated DNA is to be detected.

The slide is hybridized over night at 37°C in a humidified chamber.

The next day, the slide is rinsed in 2xSSC and the upper coverslip is removed carefully.

Then, the slide is washed in 50%FA/2xSSC for three times 5 min with vigorous shaking at room temperature.

The slide is washed with 2xSSC for three times 3 min each at room temperature.

Now, the DNA can be detected: 15µL of antibody in BlockAid is added, covered with a coverslip and incubated for 20 min at 37°C in a humidified chamber.

After each layer of antibody, the slide is washed with 2xSSC/1% - Tween 20<sup>®</sup> for three times 3 min each.

In general, three layers of antibody or more is needed for a single molecule detection of stretched DNA fibers.

After final round of antibody incubation and 2xSSC/1% Tween 20<sup>®</sup>, the slides are washed with 1xPBS for 3 min at room temperature.

Then, the DNA is dehydrated with three successive ethanol solutions (70, 90 and 100%) for 3 minutes each.

The slides are finally mounted in Vectashield<sup>®</sup> and the edges are sealed with nail polish.

After analyzing the slides, they can be stored at 4°C.

### 4.2.7 Nuclear matrix preparation

Well established methods were applied to the isolation of nuclear matrix as described in [55].

$2.5 - 5 \cdot 10^6$  cells are washed in 1 mL 1xPBS pH 7.4 (5 min, 1000 rcf centrifugation) and cells are extracted in 200  $\mu$ L cytoskeleton buffer (10 mM Pipes, pH 6.8, 100 mM NaCl, 300 mM sucrose, 3 mM MgCl<sub>2</sub>, 1 mM EGTA, supplemented with Protease Inhibitor Cocktail (Roche), 1mM TCEP and 0.5% (vol/vol) Triton X-100).

After 5 min incubation at 4°C, soluble cytoplasmic proteins are separated by centrifugation at 5,000 rcf for 3 min. (supernatant = "CP" cytoplasmic fraction).

Chromatin is solubilized by DNA digestion with 400 U of RNase-free DNase I (Roche) in 110  $\mu$ L CSK buffer plus protease inhibitors for 90 min at 37°C with shaking at 300 rpm.

Then 50  $\mu$ L ammonium sulfate was added from a 1 M stock solution in CSK buffer to a final concentration of 0.25 M and, after 5 min incubation at 4°C on rotating wheel, samples were pelleted again by centrifugation at 5,000 rcf for 3 min. (supernatant = "CHR" chromatin fraction).

The pellet is further extracted with 100  $\mu$ L 2 M NaCl in CSK buffer for 10 min at 4°C on rotating wheel, and then centrifuged at 5,000 rcf for 3 min. This treatment removes all accessible DNA and histones from the nucleus (supernatant = "2M" 2M salt wash fraction, which is diluted 1:2 with water before SDS-PAGE).

The remaining pellet is solubilized in 200  $\mu$ L 8M urea buffer (10 mM Tris-HCl, 100 mM NaH<sub>2</sub>PO<sub>4</sub> x H<sub>2</sub>O, 8M urea; pH 8.0) and is considered the nuclear matrix-containing fraction ("NM").

100  $\mu$ L of each fraction is dialysed against 1xTE buffer and subsequently digested with RNaseA (37°C, 1h) and proteinaseK (50°C o.n.), precipitated, dissolved in 50  $\mu$ L ddH<sub>2</sub>O, controlled on 1% agarose gel (10  $\mu$ L) and the "NM" fraction was subjected to qPCR analysis.

25  $\mu$ L 5x Lämmli buffer is added to the remaining 100  $\mu$ L of the fractions, and each 20  $\mu$ L were subjected to Western blot analysis or Coomassie staining after boiling the probes at 95°C, 10 min and separating on denaturing SDS-polyacrylamide gels.

## 4.2.8 Microscale thermophoresis

### 4.2.8.1 Microscale thermophoresis to quantify DNA:protein binding interactions

A method that enables this quantitative analysis of DNA:protein interactions is called Microscale thermophoresis (MST), which measures the migration of molecules in a thermal gradient (see figure 4.2) [167]. The technique is based on the thermophoresis of molecules, which provides information about molecule size, charge, and hydration shell. Typically, at least one of these parameters is affected upon binding, hence, the method can be used for the analysis of biomolecular interactions [4].

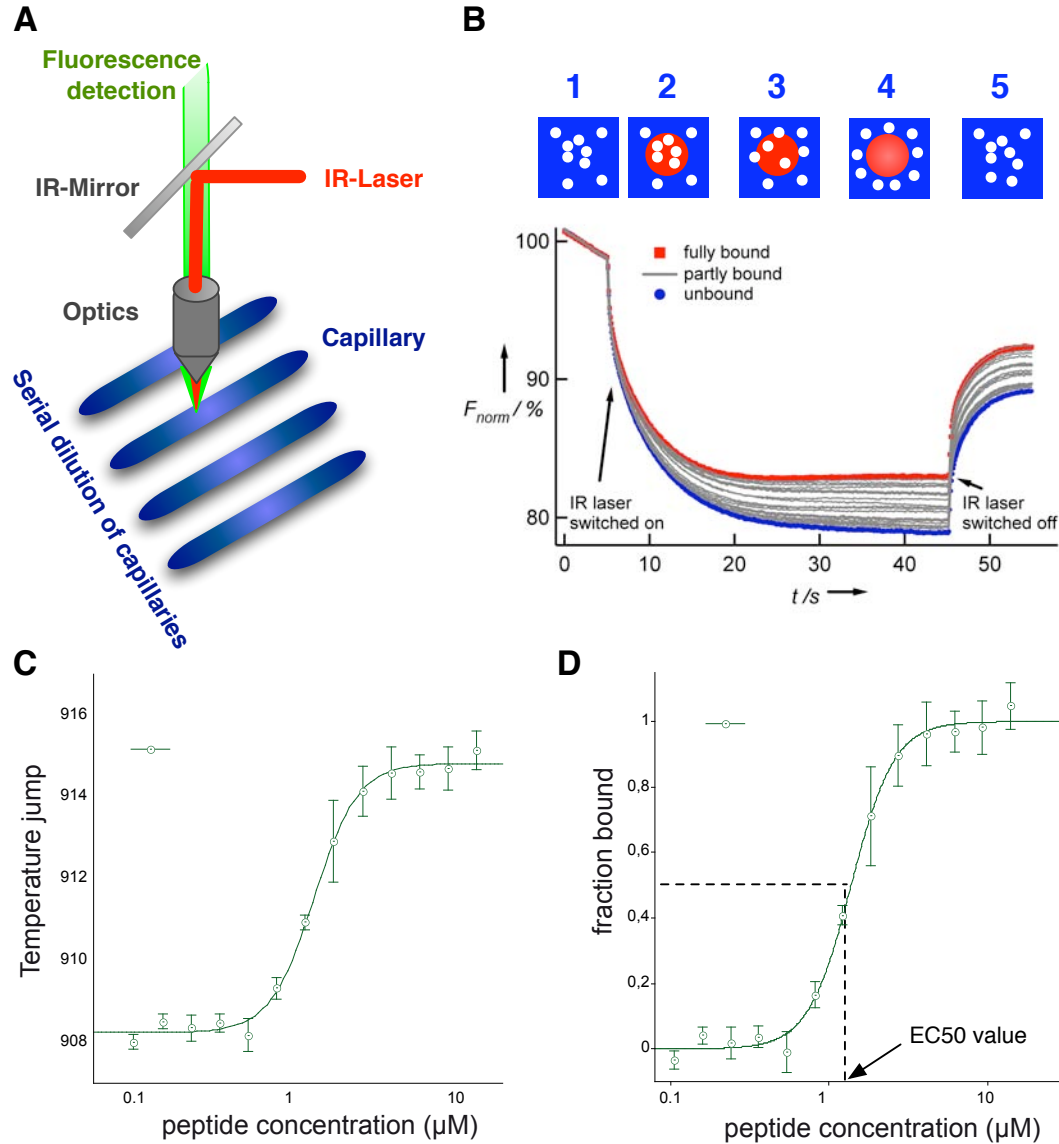
Quantitative binding parameters are obtained by using a serial dilution of the binding substrate. In a typical MST-experiment, the concentration of the fluorescently labelled molecule (in this case Cy3-labelled-dsDNA) is kept at a constant concentration and the unlabelled molecule (here the unlabeled peptide) is titrated in different capillaries until the saturation of the binding is achieved.

An IR-laser generates a precise temperature gradient in a glass capillary which is filled with the serial dilution of the titration. The temperature in this spot is increased up to 8 K and thus, a thermal gradient is created which induces the movement of molecules.

The fluorescence inside the capillary is imaged using an epifluorescence microscope, and the normalized fluorescence in the heated spot is plotted against time (see figure 4.2 B). The IR laser is switched on at  $t=10$  s, the fluorescence decreases as the temperature increases, and the labelled dsDNA molecules move away from the heated spot because of thermophoresis. When the IR laser is switched off, the molecules diffuse back to its initial state.

For data analysis, at first the different effects of Microscale Temperature Jump (MST T-Jump) and MST have to be identified according to their different time scales (see figure 4.2 B): The fast MST T-Jump equilibrates within 1 s after switching on the IR-Laser. This time scale is determined by the heat conductivity of water and the thickness of the capillaries. Thermophoresis equilibration is accomplished only after the establishment of the temperature difference. The time scale of MST equilibration takes about 10 s and is determined by the diffusion constant of the labelled molecules. The analysis software automatically detects the corresponding data points and the data can be fitted to analyze binding affinities.

The Hill-equation  $1/(1 + (EC50/c)^n)$  is used if the slope of the binding curve is too steep ( $n > 1$ ) or too smooth ( $n < 1$ ) for using a law of mass action fit. A Hill-coefficient  $n < 1$  or  $n > 1$  can indicate that the binding reaction is more complex than one molecule A in one conformation binds to one molecule B in one conformation. The binding affinity EC50 is the value at which 50% of the labelled molecules are bound to their targets, thus reasonable EC50 values have to be equal or bigger than 50% of the provided constant concentration of labelled molecule.



