Exploring the role of structured DNA elements and their specific binders in genome regulation

Dissertation zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)
der naturwissenschaftlichen Fakultät III - Biologie und vorklinische Medizin -
der Universität Regensburg

vorgelegt von
Michael Filarsky
aus Augsburg
Mai 2013

Das Promotionsgesuch wurde eingereicht am:

9. April 2013

Die Arbeit wurde angeleitet von:

Prof. Dr. Gernot Längst

(Michael Filarsky)

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ATP	Adenosine-5'-triphosphate
bp	base pair
BSA	Bovine Serum Albumine
°C	degree Celsius
C-terminal	carboxy terminal
CpG	Cytosine-phosphatidyl-guanosine
CSK	cytoskeleton
DHFR	Dihydrofolate reductase
DNA	Deoxyribonucleic Acid
DNMT	DNA-Cytosine-5-Methyltransferase
dNTP	2'-deoxynucleotide triphosphate
E. coli	Escherichia coli
EC_{50}	Equilibrium constant
EDTA	Ethyleneiaminotetraacetate
EGTA	Ethylene Glycol Tetraacetic Acid
ENCODE	Encyclopedia of DNA Elements
EtBr	Ethidium Bromide
FCS	Fetal Calf Serum
GαM	Goat anti mouse
GαRb	Goat anti rabbit
GPBP1	GC-rich Promoter Binding Protein 1
h	hour
H-DNA	Hinged DNA
H3Kxme3	tri-methylation at lysine x of histone H3
H4Kxac	acetylation at lysine x of histone H4
HDAC	Histone Acetyltransferase
HMT	Histone Methyltransferase
HP1	Heterochromatin binding Protein 1
IGS	Intergenic spacer
IPTG	Isopropylthiogalactoside
kb	kilobases

K_d	dissociation constant
kDa	Kilodalton
LB	Luria-Bertani
lncRNA	long ncRNA
M	molar
MAR	Matrix Attachment Regions
MBP	Methyl-CpG-binding Proteins
min	minute
min	minute
mM	millimolar
MST	Microscale Thermophoresis
N-terminal	amino terminal
ncRNA	non-coding RNA
nM	nanomolar
NM	nuclear matrix
NOR	Nucleolar Organizer Region
NoRC	Nucleolar Remodeling Complex
PAA	Polyacrylamide
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PHD	Plant Homeodomain
pRNA	promoter RNA
PTOV1	Prostate Tumor Overexpressed 1
qPCR	quantitative Polymerase Chain Reaction
rDNA	ribosomal DNA
RNA	Ribonucleic Acid
RNAPolI	RNA Polymerase I
rpm	rounds per minute
rRNA	ribosomal RNA
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
sec	second
SNF2H	Snf2 homolog protein

snoRNA	small nucleolar RNA
TAM	Tip5/ARBP/MBD
Taq	Thermus aquaticus
TBE	Tris Borate EDTA
TE	Tris EDTA
TFO	Triplex Forming Oligonucleotide
TIP5	TTF-I interacting protein 5
Tris	tris(hydroxymethyl)aminomethane
TTF1	Transcription Termination Factor 1
TTS	Triplex Target Site
UBF	Upstream Binding Factor
V	Volt

1 Summary

After the description of the double helix structure by Watson and Crick (Watson and Crick, 1953), the function of DNA in the cell was seen as rather static, a long stretched out database of information, which was read and interpreted by the protein machinery doing all the interesting work. That notion however, quickly changed with the discovery of the DNA's potential to adopt various alternative conformations, like Z-DNA, quadruplex DNA or triplexes (Felsenfeld and Rich, 1957; Panayotatos and Wells, 1981; Sen and Gilbert, 1988; Wang et al., 1979). Although initially characterized *in vitro*, these structures were soon implicated in the development of certain diseases and suggested to have possible roles in gene regulation and genome organization (Agazie et al., 1996; Bacolla et al., 2004; Sakamoto et al., 1999; Schroth and Ho, 1995; Wells, 2008). The major focus thereby lay on the triplexes, three stranded structures comprising the DNA duplex and a single stranded RNA or DNA that binds in the major groove of the double helix. Strikingly, the recent advancements in high-throughput sequencing technology revealed the enrichment of putative triplex forming sequences within the regulatory regions of the genome (Goñi et al., 2004). These observations combined with the discovery of an ever-growing number of functional ncRNAs, led to the assumption that triplex structures could present an interface between these ncRNAs and the chromatin (Rinn and Chang, 2012).

The data in this thesis reveal that the intergenic spacer region of the mammalian rRNA genes contains a cluster of possible triplex forming sequences in its regulatory enhancer and terminator elements. The triplex forming potential of these motifs was predicted by *in-silico* analysis of the rDNA sequence and validated using EMSA experiments. Microscale thermophoresis (MST) assays were established to investigate the formation of triplex structures in various conditions *in vitro*. These experiments presented new insights on the characteristics of the triplex formation process and suggest the complex formation to be readily possible in physiological conditions. Thus the results provide the basis for further

investigations on a possible role for enhancer-derived triplex forming ncRNAs in the regulation of the rRNA genes.

Besides their presence in regulatory gene regions, sequences prone to form non-canonical DNA structures were also shown to be part of matrix attachments regions (MARs) throughout the genome (Glazko et al., 2003). These MAR sites are tethered to the nuclear matrix and mediate the anchoring of chromatin loops on this putative protein scaffold within the nucleus. The MAR elements are implicated to play various roles in transcriptional regulation and chromatin organization (Ottaviani et al., 2008). Interestingly, MARs have also been shown to interact with specific MAR binding proteins and were implicated to form triplex structures with ncRNAs at the nuclear matrix (Ottaviani et al., 2008; Zheng et al., 2010). However, the exact nature of these processes and their implications for gene regulation and chromatin organization are so far poorly understood.

In the second part of this thesis Tip5, a subunit of NoRC, the key player in epigenetic rRNA gene repression, is revealed as nuclear matrix binding protein. The AT-hook and TAM domains of the protein are identified as motifs mediating the localization of Tip5 to the nuclear matrix. The results furthermore show a Tip5 dependent recruitment of rDNA repeats to the nuclear matrix upon rDNA repression. This suggests an additional role for the NoRC complex in the large-scale chromatin organization of the rRNA genes to the nuclear matrix. Thus the insights provided by these results are a good example for the interplay between MAR elements and their specific binders in large-scale chromatin organization.

In the course of these experiments the new extended AT-hook motif was discovered, which displays a DNA binding activity similar to canonical AT-hooks and moreover could be revealed as strong RNA binder. Initial *in vivo* experiments validated the importance of the eAT-hook domain for proper cellular localization of a candidate protein. Together these results suggest a role for the eAT-hook domain in chromatin organization, nuclear matrix binding and anchoring of proteins to structural DNA and RNA elements, similar to the canonical AT-hooks.

In summary the data presented in this thesis provides new insights on the characteristics and various roles for structured DNA elements and their specific binders in transcriptional gene regulation and nuclear matrix mediated chromatin organization.

2 Introduction

2.1 Structured DNA elements – Beyond the double helix

After the description of the double helix structure by Watson and Crick (Watson and Crick, 1953), the function of DNA in the cell was seen as rather static, a long stretched out database of information, which was read and interpreted by the protein machinery doing all the interesting work. That notion however, quickly changed with the discovery of the DNA's potential to adopt various alternative conformations. The first was the description of DNA triplex structures by Felsenfeld and colleagues in 1957 (Felsenfeld and Rich, 1957), which was followed over time by the discovery of the Z-DNA, the cruciform DNA and the quadruplex DNA structures (Panayotatos and Wells, 1981; Sen and Gilbert, 1988; Wang et al., 1979).