**Figure 4.2: Principle of Microscale thermophoresis**

**A:** Experimental set up of MST: The aqueous solution inside the capillary is locally heated with a focused IR-laser, which is coupled to an epifluorescence set up using an IR mirror. **B:** The fluorescence inside the capillary is measured for each capillary and the normalized fluorescence in the heated spot is plotted against time. The IR-laser is switched on at  $t = 10$  s and the fluorescence changes as the temperature increases. There are two effects, separated by their time scales, contributing to the new fluorescence distribution: the fast temperature jump (time scale about 1 s) and the thermophoretic movement (time scale about 10 s). Once the IR-laser is switched off, the molecules diffuse back to its initial state. The curves representing full binding and no binding are labelled in red and blue, respectively, whereas the grey curves represent partial binding. **C:** The temperature jump signals of the analysis software are plotted against peptide concentration of the titration series. A constant temperature jump signal upon increasing peptide concentration indicates saturation of binding. **D:** Data are normalized to fraction bound ranging from no bound DNA molecules ( $=0$ ) to all DNA molecules are bound by the peptide ( $=1$ ). EC50 shows the value where 50% of the DNA molecules are bound by the peptide which is an indicator for binding affinities. **C:** - **D:** All data points were plotted with the Hill equation.

For comparison between different experiments, the data can be normalized to fraction of bound molecules (FB) by the following equation:  $FB = (value(c) - free)/(complexed - free)$ , where  $value(c)$  is the MST-value measured for the concentration  $c$ ,  $free$  is the MST value for the unbound state (lowest concentration) and  $complexed$  is the MST-value for the fully bound state. The value 1 denotes that 100% of fluorescently labelled dsDNA is bound by the peptide and saturation is reached (out of [167]).

This method enables a specific and sensitive measurement of DNA:protein binding interactions in aqueous solution and is able to determine the DNA binding domains of Tip5.

#### 4.2.8.2 Preparation of DNA template

Double stranded DNA substrate molecules are annealed from single-stranded oligonucleotides. It is crucial for the experiment that the fluorescently labeled oligonucleotide is quantitatively incorporated into the DNA substrate. This is achieved by adding the unlabelled oligonucleotide at a 1.1 fold molar ratio with respect to the labeled oligonucleotide to the annealing reaction. The efficiency of the annealing reaction can be determined on a 15% native Polyacrylamide (PAA) gel that is first analysed on a fluorescence imager to reveal non-incorporated, fluorescently labeled oligonucleotides and second, post-stained with ethidium bromide. Oligonucleotides are dissolved according to the manufacturer's instructions and the nucleic acid concentration is measured using a UV/VIS Spectrophotometer.

550 pmol unlabelled oligonucleotides are mixed with 500 pmol Cy3-labelled oligonucleotide. Then, 5  $\mu$ L annealing buffer (10 x) is added and the volume is adjusted to 50  $\mu$ L with ddH<sub>2</sub>O to finally obtain a 10  $\mu$ M solution of double-stranded DNA.

The mixture is incubated for 15 min at 95 °C on a thermoblock, then the thermoblock is switched off to allow the reaction to slowly cool down until it reaches room temperature. The reaction can now be stored at -20 °C.

A 15% native PAA gel is prepared by the following scheme and quickly poured into an assembled gel chamber: 9 ml 30% Bis-Acrylamide, 9 ml 0.4 x TBE, 25  $\mu$ L APS, 5  $\mu$ L TEMED. A 10-well comb is positioned in the top of the gel. After the gel polymerized (60 min), the chamber is placed into the running cell, the comb is removed and the running cell is filled with 0.4 x TBE running buffer. To remove unpolymerized acrylamide, the gel is pre-run for 60 min at 120 V.

15 pmol of the annealing reaction, as well as 15 pmol of the single-stranded oligonucleotides are individually mixed with glycerol to reach a final concentration of 5% (v/v) glycerol. This will weigh down the sample and prevent the solution to mix with the buffer in the well.

All samples are carefully loaded together with the DNA ladder onto the pre-run gel, connected to a power supply and run at 4gc for 90 min at 120 V. Bromphenol Blue (usually present in the DNA marker) can be used as an indicator, as it migrates ahead of the single stranded oligonucleotides with an apparent molecular weight corresponding to an oligonucleotide of about 10 nt in length.



The gel is visualized with a fluorescence imager. The fuzzy oligonucleotide band has to be quantitatively shifted up in the annealing reaction, migrating as a defined band representing the double-stranded oligonucleotide.

Optionally, the efficiency of the annealing reaction is monitored by ethidium bromide staining. The gel is placed in the aqueous ethidium bromide solution and shaken for 10 min at room temperature. The gel can subsequently be visualized by a UV screen.

If free, labeled oligonucleotides are visible, the annealing reaction has to be repeated with an increased ratio of the fluorescently labeled oligonucleotide.

#### **4.2.8.3 Preparation of the titration series**

A titration series consists of up to 16 capillaries which are measured in a single thermophoresis run. Pipetting the samples and filling the capillaries will take about 30 min in total. In order to ease pipetting, the DNA substrate is diluted to a final concentration of 750 nM. The binding reaction contains 50 nM fluorescently labeled DNA with varying protein concentrations as indicated.

The concentration of the fluorescently labeled molecule should be close to the expected KD and in the range of 80-1,500 fluorescence counts. Dilutions of the unlabelled peptide should start at a concentration about 20-fold higher as the expected KD, being diluted to concentrations of about 0.01-fold of the expected KD.

The individual binding reactions should be prepared with a volume of 15  $\mu$ L, for the ease of pipetting and the minimization of experimental errors. However, a volume of only 5  $\mu$ L would be sufficient to fill the capillary. All solutions are mixed carefully by pipetting the reaction up and down, rather than vortexing the solution.

To ensure high accuracy between the individual samples, a Master Mix (MM) is prepared, just lacking the protein. An individual reaction mix contains 3  $\mu$ L MST buffer (5 x), 1  $\mu$ L annealed oligonucleotides (750 nM) and 1  $\mu$ L water to adjust the volume to 5  $\mu$ L. The MM has to be protected from light and can be stored on ice for several days. Prepare a serial dilution reducing the protein concentration 1.5 fold with every dilution step.

It is crucial to assure that the correct protein amount is pipetted because there is no possibility to normalize for protein concentrations in contrast to the fluorescent DNA. Therefore, it is beneficial to add firstly 5  $\mu$ L of MM and then 10  $\mu$ L of the protein solution for each dilution step. The reaction is mixed well by pipetting up and down.

Before filling a capillary, the reaction must be in equilibrium which only takes one minute for the presented kind of interaction. From now on, all steps including the work with capillaries are performed with powder free gloves to prevent impurities and adverse effects on the glass surface. In addition, only the ends and not the middle part of the capillaries where the observation field is located are touched.

To fill a capillary, the capillary is dipped into the sample. Care should be taken that the capillary does not touch the surface of the reaction tube, since adhering

molecules may falsify the measurement. Loaded capillaries are sealed on both ends by shortly sticking them into wax (provided with the capillaries).

All capillaries are then inserted into the metal tray and analyzed using the NanoTemper Monolith and the MST-data acquisition software ("Titration").

#### 4.2.8.4 Measurement and analysis of the binding affinity

After starting the NanoTemper Software "Titration", the LED "green" is selected for Cy3-dyes and the capillaries are automatically identified by clicking on the button "find capillaries" (initial settings: LED Power: 50%, Laser Voltage 10%). During capillary scan, the fluorescence signal should be between 100 and 2,000 fluorescence units.

Since all samples contain the same concentration of fluorescently labeled DNA, individual differences in intensity should be maximally 10%. After the identification of capillary positions, the Laser is set to "off" for 10 s and "on" for 40 s, guaranteeing sufficient time for thermophoretic movement and optimal thermophoretic resolution.

After updating the software with the unlabelled protein concentrations, the destination folder is selected, the experiment is saved and the measurement is initiated. With the stated settings, a run will be completed in 10 to 15 minutes. When opening the Start-window, it is possible to adjust the number of measurements and the strength/level of the temperature field induced by the IR-laser: 10% (low IR-laser power), 40% (moderate), 100% (high).

Immediately after starting the measurement, the data are ready for analysis with the provided MST-analysis software. A binding curve is plotted using the normalized fluorescence of the labeled dsDNA at different concentrations of the unlabelled peptide.

The analysis software automatically detects the Temperature Jump signal and the Thermophoresis signal and the respective data points are plotted against the concentration of the binding partner.

For quantifying the interaction affinity, the value of the dissociation constant  $K_D$  or the  $EC_{50}$  value are calculated using the NanoTemper Analysis software: The "load folder" button is pushed to import the measured data.

The "select all" button is pushed.

By selecting Thermophoresis, Thermophoresis and Temperature jump, Temperature jump, Fluorescence, the appropriate parameter will appear and analyzed.

By using the Fit-window, the data using the parameter of choice are fitted. To determine the dissociation constant  $K_D$ , the data are fitted using MST-standard fit algorithms (law of mass action) or Hill-algorithms.

The data are mostly best fitted using the Hill-equation, because the Hill-equation provides one more free parameter  $n$  (the Hill-coefficient). From the Hill equation the  $EC_{50}$ -affinity value can be determined.

The fitted data can be saved as an image or alternatively be exported as a text file. For comparison, the data can be normalized to the fraction of complexed molecules (FB) by the following equation:

$$FB = \frac{(value(c) - free)}{(complexed - free)} \quad (4.4)$$

where  $value(c)$  is the MST-value measured for the concentration  $c$ ,  $free$  is the MST-value for the unbound state (lowest concentration) and  $complexed$  is the MST-value for the fully bound state.