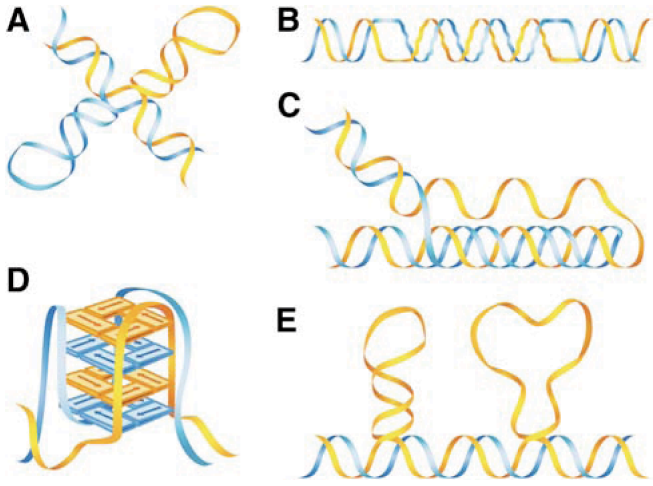


Figure 2.1: Examples of non-canonical DNA structures
A) cruciform DNA **B**) Z-DNA **C**) H-DNA (triplex) **D**) quadruplex DNA **E**) slipped DNA

All in all roughly ten different non-canonical DNA structures have been described up to date (Zhao et al., 2010). Although all of them were initially characterized *in vitro*, a growing body of evidence suggests their formation under physiological conditions and most of them are hypothesized to have some functional roles *in vivo*

(Bochman et al., 2012). Furthermore, the presence of non-canonical DNA motifs in the genome has been shown to affect transcription, replication, and to be able to create hotspots of genome instability (Bacolla et al., 2004; Bacolla and Wells, 2004; Wang and Vasquez, 2006). For all these structures to occur, certain sequence elements in the underlying DNA are a necessity. Surprisingly, sequencing analysis of the mouse and humane genomes have revealed a widespread distribution of such structured DNA elements and illustrated their enrichment in regulatory regions like promoters or enhancers (Goñi et al., 2004; Huppert and Balasubramanian, 2007; Schroth and Ho, 1995; Todd et al., 2005; Wang and Vasquez, 2006). All together these observations suggest multiple roles for non-canonical DNA structures in gene regulation and genome organization.

For this work, possible functions in gene regulation and large-scale chromatin organization for triplex DNA structures and the association of structured DNA elements with their specific binders to the nuclear matrix were explored. The aim of this project was thereby, to gain a better insight on the functional roles of these structured DNA elements in the regulation of our genome.

2.2 Triplex DNA structures

2.2.1 The concept of triplex DNA

The concept of DNA triple helices or triplex DNA has come a long way. Almost 60 years ago Felsenfeld and colleagues described the potential of double stranded nucleic acids to form triple helical complexes with a third single stranded molecule (Felsenfeld and Rich, 1957). Triplex formation has been shown to be possible on long homopolypurine or homopolypyrimidine stretches in the underlying DNA sequence. The third strand is thereby located within the major groove of the double helix, binding to the purine rich DNA strand in either a parallel or an anti-parallel orientation, via so called Hoogsteen hydrogen bonds (Hoogsteen, 1959). This results in two different triplex motifs, named after the nature of the third strand. With a pyrimidine rich sequence the annealing of the third strand occurs in

a parallel orientation via the Hoogsteen hydrogen bonds and is called the pyrimidine motif. Whereas a polypurine strand binds to the double helix in an anti-parallel fashion through reverse-Hoogsteen hydrogen bonds, creating the purine motif (Beal and Dervan, 1991; Cooney et al., 1988; Letai et al., 1988; Morgan and Wells, 1968).