## 5 Appendix

### 5.1 Curriculum vitae

---

#### *Personal details*

Name, Surname	Karina Martina, Zillner
Date of birth	26. August 1984
Place of birth	Munich, Germany

---

#### *Education*

From 01/2013	Collège des Ingénieurs, Paris, France Management MBA Program for scientists
03/2011 - 11/2012	Bavarian EliteAcademy, Munich, Germany Education in Management, Finance and Leadership
Since 10/2009	RIGeL school, Regensburg, Germany PhD student in Cellular Biochemistry and Biophysics
10/2004 - 07/2009	University Regensburg, Germany  Studies in Biochemistry (Diploma), grade 1.1 (excellent)  Minor subject: Biophysics, grade 1.0 (excellent)  Pre-Diploma as best in class, grade 1.5 (excellent)
07/2004	Auersperg Gymnasium, Passau, Germany  high school graduation (A-levels) as best female and second best student, grade 1.1 (excellent)

**Work experience**

10/2008 - 07/2009	University Regensburg, Germany student research assistant in epigenetics, group Prof. Dr. Längst
04/2008 - 05/2008	University Regensburg, Germany internship in epigenetics, group Prof. Dr. Längst
02/2008 - 03/2008	University Regensburg, Germany internship in molecular biology, group Prof. Dr. Sumper
01/2008 - 02/2008	University Regensburg, Germany internship in radiation biology, group Prof. Dr. Dr. Kalbitzer
08/2007 - 12/2007	Shanghai Institute of Organic Chemistry, Shanghai, P.R. China research assistant in Medicinal Chemistry, group Prof. Dr. Ma
08/2005 - 10/2005	Center of Excellence for Fluorescent Bioanalytics, Germany internship in molecular biology, group Prof. Dr. Hegemann

---

**Skills and Expertise**

Languages	<b>German</b> (native fluency) <b>English</b> (fluency) 11/2008: TOEFL (111 of 120 points) 03/2004: CAE (grade B) <b>French</b> (conversational ability) <b>Mandarin</b> (conversational ability) 06/2008 - 07/2008: language stay in P.R. China
11/2008	<b>GRE subject test in Biochemistry, Molecular Biology and Cell Biology</b> result: 92% (among top 8% worldwide)

---

### ***Extracurricular Activities***

<b>Sports</b>	Volleyball (4th German league), Golf, Beachvolleyball
<b>Honorary Office</b>	2010 - 2012: Invited speaker for science education in high schools Organisation of several Biochemistry summer schools

## **5.2 Scientific contributions**

---

### ***List of publications***

<b>01/2013</b>	<b><i>Zillner K.</i></b> , Filarsky M., Rachow K., Weinberger M., Längst G., Németh A., Large-scale organization of ribosomal DNA chromatin is regulated by Tip5, Nucleic Acids Research (under revision)
<b>01/2013</b>	<b><i>Zillner K.</i></b> , Németh A., Psoralen-combing for analysis of nucleosomal patterns on single DNA fibers, Methods in Molecular Biology - Molecular Combing (accepted manuscript)
<b>08/2012</b>	<b><i>Zillner K.</i></b> , Németh A., Single molecule, genome-scale analyses of DNA modifications: Exposing the epigenome with next generation technologies, Epigenomics 4
<b>01/2012</b>	<b><i>Zillner K.</i></b> , Jerabek-Willemsen M., Duhr S., Braun D., Längst G., Baaske P., Microscale Thermophoresis as a sensitive method to quantify protein - nucleic acid interactions in solution, Methods in Molecular Biology - Functional genomics
<b>05/2008</b>	Ruecker O., <b><i>Zillner K.</i></b> , Groebner-Ferreira R., Heitzer M. , Gaussia-luciferase as a sensitive reporter gene for monitoring promoter activity in the nucleus of the green algae <i>Chlamydomonas reinhardtii</i> , Mol Genet Genomics

---

## ***Presentations***

### **Oral presentations**

- |                |   |
|----------------|---|
| <b>09/2012</b> | "Deciphering DNA methylation patterns on single DNA fibers, RIGeL summer school "Cellular Biochemistry and Biophysics", Kostenz, Germany                    |
| <b>05/2012</b> | "TeilMahl - sharing is caring", be.project Business competition German/European final, Berlin, Germany/Brussels Belgium                                     |
| <b>04/2012</b> | "Epi-combing, single molecule analysis of DNA methylation patterns", National Symposium on "Chromatin architecture and dynamics", Oppurg, Germany           |
| <b>09/2012</b> | "TeilMahl - sharing is caring", Conference "Strengthening Europe: Strategic partnerships and research focus in the age of globalization", Brussels, Belgium |
| <b>06/2011</b> | "Epi-combing - single molecule, genome scale analysis of DNA methylation patterns", 61st Lindau Nobel Laureate Meeting, Lindau Germany                      |

### **Poster presentations**

- |                |   |
|----------------|---|
| <b>10/2012</b> | "Deciphering DNA methylation patterns of repetitive DNA on single DNA fibers", Wellcome Trust Conference "Epigenomics of common diseases", Baltimore, Maryland, USA |
| <b>09/2012</b> | "Deciphering DNA methylation patterns on single DNA fibers - Epi-combing", IUBMB/FEBS congress, Seville, Spain  |
| <b>09/2012</b> | "Deciphering DNA methylation patterns on single DNA fibers", "From Single molecules to systems biology", Cadiz, Spain   |
| <b>10/2011</b> | "Epi-combing - single molecule analysis of DNA methylation patterns", "EMBO nuclear structure and dynamics", Isle sur la Sorgue, France                             |
| <b>10/2010</b> | "NoRC regulates the higher order structure of rDNA", 3rd SFB Symposium - "CHROMATIN - Assembly and Inheritance of Functional States" Munich, Germany                |

---

09/2010	"NoRC regulates higher order structures of rDNA, "RIGeL summer school in Cellular Biochemistry and Biophysics", Kostenz, Germany
08/2010	"NoRC regulates higher order structures of rDNA, "EMBO transcription and chromatin", Heidelberg, Germany

---

### ***Other scientific contributions***

03/2012	Interview and scientific article "Epigenetics - the fight against cancer", magazine "Arzt und Karriere"
10/2011	Nature video " A life in science - with Elizabeth H. Blackburn" by Nature video publishing group

## **5.3 Grants, awards and fellowships**

### ***Fellowships***

Since 10/2012	Academy for political Education, Tutzing, Germany fellowship "good science"
09/ 2012	International Union of Biochemistry and Molecular Biology fellowship for Young Scientist program (YSP) 2012
06/2011	Council of Lindau Nobel Laureate Meetings, Germany fellowship for the 61st Lindau Nobel Laureate Meeting
03/2011 - 11/2012	Bavarian EliteAcademy, Germany fellowship by the foundation of Bavarian economy and industry
Since 12/2010	e-fellows.net, Germany Online fellowship and career network
11/2010 - 10/2012	Elitenetwork of Bavaria, Germany Research fellowship for PhD studies



**Awards**

10/2012	Wellcome Trust Best poster presentation at conference "Epigenomics of common diseases"
09/2012	Federation of European Biochemical Societies Best meeting report of conference YSP/YSF 2012
09/2012	International Union of Biochemistry Best poster presentation at conference YSP/YSF 2012
09/2012	BIO-SECURITY 3rd prize business competition with project "Teil-Mahl"
05/2012	be.project European Business competition Winner of the German national final with the team "Teil-Mahl"
10/2007 - 09/2008	University Regensburg, Germany tuition fee waiver for best pre-diploma in class
07/2004	Auersperg Gymnasium Passau, Germany award for best female student of the year

---

**Grants**

09/2012 - 12/2012	graduate research academy "RNA biology" Regensburg, Germany grant for hiring a student research assistant
10/2012	Elitenetwork of Bavaria, Germany travel grant for Wellcome Trust conference "Epigenomics of common diseases"
09/2012	German society for Biochemistry (GBM) travel grant for participation in IUBMB/FEBS congress in Seville, Spain

---

**08/2007**

German Academic Exchange Service

travel grant for research stay in P.R. China

---

## 6 Acknowledgements

Here, I would like to seize the opportunity to thank all the people who made this PhD thesis possible. Sincere thanks are given to all.

First of all, my special thanks go to PD Dr. habil. Attila Németh who supported me during all these years. I am very grateful for the "fancy and hot" topic I was able to work on. "I have to thank YOU.... No, really, I have to thank YOU". Thank you for all the intellectual, the scientific, the technical and the personal help you gave me.

Second, I am very thankful to Prof. Dr. Gernot Längst who was an amazing patron during all the time. His aid gave me a great and fruitful education not only scientifically, but also personally. Additionally, I owe earnest thankfulness to Prof. Dr. Herbert Tschochner, who always had an open ear and an open door when I needed him. I would also like to thank my mentor Dr. Hermann Franz from Siemens who had the ability to listen carefully but also to ask the right questions at the right point of time.

Furthermore, I would like to show my gratitude to all the members of the house of the ribosome. You made me feel welcome from the first day on and I am proud to be a part of that team. Thank you for having many scientific discussions and even more cakes together.

I am also very thankful to the Längst lab who gave me in April 2008 a very warm welcome and a great time. You were one of the reasons why I stayed not only for the Diploma but also for the PhD thesis.

In particular, my sincere thanks go to Sarah. It was simply amazing to have my best friend working next door and I miss this atmosphere and above all, miss you, already. In addition, it is a great pleasure to thank the members of the command bridge: Josef, Katharina, Tom, Michael and Stefan as well as Raj. I had so many unforgettable moments with you, or to describe it with simple words: **Best time of my life.**

Last, I would like to mention my family. I owe you so much and it is impossible to list all of it. Let me just say: **Thank you for everything.**

# Bibliography

- [1] ANANIEV, Gene E. ; GOLDSTEIN, Steve ; RUNNHEIM, Rod ; FORREST, Dan K. ; ZHOU, Shiguo ; POTAMOUSIS, Konstantinos ; CHURAS, Chris P. ; BERGENDAHL, Veit ; THOMSON, James A. ; SCHWARTZ, David C.: Optical mapping discerns genome wide DNA methylation profiles. In: *BMC Mol Biol* 9 (2008), S. 68
- [2] ARAVIND, L ; LANDSMAN, D: AT-hook motifs identified in a wide variety of DNA-binding proteins. In: *Nucleic Acids Res* 26 (1998), Oct, Nr. 19, S. 4413–21
- [3] AVERY, O T. ; MACLEOD, C M. ; MCCARTY, M: STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. In: *J Exp Med* 79 (1944), Feb, Nr. 2, S. 137–58
- [4] BAASKE, Philipp ; WIENKEN, Christoph J. ; REINECK, Philipp ; DUHR, Stefan ; BRAUN, Dieter: Optical thermophoresis for quantifying the buffer dependence of aptamer binding. In: *Angew Chem Int Ed Engl* 49 (2010), Mar, Nr. 12, S. 2238–41
- [5] BALLESTAR, Esteban ; ESTELLER, Manel: SnapShot: the human DNA methylome in health and disease. In: *Cell* 135 (2008), Dec, Nr. 6, S. 1144–1144.e1
- [6] BANDYOPADHYAY, Debdutta ; GATZA, Catherine ; DONEHOWER, Lawrence A. ; MEDRANO, Estela E.: Analysis of cellular senescence in culture in vivo: the senescence-associated beta-galactosidase assay. In: *Curr Protoc Cell Biol* Chapter 18 (2005), Jul, S. Unit 18.9
- [7] BEREZNEY, R ; COFFEY, D S.: The nuclear protein matrix: isolation, structure, and functions. In: *Adv Enzyme Regul* 14 (1976), S. 63–100
- [8] BERMÚDEZ, Ignacio ; GARCÍA-MARTÍNEZ, José ; PÉREZ-ORTÍN, José E ; ROCA, Joaquim: A method for genome-wide analysis of DNA helical tension by means of psoralen-DNA photobinding. In: *Nucleic Acids Res* 38 (2010), Oct, Nr. 19, S. e182
- [9] BERNSTEIN, Bradley E. ; MEISSNER, Alexander ; LANDER, Eric S.: The mammalian epigenome. In: *Cell* 128 (2007), Feb, Nr. 4, S. 669–81
- [10] BIRCH, J. L. ; ZOMERDIJK, J. C.: Structure and function of ribosomal RNA gene chromatin. In: *Biochem. Soc. Trans.* 36 (2008), Aug, S. 619–624

- [11] BIRD, A P. ; SOUTHERN, E M.: Use of restriction enzymes to study eukaryotic DNA methylation: I. The methylation pattern in ribosomal DNA from *Xenopus laevis*. In: *J Mol Biol* 118 (1978), Jan, Nr. 1, S. 27–47
- [12] BIRD, Adrian: DNA methylation patterns and epigenetic memory. In: *Genes Dev* 16 (2002), Jan, Nr. 1, S. 6–21
- [13] BODE, J ; KOHWI, Y ; DICKINSON, L ; JOH, T ; KLEHR, D ; MIELKE, C ; KOHWI-SHIGEMATSU, T: Biological significance of unwinding capability of nuclear matrix-associating DNAs. In: *Science* 255 (1992), Jan, Nr. 5041, S. 195–7
- [14] BOLLA, R I. ; BRAATEN, D C. ; SHIOMI, Y ; HEBERT, M B. ; SCHLESSINGER, D: Localization of specific rDNA spacer sequences to the mouse L-cell nucleolar matrix. In: *Mol Cell Biol* 5 (1985), Jun, Nr. 6, S. 1287–94
- [15] BOLLATI, V ; GALIMBERTI, D ; PERGOLI, L ; DALLA VALLE, E ; BARRETTA, F ; CORTINI, F ; SCARPINI, E ; BERTAZZI, P A. ; BACCARELLI, A: DNA methylation in repetitive elements and Alzheimer disease. In: *Brain Behav Immun* 25 (2011), Aug, Nr. 6, S. 1078–83
- [16] BORST, Piet ; SABATINI, Robert: Base J: discovery, biosynthesis, and possible functions. In: *Annu Rev Microbiol* 62 (2008), S. 235–51
- [17] BOURACHOT, B ; YANIV, M ; MUCHARDT, C: The activity of mammalian brm/SNF2alpha is dependent on a high-mobility-group protein I/Y-like DNA binding domain. In: *Mol Cell Biol* 19 (1999), Jun, Nr. 6, S. 3931–9
- [18] BROCK, G J. ; BIRD, A: Mosaic methylation of the repeat unit of the human ribosomal RNA genes. In: *Hum Mol Genet* 6 (1997), Mar, Nr. 3, S. 451–6
- [19] BUSCH, H ; MURAMATSU, M ; ADAMS, H ; STEELE, W J. ; LIAU, M C. ; SMETANA, K: ISOLATION OF NUCLEOLI. In: *Exp Cell Res* 24 (1963), S. SUPPL9:150–63
- [20] CABURET, Sandrine ; CONTI, Chiara ; SCHURRA, Catherine ; LEBOWSKY, Ronald ; EDELSTEIN, Stuart J. ; BENSIMON, Aaron: Human ribosomal RNA gene arrays display a broad range of palindromic structures. In: *Genome Res* 15 (2005), Aug, Nr. 8, S. 1079–85
- [21] CALLINAN, Pauline A. ; FEINBERG, Andrew P.: The emerging science of epigenomics. In: *Hum Mol Genet* 15 Spec No 1 (2006), Apr, S. R95–101
- [22] CECH, T ; PARDUE, M L.: Cross-linking of DNA with trimethylpsoralen is a probe for chromatin structure. In: *Cell* 11 (1977), Jul, Nr. 3, S. 631–40
- [23] CECH, T ; POTTER, D ; PARDUE, M L.: Electron microscopy of DNA cross-linked with trimethylpsoralen: a probe for chromatin structure. In: *Biochemistry* 16 (1977), Nov, Nr. 24, S. 5313–21
- [24] CHAN, P K. ; ALDRICH, M ; BUSCH, H: Alterations in immunolocalization of the phosphoprotein B23 in HeLa cells during serum starvation. In: *Exp Cell Res* 161 (1985), Nov, Nr. 1, S. 101–10
- [25] CHENG, X: Structure and function of DNA methyltransferases. In: *Annu Rev Biophys Biomol Struct* 24 (1995), S. 293–318

- 
- [26] CHOI, Si H. ; WORSWICK, Scott ; BYUN, Hyang-Min ; SHEAR, Talia ; SOUSSA, John C. ; WOLFF, Erika M. ; DOUER, Dan ; GARCIA-MANERO, Guillermo ; LIANG, Gangning ; YANG, Allen S.: Changes in DNA methylation of tandem DNA repeats are different from interspersed repeats in cancer. In: *Int J Cancer* 125 (2009), Aug, Nr. 3, S. 723–9
- [27] COLLIER, Justine: Epigenetic regulation of the bacterial cell cycle. In: *Curr Opin Microbiol* 12 (2009), Dec, Nr. 6, S. 722–9
- [28] CONCONI, A ; WIDMER, R M. ; KOLLER, T ; SOGO, J M.: Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. In: *Cell* 57 (1989), Jun, Nr. 5, S. 753–61
- [29] CONTI, C ; CABURET, S ; SCHURRA, C ; BENSIMON, A: Molecular combing. In: *Curr Protoc Cytom* Chapter 8 (2001), May, S. Unit 8.10
- [30] CRAIG, J M. ; BOYLE, S ; PERRY, P ; BICKMORE, W A.: Scaffold attachments within the human genome. In: *J Cell Sci* 110 ( Pt 21) (1997), Nov, S. 2673–82
- [31] CREMER, Marion ; GRASSER, Florian ; LANCTÔT, Christian ; MÜLLER, Stefan ; NEUSSER, Michaela ; ZINNER, Roman ; SOLOVEI, Irina ; CREMER, Thomas: Multicolor 3D Fluorescence In Situ Hybridization for Imaging Interphase Chromosomes. In: *Methods Mol Biol* 463 (2008), S. 205–39
- [32] DIMRI, G P. ; LEE, X ; BASILE, G ; ACOSTA, M ; SCOTT, G ; ROSKELLEY, C ; MEDRANO, E E. ; LINSKENS, M ; RUBELJ, I ; PEREIRA-SMITH, O: A biomarker that identifies senescent human cells in culture and in aging skin in vivo. In: *Proc Natl Acad Sci U S A* 92 (1995), Sep, Nr. 20, S. 9363–7
- [33] DRUKER, R ; WHITELOW, E: Retrotransposon-derived elements in the mammalian genome: a potential source of disease. In: *J Inherit Metab Dis* 27 (2004), Nr. 3, S. 319–30
- [34] EKANAYAKE, Dilrukshi K. ; MINNING, Todd ; WEATHERLY, Brent ; GUNASEKERA, Kapila ; NILSSON, Daniel ; TARLETON, Rick ; OCHSENREITER, Torsten ; SABATINI, Robert: Epigenetic regulation of transcription and virulence in *Trypanosoma cruzi* by O-linked thymine glucosylation of DNA. In: *Mol Cell Biol* 31 (2011), Apr, Nr. 8, S. 1690–700
- [35] ESPADA, J. ; BALLESTAR, E. ; SANTORO, R. ; FRAGA, M. F. ; VILLAGAREA, A. ; NÉMETH, A. ; LOPEZ-SERRA, L. ; ROPERO, S. ; ARANDA, A. ; OROZCO, H. ; MORENO, V. ; JUARRANZ, A. ; STOCKERT, J. C. ; LÄNGST, G. ; GRUMMT, I. ; BICKMORE, W. ; ESTELLER, M.: Epigenetic disruption of ribosomal RNA genes and nucleolar architecture in DNA methyltransferase 1 (Dnmt1) deficient cells. In: *Nucleic Acids Res.* 35 (2007), S. 2191–2198
- [36] ESTELLER, Manel: Epigenetics in cancer. In: *N Engl J Med* 358 (2008), Mar, Nr. 11, S. 1148–59
- [37] FEINBERG, Andrew P.: Genome-scale approaches to the epigenetics of common human disease. In: *Virchows Arch* 456 (2010), Jan, Nr. 1, S. 13–21