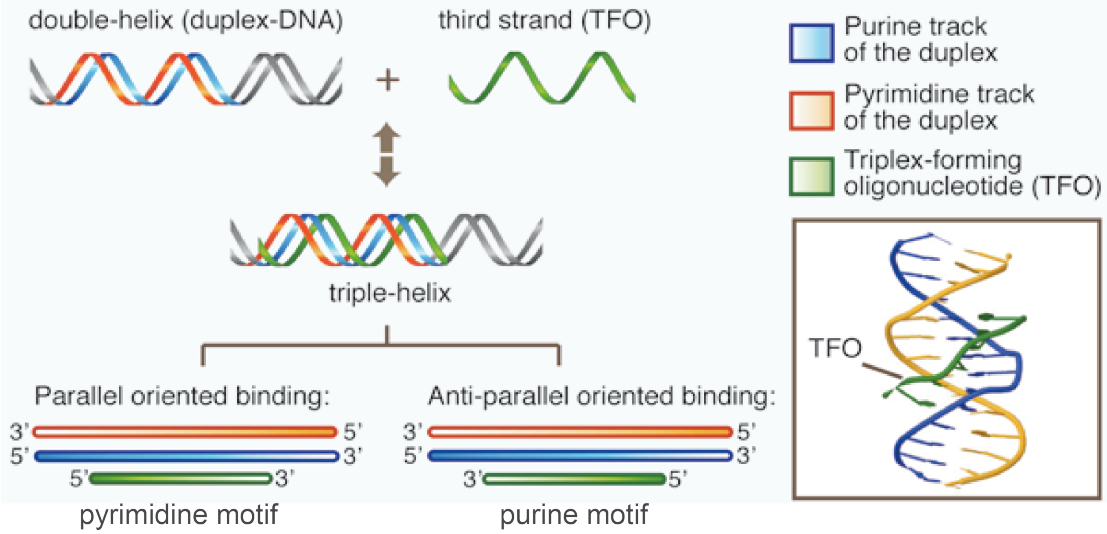


Figure 2.2: The concept of triplex formation
A third strand binds in the major groove of the DNA double helix via Hoogsteen or reverse-Hoogsteen hydrogen bonds. This results in the parallel pyrimidine and the anti-parallel purine triplex, respectively. The inlet on the right displays a structural model of the resulting triplex. (modified from (Buske et al., 2011))

The purine motif triplex readily assembles in neutral pH and results in the triplets G:G-C, A:A-T and T:A-T. The parallel annealing of the third strand in the pyrimidine motif produces the triplets C:G-C and T:A-T. Here the cytosine in the C:G-C triplet needs to be protonated at the N3 position in order to achieve a stable interaction therefore the binding is favoured at acidic pH (Lee et al., 1979). In general two different types of triple helices are possible, the intermolecular triplex, consisting of a duplex and a separate third strand and the intramolecular triplex, where all three strands are part of the same molecule. These so-called H-DNA structures are supposed to form at mirror repeat regions within the genome and have been implicated in gene regulatory processes and the development of certain diseases (reviewed in (Jain et al., 2008; Wells, 2008).