- [38] FRAGA, M F. ; RODRÍGUEZ, R ; CAÑAL, M J.: Rapid quantification of DNA methylation by high performance capillary electrophoresis. In: *Electrophoresis* 21 (2000), Aug, Nr. 14, S. 2990–4
- [39] FRAGA, Mario F. ; URIOL, Esther ; BORJA DIEGO, L ; BERDASCO, María ; ESTELLER, Manel ; CAÑAL, María Jesús ; RODRÍGUEZ, Roberto: High-performance capillary electrophoretic method for the quantification of 5-methyl 2'-deoxycytidine in genomic DNA: application to plant, animal and human cancer tissues. In: *Electrophoresis* 23 (2002), Jun, Nr. 11, S. 1677–81
- [40] FROMMER, M ; McDONALD, L E. ; MILLAR, D S. ; COLLIS, C M. ; WATT, F ; GRIGG, G W. ; MOLLOY, P L. ; PAUL, C L.: A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. In: *Proc Natl Acad Sci U S A* 89 (1992), Mar, Nr. 5, S. 1827–31
- [41] FROMMER, M ; PROSSER, J ; TKACHUK, D ; REISNER, A H. ; VINCENT, P C.: Simple repeated sequences in human satellite DNA. In: *Nucleic Acids Res* 10 (1982), Jan, Nr. 2, S. 547–63
- [42] GHOSHAL, Kalpana ; MAJUMDER, Sarmila ; DATTA, Jharna ; MOTIWALA, Tasneem ; BAI, Shoumei ; SHARMA, Sudarshana M. ; FRANKEL, Wendy ; JACOB, Samson T.: Role of human ribosomal RNA (rRNA) promoter methylation and of methyl-CpG-binding protein MBD2 in the suppression of rRNA gene expression. In: *J Biol Chem* 279 (2004), Feb, Nr. 8, S. 6783–93
- [43] GIACOMODONATO, Mónica N ; SARNACKI, Sebastián H ; LLANA, Mariángel Noto ; CERQUETTI, M C.: Dam and its role in pathogenicity of *Salmonella enterica*. In: *J Infect Dev Ctries* 3 (2009), Nr. 7, S. 484–90
- [44] GOETZE, Hannah ; WITTNER, Manuel ; HAMPERL, Stephan ; HONDELE, Maria ; MERZ, Katharina ; STOECKL, Ulrike ; GRIESENBECK, Joachim: Alternative chromatin structures of the 35S rRNA genes in *Saccharomyces cerevisiae* provide a molecular basis for the selective recruitment of RNA polymerases I and II. In: *Mol Cell Biol* 30 (2010), Apr, Nr. 8, S. 2028–45
- [45] GRANICK, D: Nucleolar necklaces in chick embryo fibroblast cells. II. Microscope observations of the effect of adenosine analogues on nucleolar necklace formation. In: *J Cell Biol* 65 (1975), May, Nr. 2, S. 418–27
- [46] GREEF, Jessica C. de ; LEMMERS, Richard J L F. ; ENGELN, Baziel G M. van ; SACCONI, Sabrina ; VENANCE, Shannon L. ; FRANTS, Rune R. ; TAWIL, Rabi ; MAAREL, Silvére M van der: Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. In: *Hum Mutat* 30 (2009), Oct, Nr. 10, S. 1449–59
- [47] GRIESENBECK, Joachim ; WITTNER, Manuel ; CHARTON, Romain ; CONCONI, Antonio: Chromatin endogenous cleavage and psoralen crosslinking assays to analyze rRNA gene chromatin in vivo. In: *Methods Mol Biol* 809 (2012), S. 291–301
- [48] GRUMMT, I.: Different epigenetic layers engage in complex crosstalk to define the epigenetic state of mammalian rRNA genes. In: *Hum. Mol. Genet.* 16 Spec No 1 (2007), Apr, S. R21–27

- 
- [49] GWYNN, B ; LUEDERS, K ; SANDS, M S. ; BIRKENMEIER, E H.: Intracisternal A-particle element transposition into the murine beta-glucuronidase gene correlates with loss of enzyme activity: a new model for beta-glucuronidase deficiency in the C3H mouse. In: *Mol Cell Biol* 18 (1998), Nov, Nr. 11, S. 6474–81
- [50] HAAF, T ; WARD, D C.: Inhibition of RNA polymerase II transcription causes chromatin decondensation, loss of nucleolar structure, and dispersion of chromosomal domains. In: *Exp Cell Res* 224 (1996), Apr, Nr. 1, S. 163–73
- [51] HANCOCK, R: A new look at the nuclear matrix. In: *Chromosoma* 109 (2000), Jul, Nr. 4, S. 219–25
- [52] HANSEN, R S. ; STÖGER, R ; WIJMENGA, C ; STANEK, A M. ; CANFIELD, T K. ; LUO, P ; MATARAZZO, M R. ; D’ESPOSITO, M ; FEIL, R ; GIMELLI, G ; WEEMAES, C M. ; LAIRD, C D. ; GARTLER, S M.: Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. In: *Hum Mol Genet* 9 (2000), Nov, Nr. 18, S. 2575–87
- [53] HANSEN, R S. ; WIJMENGA, C ; LUO, P ; STANEK, A M. ; CANFIELD, T K. ; WEEMAES, C M. ; GARTLER, S M.: The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. In: *Proc Natl Acad Sci U S A* 96 (1999), Dec, Nr. 25, S. 14412–7
- [54] HASSAN, K M. ; NORWOOD, T ; GIMELLI, G ; GARTLER, S M. ; HANSEN, R S.: Satellite 2 methylation patterns in normal and ICF syndrome cells and association of hypomethylation with advanced replication. In: *Hum Genet* 109 (2001), Oct, Nr. 4, S. 452–62
- [55] HE, D C. ; NICKERSON, J A. ; PENMAN, S: Core filaments of the nuclear matrix. In: *J Cell Biol* 110 (1990), Mar, Nr. 3, S. 569–80
- [56] HORARD, Béatrice ; EYMERY, Angéline ; FOUREL, Geneviève ; VASSETZKY, Nikita ; PUECHBERTY, Jacques ; ROIZES, Gérard ; LEBRIGAND, Kevin ; BARBRY, Pascal ; LAUGRAUD, Aurélie ; GAUTIER, Christian ; SIMON, Elsa B. ; DEVAUX, Frédéric ; MAGDINIER, Frédérique ; VOURC’H, Claire ; GILSON, Eric: Global analysis of DNA methylation and transcription of human repetitive sequences. In: *Epigenetics* 4 (2009), Jul, Nr. 5, S. 339–50
- [57] HUANG, Yun ; PASTOR, William A. ; SHEN, Yinghua ; TAHILIANI, Mamta ; LIU, David R. ; RAO, Anjana: The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. In: *PLoS One* 5 (2010), Nr. 1, S. e8888
- [58] HUTH, J R. ; BEWLEY, C A. ; NISSEN, M S. ; EVANS, J N. ; REEVES, R ; GRONENBORN, A M. ; CLORE, G M.: The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. In: *Nat Struct Biol* 4 (1997), Aug, Nr. 8, S. 657–65
- [59] IATROPOULOS, M J. ; WILLIAMS, G M.: Proliferation markers. In: *Exp Toxicol Pathol* 48 (1996), Feb, Nr. 2-3, S. 175–81
- [60] ITO, Shinsuke ; SHEN, Li ; DAI, Qing ; WU, Susan C. ; COLLINS, Leonard B. ; SWENBERG, James A. ; HE, Chuan ; ZHANG, Yi: Tet proteins can convert



- 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. In: *Science* 333 (2011), Sep, Nr. 6047, S. 1300–3
- [61] JACKSON, D A. ; YUAN, J ; COOK, P R.: A gentle method for preparing cyto- and nucleo-skeletons and associated chromatin. In: *J Cell Sci* 90 ( Pt 3) (1988), Jul, S. 365–78
- [62] JACKSON, Kesmic ; YU, Mimi C. ; ARAKAWA, Kazuko ; FIALA, Emerich ; YOUN, Byungwoo ; FIEGL, Heidi ; MÜLLER-HOLZNER, Elisabeth ; WIDSCHWENDTER, Martin ; EHRLICH, Melanie: DNA hypomethylation is prevalent even in low-grade breast cancers. In: *Cancer Biol Ther* 3 (2004), Dec, Nr. 12, S. 1225–31
- [63] JEANPIERRE, M: Human satellites 2 and 3. In: *Ann Genet* 37 (1994), Nr. 4, S. 163–71
- [64] JEGOU, Thibaud ; CHUNG, Inn ; HEUVELMAN, Gerrit ; WACHSMUTH, Malte ; GÖRISCH, Sabine M. ; GREULICH-BODE, Karin M. ; BOUKAMP, Petra ; LICHTER, Peter ; RIPPE, Karsten: Dynamics of telomeres and promyelocytic leukemia nuclear bodies in a telomerase-negative human cell line. In: *Mol Biol Cell* 20 (2009), Apr, Nr. 7, S. 2070–82
- [65] JURKA, Jerzy ; KAPITONOV, Vladimir V. ; KOHANY, Oleksiy ; JURKA, Michael V.: Repetitive sequences in complex genomes: structure and evolution. In: *Annu Rev Genomics Hum Genet* 8 (2007), S. 241–59
- [66] KEPPEL, F: Transcribed human ribosomal RNA genes are attached to the nuclear matrix. In: *J Mol Biol* 187 (1986), Jan, Nr. 1, S. 15–21
- [67] KLOSE, R. J. ; BIRD, A. P.: Genomic DNA methylation: the mark and its mediators. In: *Trends Biochem. Sci.* 31 (2006), Feb, S. 89–97
- [68] KOMURA, J ; IKEHATA, H ; HOSOI, Y ; RIGGS, A D. ; ONO, T: Mapping psoralen cross-links at the nucleotide level in mammalian cells: suppression of cross-linking at transcription factor- or nucleosome-binding sites. In: *Biochemistry* 40 (2001), Apr, Nr. 13, S. 4096–105
- [69] KRAMER, J A. ; SINGH, G B. ; KRAWETZ, S A.: Computer-assisted search for sites of nuclear matrix attachment. In: *Genomics* 33 (1996), Apr, Nr. 2, S. 305–8
- [70] KRIAUCIONIS, Skirmantas ; HEINTZ, Nathaniel: The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. In: *Science* 324 (2009), May, Nr. 5929, S. 929–30
- [71] KUO, K C. ; MCCUNE, R A. ; GEHRKE, C W. ; MIDGETT, R ; EHRLICH, M: Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. In: *Nucleic Acids Res* 8 (1980), Oct, Nr. 20, S. 4763–76
- [72] KUSTER, J E. ; GUARNIERI, M H. ; AULT, J G. ; FLAHERTY, L ; SWIATEK, P J.: IAP insertion in the murine LamB3 gene results in junctional epidermolysis bullosa. In: *Mamm Genome* 8 (1997), Sep, Nr. 9, S. 673–81

- 
- [73] LAMPRECHT, Björn ; BONIFER, Constanze ; MATHAS, Stephan: Repeat-element driven activation of proto-oncogenes in human malignancies. In: *Cell Cycle* 9 (2010), Nov, Nr. 21, S. 4276–81
- [74] LISTER, Ryan ; PELIZZOLA, Mattia ; DOWEN, Robert H. ; HAWKINS, R D. ; HON, Gary ; TONTI-FILIPPINI, Julian ; NERY, Joseph R. ; LEE, Leonard ; YE, Zhen ; NGO, Que-Minh ; EDSALL, Lee ; ANTOSIEWICZ-BOURGET, Jessica ; STEWART, Ron ; RUOTTI, Victor ; MILLAR, A H. ; THOMSON, James A. ; REN, Bing ; ECKER, Joseph R.: Human DNA methylomes at base resolution show widespread epigenomic differences. In: *Nature* (2009), Oct
- [75] LISTER, Ryan ; PELIZZOLA, Mattia ; DOWEN, Robert H. ; HAWKINS, R D. ; HON, Gary ; TONTI-FILIPPINI, Julian ; NERY, Joseph R. ; LEE, Leonard ; YE, Zhen ; NGO, Que-Minh ; EDSALL, Lee ; ANTOSIEWICZ-BOURGET, Jessica ; STEWART, Ron ; RUOTTI, Victor ; MILLAR, A H. ; THOMSON, James A. ; REN, Bing ; ECKER, Joseph R.: Human DNA methylomes at base resolution show widespread epigenomic differences. In: *Nature* 462 (2009), Nov, Nr. 7271, S. 315–22
- [76] LITTLE, R D. ; PLATT, T H. ; SCHILDKRAUT, C L.: Initiation and termination of DNA replication in human rRNA genes. In: *Mol Cell Biol* 13 (1993), Oct, Nr. 10, S. 6600–13
- [77] LYLE, R ; WRIGHT, T J. ; CLARK, L N. ; HEWITT, J E.: The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. In: *Genomics* 28 (1995), Aug, Nr. 3, S. 389–97
- [78] MACHWE, A ; ORREN, D K. ; BOHR, V A.: Accelerated methylation of ribosomal RNA genes during the cellular senescence of Werner syndrome fibroblasts. In: *FASEB J* 14 (2000), Sep, Nr. 12, S. 1715–24
- [79] MALONE, Colin D. ; HANNON, Gregory J.: Small RNAs as guardians of the genome. In: *Cell* 136 (2009), Feb, Nr. 4, S. 656–68
- [80] MARINUS, Martin G. ; CASADESUS, Josep: Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. In: *FEMS Microbiol Rev* 33 (2009), May, Nr. 3, S. 488–503
- [81] MAYER, C. ; SCHMITZ, K. M. ; LI, J. ; GRUMMT, I. ; SANTORO, R.: Inter-genic transcripts regulate the epigenetic state of rRNA genes. In: *Mol. Cell* 22 (2006), May, S. 351–361
- [82] MAYER, Christine ; NEUBERT, Melanie ; GRUMMT, Ingrid: The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. In: *EMBO Rep* 9 (2008), Aug, Nr. 8, S. 774–80
- [83] MCCARTY, M ; AVERY, O T.: STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : III. AN IMPROVED METHOD FOR THE ISOLATION OF THE TRANSFORMING SUBSTANCE AND ITS APPLICATION TO

- PNEUMOCOCCUS TYPES II, III, AND VI. In: *J Exp Med* 83 (1946), Jan, Nr. 2, S. 97–104
- [84] MCFEE, A F. ; SAYER, A M. ; SALOMAA, S I. ; LINDHOLM, C ; LITTLEFIELD, L G.: Methods for improving the yield and quality of metaphase preparations for FISH probing of human lymphocyte chromosomes. In: *Environ Mol Mutagen* 29 (1997), Nr. 1, S. 98–104
- [85] MCSTAY, B. ; GRUMMT, I.: The epigenetics of rRNA genes: from molecular to chromosome biology. In: *Annu. Rev. Cell Dev. Biol.* 24 (2008), S. 131–157
- [86] MEDA, Francesca ; FOLCI, Marco ; BACCARELLI, Andrea ; SELMI, Carlo: The epigenetics of autoimmunity. In: *Cell Mol Immunol* 8 (2011), May, Nr. 3, S. 226–36
- [87] MEHLER, Mark F.: Epigenetics and the nervous system. In: *Ann Neurol* 64 (2008), Dec, Nr. 6, S. 602–17
- [88] MERZ, Katharina ; HONDELE, Maria ; GOETZE, Hannah ; GMELCH, Katharina ; STOECKL, Ulrike ; GRIESENBECK, Joachim: Actively transcribed rRNA genes in *S. cerevisiae* are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules. In: *Genes Dev* 22 (2008), May, Nr. 9, S. 1190–204
- [89] MICHAEL R. GREEN, Joseph S.: *Molecular Cloning: A Laboratory Manual (Fourth Edition): Three-Volume Set*. Cold Spring Harbor Laboratory, 2012
- [90] MICHALET, X ; EKONG, R ; FOUGEROUSSE, F ; ROUSSEAUX, S ; SCHURRA, C ; HORNIGOLD, N ; SLEGTENHORST, M van ; WOLFE, J ; POVEY, S ; BECKMANN, J S. ; BENSIMON, A: Dynamic molecular combing: stretching the whole human genome for high-resolution studies. In: *Science* 277 (1997), Sep, Nr. 5331, S. 1518–23
- [91] MICHAUD, E J. ; VUGT, M J. van ; BULTMAN, S J. ; SWEET, H O. ; DAVISON, M T. ; WOYCHIK, R P.: Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage. In: *Genes Dev* 8 (1994), Jun, Nr. 12, S. 1463–72
- [92] MIRKOVITCH, J ; MIRAULT, M E. ; LAEMMLI, U K.: Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. In: *Cell* 39 (1984), Nov, Nr. 1, S. 223–32
- [93] MOORE, Lisa D. ; LE, Thuc ; FAN, Guoping: DNA methylation and its basic function. In: *Neuropsychopharmacology* 38 (2013), Jan, Nr. 1, S. 23–38
- [94] NÉMETH, A. ; STROHNER, R. ; GRUMMT, I. ; LÄNGST, G.: The chromatin remodeling complex NoRC and TTF-I cooperate in the regulation of the mammalian rRNA genes in vivo. In: *Nucleic Acids Res.* 32 (2004), S. 4091–4099
- [95] NÉMETH, Attila ; CONESA, Ana ; SANTOYO-LOPEZ, Javier ; MEDINA, Ignacio ; MONTANER, David ; PÉTERFIA, Bálint ; SOLOVEI, Irina ; CREMER, Thomas ; DOPAZO, Joaquín ; LÄNGST, Gernot: Initial genomics of the human nucleolus. In: *PLoS Genet* 6 (2010), Mar, Nr. 3, S. e1000889

- 
- [96] NÉMETH, Attila ; GUIBERT, Sylvain ; TIWARI, Vijay K. ; OHLSSON, Rolf ; LÄNGST, Gernot: Epigenetic regulation of TTF-I-mediated promoter-terminator interactions of rRNA genes. In: *EMBO J* 27 (2008), Apr, Nr. 8, S. 1255–65
- [97] NÉMETH, Attila ; LÄNGST, Gernot: Genome organization in and around the nucleolus. In: *Trends Genet* 27 (2011), Apr, Nr. 4, S. 149–56
- [98] NEUBERGER, Michael S. ; HARRIS, Reuben S. ; DI NOIA, Javier ; PETERSEN-MAHRT, Svend K.: Immunity through DNA deamination. In: *Trends Biochem Sci* 28 (2003), Jun, Nr. 6, S. 305–12
- [99] OKANO, M ; BELL, D W. ; HABER, D A. ; LI, E: DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. In: *Cell* 99 (1999), Oct, Nr. 3, S. 247–57
- [100] OTTAVIANI, Diego ; LEVER, Elliott ; TAKOUSHIS, Petros ; SHEER, Denise: Anchoring the genome. In: *Genome Biol* 9 (2008), Nr. 1, S. 201
- [101] PARDOLL, D M. ; VOGELSTEIN, B: Sequence analysis of nuclear matrix associated DNA from rat liver. In: *Exp Cell Res* 128 (1980), Aug, Nr. 2, S. 466–70
- [102] PEARSON, Christopher E. ; NICHOL EDAMURA, Kerrie ; CLEARY, John D.: Repeat instability: mechanisms of dynamic mutations. In: *Nat Rev Genet* 6 (2005), Oct, Nr. 10, S. 729–42
- [103] PEDERSON, T: Half a century of "the nuclear matrix". In: *Mol Biol Cell* 11 (2000), Mar, Nr. 3, S. 799–805
- [104] PENG, Jamy C. ; KARPEN, Gary H.: Epigenetic regulation of heterochromatic DNA stability. In: *Curr Opin Genet Dev* 18 (2008), Apr, Nr. 2, S. 204–11
- [105] PHOKAEW, Chureerat ; KOWUDTITHAM, Supakit ; SUBBALEKHA, Keskanya ; SHUANGSHOTI, Shanop ; MUTIRANGURA, Apiwat: LINE-1 methylation patterns of different loci in normal and cancerous cells. In: *Nucleic Acids Res* 36 (2008), Oct, Nr. 17, S. 5704–12
- [106] PIETRZAK, Maciej ; REMPALA, Grzegorz ; NELSON, Peter T. ; ZHENG, Jing-Juan ; HETMAN, Michal: Epigenetic silencing of nucleolar rRNA genes in Alzheimer's disease. In: *PLoS One* 6 (2011), Nr. 7, S. e22585
- [107] PORTELA, Anna ; ESTELLER, Manel: Epigenetic modifications and human disease. In: *Nat Biotechnol* 28 (2010), Oct, Nr. 10, S. 1057–68
- [108] POTTER, D A. ; FOSTEL, J M. ; BERNINGER, M ; PARDUE, M L. ; CECHE, T R.: DNA-protein interactions in the *Drosophila melanogaster* mitochondrial genome as deduced from trimethylpsoralen crosslinking patterns. In: *Proc Natl Acad Sci U S A* 77 (1980), Jul, Nr. 7, S. 4118–22
- [109] PRADA, Diddier ; GONZÁLEZ, Rodrigo ; SÁNCHEZ, Lisandro ; CASTRO, Clementina ; FABIÁN, Eunice ; HERRERA, Luis A.: Satellite 2 demethylation induced by 5-azacytidine is associated with missegregation of chromosomes 1 and 16 in human somatic cells. In: *Mutat Res* 729 (2012), Jan, Nr. 1-2, S. 100–5

- [110] PROSSER, J ; FROMMER, M ; PAUL, C ; VINCENT, P C.: Sequence relationships of three human satellite DNAs. In: *J Mol Biol* 187 (1986), Jan, Nr. 2, S. 145–55
- [111] RAKYAN, Vardhman K. ; DOWN, Thomas A. ; BALDING, David J. ; BECK, Stephan: Epigenome-wide association studies for common human diseases. In: *Nat Rev Genet* 12 (2011), Aug, Nr. 8, S. 529–41
- [112] RASKA, I. ; KOBERNA, K. ; MALÍNSKÝ, J. ; FIDLEROVÁ, H. ; MASATA, M.: The nucleolus and transcription of ribosomal genes. In: *Biol. Cell* 96 (2004), Oct, S. 579–594
- [113] RATEL, David ; RAVANAT, Jean-Luc ; BERGER, François ; WION, Didier: N6-methyladenine: the other methylated base of DNA. In: *Bioessays* 28 (2006), Mar, Nr. 3, S. 309–15
- [114] REYES, J C. ; MUCHARDT, C ; YANIV, M: Components of the human SWI/SNF complex are enriched in active chromatin and are associated with the nuclear matrix. In: *J Cell Biol* 137 (1997), Apr, Nr. 2, S. 263–74
- [115] ROCHE: *Lab FAQs Find a Quick Solution 4th Edition*. 2011
- [116] ROMAN-GOMEZ, Jose ; JIMENEZ-VELASCO, Antonio ; AGIRRE, Xabier ; CASTILLEJO, Juan A. ; NAVARRO, German ; SAN JOSE-ENERIZ, Eurne ; GARATE, Leire ; CORDEU, Lucia ; CERVANTES, Francisco ; PROSPER, Felipe ; HEINIGER, Anabel ; TORRES, Antonio: Repetitive DNA hypomethylation in the advanced phase of chronic myeloid leukemia. In: *Leuk Res* 32 (2008), Mar, Nr. 3, S. 487–90
- [117] RUSSELL, J. ; ZOMERDIJK, J. C.: RNA-polymerase-I-directed rDNA transcription, life and works. In: *Trends Biochem. Sci.* 30 (2005), Feb, S. 87–96
- [118] SANTORO, R. ; GRUMMT, I.: Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. In: *Mol. Cell* 8 (2001), Sep, S. 719–725
- [119] SANTORO, R. ; GRUMMT, I.: Epigenetic mechanism of rRNA gene silencing: temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation. In: *Mol. Cell. Biol.* 25 (2005), Apr, S. 2539–2546
- [120] SANTORO, R. ; LI, J. ; GRUMMT, I.: The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. In: *Nat. Genet.* 32 (2002), Nov, S. 393–396
- [121] SCHMITZ, Kerstin-Maike ; MAYER, Christine ; POSTEPSKA, Anna ; GRUMMT, Ingrid: Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. In: *Genes Dev* 24 (2010), Oct, Nr. 20, S. 2264–9
- [122] SCHNAPP, G. ; SANTORI, F. ; CARLES, C. ; RIVA, M. ; GRUMMT, I.: The HMG box-containing nucleolar transcription factor UBF interacts with a specific subunit of RNA polymerase I. In: *EMBO J.* 13 (1994), Jan, S. 190–199

- 
- [123] SCHULZ, W A. ; STEINHOFF, C ; FLORL, A R.: Methylation of endogenous human retroelements in health and disease. In: *Curr Top Microbiol Immunol* 310 (2006), S. 211–50
- [124] SCHWANBECK, R ; MANFIOLETTI, G ; WIŚNIEWSKI, J R.: Architecture of high mobility group protein I-C.DNA complex and its perturbation upon phosphorylation by Cdc2 kinase. In: *J Biol Chem* 275 (2000), Jan, Nr. 3, S. 1793–801
- [125] SCHWARZACHER, H G. ; MOSGOELLER, W: Ribosome biogenesis in man: current views on nucleolar structures and function. In: *Cytogenet Cell Genet* 91 (2000), Nr. 1-4, S. 243–52
- [126] SGARRA, Riccardo ; LEE, Jaeho ; TESSARI, Michela A. ; ALTAMURA, Sandro ; SPOLAORE, Barbara ; GIANCOTTI, Vincenzo ; BEDFORD, Mark T. ; MANFIOLETTI, Guidalberto: The AT-hook of the chromatin architectural transcription factor high mobility group A1a is arginine-methylated by protein arginine methyltransferase 6. In: *J Biol Chem* 281 (2006), Feb, Nr. 7, S. 3764–72
- [127] SHARMA, Shikhar ; KELLY, Theresa K. ; JONES, Peter A.: Epigenetics in cancer. In: *Carcinogenesis* 31 (2010), Jan, Nr. 1, S. 27–36
- [128] SHI, Y B. ; HEARST, J E.: Wavelength dependence for the photoreactions of DNA-psoralen monoadducts. 1. Photoreversal of monoadducts. In: *Biochemistry* 26 (1987), Jun, Nr. 13, S. 3786–92
- [129] SIRRI, Valentina ; URCUQUI-INCHIMA, Silvio ; ROUSSEL, Pascal ; HERNANDEZ-VERDUN, Danièle: Nucleolus: the fascinating nuclear body. In: *Histochem Cell Biol* 129 (2008), Jan, Nr. 1, S. 13–31
- [130] SMIRNOV, E ; KALMÁROVÁ, M ; KOBERNA, K ; ZEMANOVÁ, Z ; MALÍNSKÝ, J ; MASATA, M ; CVACKOVÁ, Z ; MICHALOVÁ, K ; RASKA, I: NORs and their transcription competence during the cell cycle. In: *Folia Biol (Praha)* 52 (2006), Nr. 3, S. 59–70
- [131] SMITH, H C. ; ROTHBLUM, L I.: Ribosomal DNA sequences attached to the nuclear matrix. In: *Biochem Genet* 25 (1987), Dec, Nr. 11-12, S. 863–79
- [132] SOGO, J M. ; NESS, P J. ; WIDMER, R M. ; PARISH, R W. ; KOLLER, T: Psoralen-crosslinking of DNA as a probe for the structure of active nucleolar chromatin. In: *J Mol Biol* 178 (1984), Oct, Nr. 4, S. 897–919
- [133] SOLOVEI, Irina ; CAVALLO, Antonio ; SCHERMELLEH, Lothar ; JAUNIN, Françoise ; SCASSELATI, Catia ; CMARKO, Dusan ; CREMER, Christoph ; FAKAN, Stanislav ; CREMER, Thomas: Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). In: *Exp Cell Res* 276 (2002), May, Nr. 1, S. 10–23
- [134] STEPHANOVA, E ; STANCHEVA, R ; AVRAMOVA, Z: Binding of sequences from the 5'- and 3'-nontranscribed spacers of the rat rDNA locus to the nucleolar matrix. In: *Chromosoma* 102 (1993), Mar, Nr. 4, S. 287–95

- [135] STRATAGENE: *Stratalinker UV Crosslinker*. Revision 122003 and IN 70034-06. La Jolla, CA, USA: Agilent technologies (Veranst.), 2002. – Instruction manual
- [136] STROHNER, R. ; NEMETH, A. ; JANSÁ, P. ; HOFMANN-ROHRER, U. ; SANTORO, R. ; LÄNGST, G. ; GRUMMT, I.: NoRC—a novel member of mammalian ISWI-containing chromatin remodeling machines. In: *EMBO J.* 20 (2001), Sep, S. 4892–4900
- [137] STROHNER, R. ; NÉMETH, A. ; NIGHTINGALE, K. P. ; GRUMMT, I. ; BECKER, P. B. ; LÄNGST, G.: Recruitment of the nucleolar remodeling complex NoRC establishes ribosomal DNA silencing in chromatin. In: *Mol. Cell. Biol.* 24 (2004), Feb, S. 1791–1798
- [138] STULTS, D. M. ; KILLEN, M. W. ; PIERCE, H. H. ; PIERCE, A. J.: Genomic architecture and inheritance of human ribosomal RNA gene clusters. In: *Genome Res.* 18 (2008), Jan, S. 13–18
- [139] SUZUKI, Toshikazu ; FUJII, Michihiko ; AYUSAWA, Dai: Demethylation of classical satellite 2 and 3 DNA with chromosomal instability in senescent human fibroblasts. In: *Exp Gerontol* 37 (2002), Nr. 8-9, S. 1005–14
- [140] TAHILIANI, Mamta ; KOH, Kian P. ; SHEN, Yinghua ; PASTOR, William A. ; BANDUKWALA, Hozefa ; BRUDNO, Yevgeny ; AGARWAL, Suneet ; IYER, Lakshminarayan M. ; LIU, David R. ; ARAVIND, L ; RAO, Anjana: Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. In: *Science* 324 (2009), May, Nr. 5929, S. 930–5
- [141] TILMAN, Gaëlle ; ARNOULT, Nausica ; LENGLEZ, Sandrine ; VAN BENEDEN, Amandine ; LORIOT, Axelle ; DE SMET, Charles ; DECOTTIGNIES, Anabelle: Cancer-linked satellite 2 DNA hypomethylation does not regulate Sat2 non-coding RNA expression and is initiated by heat shock pathway activation. In: *Epigenetics* 7 (2012), Aug, Nr. 8, S. 903–13
- [142] UKAI, Hideki ; ISHII-OBA, Hiroko ; UKAI-TADENUMA, Maki ; OGIU, Toshiaki ; TSUJI, Hideo: Formation of an active form of the interleukin-2/15 receptor beta-chain by insertion of the intracisternal A particle in a radiation-induced mouse thymic lymphoma and its role in tumorigenesis. In: *Mol Carcinog* 37 (2003), Jun, Nr. 2, S. 110–9
- [143] VALINLUCK, Victoria ; TSAI, Hsin-Hao ; ROGSTAD, Daniel K. ; BURDZY, Artur ; BIRD, Adrian ; SOWERS, Lawrence C.: Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). In: *Nucleic Acids Res* 32 (2004), Nr. 14, S. 4100–8
- [144] VENGROVA, Sonya ; DALGAARD, Jacob Z.: The wild-type *Schizosaccharomyces pombe* mat1 imprint consists of two ribonucleotides. In: *EMBO Rep* 7 (2006), Jan, Nr. 1, S. 59–65
- [145] VISCHER, E ; CHARGAFF, E: The separation and quantitative estimation of purines and pyrimidines in minute amounts. In: *J Biol Chem* 176 (1948), Nov, Nr. 2, S. 703–14

- 
- [146] WALSH, C P. ; CHAILLET, J R. ; BESTOR, T H.: Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. In: *Nat Genet* 20 (1998), Oct, Nr. 2, S. 116–7
- [147] WANG, Lianrong ; CHEN, Shi ; VERGIN, Kevin L. ; GIOVANNONI, Stephen J. ; CHAN, Simon W. ; DEMOTT, Michael S. ; TAGHIZADEH, Koli ; CORDERO, Otto X. ; CUTLER, Michael ; TIMBERLAKE, Sonia ; ALM, Eric J. ; POLZ, Martin F. ; PINHASSI, Jarone ; DENG, Zixin ; DEDON, Peter C.: DNA phosphorothioation is widespread and quantized in bacterial genomes. In: *Proc Natl Acad Sci U S A* 108 (2011), Feb, Nr. 7, S. 2963–8
- [148] WANG, Lianrong ; CHEN, Shi ; XU, Tiegang ; TAGHIZADEH, Koli ; WISHNOK, John S. ; ZHOU, Xiufen ; YOU, Delin ; DENG, Zixin ; DEDON, Peter C.: Phosphorothioation of DNA in bacteria by dnd genes. In: *Nat Chem Biol* 3 (2007), Nov, Nr. 11, S. 709–10
- [149] WEBER, Michael ; DAVIES, Jonathan J. ; WITTIG, David ; OAKELEY, Edward J. ; HAASE, Michael ; LAM, Wan L. ; SCHÜBELER, Dirk: Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. In: *Nat Genet* 37 (2005), Aug, Nr. 8, S. 853–62
- [150] WEISENBERGER, Daniel J. ; CAMPAN, Mihaela ; LONG, Tiffany I. ; KIM, Myungjin ; WOODS, Christian ; FIALA, Emerich ; EHRLICH, Melanie ; LAIRD, Peter W.: Analysis of repetitive element DNA methylation by MethyLight. In: *Nucleic Acids Res* 33 (2005), Nr. 21, S. 6823–36
- [151] WELLINGER, R E. ; LUCCHINI, R ; DAMMANN, R ; SOGO, J M.: In vivo mapping of nucleosomes using psoralen-DNA crosslinking and primer extension. In: *Methods Mol Biol* 119 (1999), S. 161–73
- [152] WELLINGER, R E. ; SOGO, J M.: In vivo mapping of nucleosomes using psoralen-DNA crosslinking and primer extension. In: *Nucleic Acids Res* 26 (1998), Mar, Nr. 6, S. 1544–5
- [153] WIJMEGA, C ; HANSEN, R S. ; GIMELLI, G ; BJÖRCK, E J. ; DAVIES, E G. ; VALENTINE, D ; BELOHRADSKY, B H. ; DONGEN, J J. van ; SMEETS, D F. ; HEUVEL, L P. van den ; LUYTEN, J A. ; STRENGMAN, E ; WEEMAES, C ; PEARSON, P L.: Genetic variation in ICF syndrome: evidence for genetic heterogeneity. In: *Hum Mutat* 16 (2000), Dec, Nr. 6, S. 509–17
- [154] WILSON, Ann S. ; POWER, Barbara E. ; MOLLOY, Peter L.: DNA hypomethylation and human diseases. In: *Biochim Biophys Acta* 1775 (2007), Jan, Nr. 1, S. 138–62
- [155] WITTNER, Manuel ; HAMPERL, Stephan ; STÖCKL, Ulrike ; SEUFERT, Wolfgang ; TSCHOCHNER, Herbert ; MILKEREIT, Philipp ; GRIESENBECK, Joachim: Establishment and maintenance of alternative chromatin states at a multicopy gene locus. In: *Cell* 145 (2011), May, Nr. 4, S. 543–54
- [156] WU, Hao ; ZHANG, Yi: Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. In: *Genes Dev* 25 (2011), Dec, Nr. 23, S. 2436–52



- [157] WU, M ; RINCHIK, E M. ; WILKINSON, E ; JOHNSON, D K.: Inherited somatic mosaicism caused by an intracisternal A particle insertion in the mouse tyrosinase gene. In: *Proc Natl Acad Sci U S A* 94 (1997), Feb, Nr. 3, S. 890–4
- [158] WYATT, G R.: Occurrence of 5-methylcytosine in nucleic acids. In: *Nature* 166 (1950), Aug, Nr. 4214, S. 237–8
- [159] XU, G L. ; BESTOR, T H. ; BOURC'HIS, D ; HSIEH, C L. ; TOMMERUP, N ; BUGGE, M ; HULTEN, M ; QU, X ; RUSSO, J J. ; VIEGAS-PÉQUIGNOT, E: Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. In: *Nature* 402 (1999), Nov, Nr. 6758, S. 187–91
- [160] YAN, P S. ; RODRIGUEZ, F J. ; LAUX, D E. ; PERRY, M R. ; STANDIFORD, S B. ; HUANG, T H.: Hypermethylation of ribosomal DNA in human breast carcinoma. In: *Br J Cancer* 82 (2000), Feb, Nr. 3, S. 514–7
- [161] YANG, Allen S. ; ESTÉCIO, Marcos R H. ; DOSHI, Ketan ; KONDO, Yutaka ; TAJARA, Eloiza H. ; ISSA, Jean-Pierre J.: A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. In: *Nucleic Acids Res* 32 (2004), Nr. 3, S. e38
- [162] YU, C E. ; OSHIMA, J ; FU, Y H. ; WIJSMAN, E M. ; HISAMA, F ; ALISCH, R ; MATTHEWS, S ; NAKURA, J ; MIKI, T ; OUAIS, S ; MARTIN, G M. ; MULLIGAN, J ; SCHELLENBERG, G D.: Positional cloning of the Werner's syndrome gene. In: *Science* 272 (1996), Apr, Nr. 5259, S. 258–62
- [163] YUSUFZAI, Timur M. ; TAGAMI, Hideaki ; NAKATANI, Yoshihiro ; FELSENFELD, Gary: CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. In: *Mol Cell* 13 (2004), Jan, Nr. 2, S. 291–8
- [164] ZHEN, W P. ; JEPPESEN, C ; NIELSEN, P E.: Psoralen photofootprinting of protein-binding sites on DNA. In: *FEBS Lett* 229 (1988), Feb, Nr. 1, S. 73–6
- [165] ZHOU, Yonggang ; SCHMITZ, Kerstin-Maike ; MAYER, Christine ; YUAN, Xuejun ; AKHTAR, Asifa ; GRUMMT, Ingrid: Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. In: *Nat Cell Biol* 11 (2009), Aug, Nr. 8, S. 1010–6
- [166] ZILLNER, Karina ; FILARSKY, Michael ; RACHOW, Katrin ; WEINBERGER, Michael ; LÄNGST, Gernot ; NÉMETH, Attila: *Large-scale organization of ribosomal DNA chromatin is regulated by Tip5*. Jan 2013. – unpublished; under revision at Nucleic Acid Research
- [167] ZILLNER, Karina ; JERABEK-WILLEMSSEN, Moran ; DUHR, Stefan ; BRAUN, Dieter ; LÄNGST, Gernot ; BAASKE, Philipp: Microscale thermophoresis as a sensitive method to quantify protein: nucleic acid interactions in solution. In: *Methods Mol Biol* 815 (2012), S. 241–52
- [168] ZILLNER, Karina ; NÉMETH, Attila: Single-molecule, genome-scale analyses of DNA modifications: exposing the epigenome with next-generation technologies. In: *Epigenomics* 4 (2012), Aug, Nr. 4, S. 403–14

- 
- [169] ZILLNER, Karina ; NÉMETH, Attila: *Psoralen-combing for Sequence-specific Single Molecule Chromatin Analysis*. Jan 2013. – unpublished, under revision at Methods in Molecular Biology
- [170] ZOU, Yan ; WEBB, Kristofor ; PERNA, Avi D. ; ZHANG, Qingchun ; CLARKE, Steven ; WANG, Yinsheng: A mass spectrometric study on the in vitro methylation of HMGA1a and HMGA1b proteins by PRMTs: methylation specificity, the effect of binding to AT-rich duplex DNA, and the effect of C-terminal phosphorylation. In: *Biochemistry* 46 (2007), Jul, Nr. 26, S. 7896–906