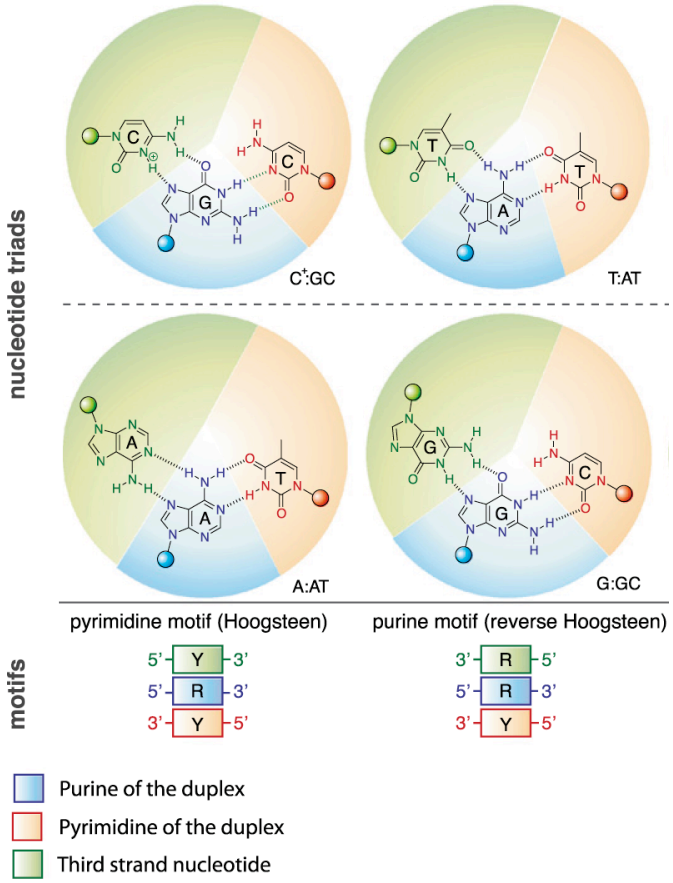


Figure 2.3: Molecular mechanism of triplex formation
The figure illustrates the nucleotide triads, which are formed by the Hoogsteen and reverse-Hoogsteen hydrogen bonds in the pyrimidine (left) and purine (right) motif, respectively. (modified from (Buske et al., 2012))

2.2.2 Encoding the triplex – Factors determining triplex formation and stability

The biochemical and biophysical factors, which determine the formation and stability of DNA triplexes, have been studied extensively by a variety of different methods, ranging from rather simple techniques like EMSA and footprinting assays (Cheng and Van Dyke, 1994; Hampel et al., 1991; Roberts and Crothers, 1991), to more complex, state of the art methods like CD spectroscopy (Ellouze et al., 1997) and SPR (Alberti et al., 2002; Sugimoto et al., 2001).

This extensive *in vitro* characterization showed, that triplex formation is a rather slow process, which follows a so called nucleation-zipper mechanism starting at the 5'-end of the duplex DNA (Alberti et al., 2002; Paes and Fox, 1997). Once formed the triplex is a stable compound exhibiting half-lives of days (James et al.,

2003; Vasquez et al., 1995), with the combination of a DNA double strand and a single stranded RNA molecule representing the most stable triplex (Escudé et al., 1993; Roberts and Crothers, 1992). Besides the length and base composition, the pH and the concentration of available cations have been illustrated as the major factors to determine triplex formation and stability (Kiessling et al., 1992; Moser and Dervan, 1987).

As mentioned earlier, the pyrimidine motif triplex relies on a lower pH, which limits the stability of this motif in physiological conditions (Sugimoto et al., 2001). The positive charge added by the protonation at lower pH has the additional advantage of compensating for the repulsion induced by the polyanionic oligonucleotide backbones. On the other hand continuous runs of protonated cytosines result in a unfavourable charge repulsion between them (Völker and Klump, 1994). While the purine motif forms readily at physiological pH, it has been shown, that the triplex formation at long guanine stretches competes with quadruplex formation, thereby limiting its effectiveness under such conditions (Floris et al., 1999; Olivas and Maher, 1995a).

The greatest obstacle for triplex formation is the electrostatic repulsion, caused by the three negatively charged oligonucleotide backbones. The major factor in overcoming this problem is the interaction with different cation species. The influence on triplex stability was shown for monovalent (Chen and Chen, 2011; Durand et al., 1992; James et al., 2003; Pilch et al., 1990; Rougée et al., 1992), divalent (Chen and Chen, 2011; Ellouze et al., 1997; James et al., 2003; Wu et al., 2002) and also polyvalent cations, such as spermine (Hampel et al., 1991; Singleton and Dervan, 1993). Interestingly, in monocationic solutions increasing cation concentrations stabilize triplex formation, with T-rich triplexes being much more salt depended than complexes comprising a higher GC content (James et al., 2003; Volker et al., 1993). Thereby the effect of divalent magnesium cations has been shown to be much larger than the one observed for the monovalent sodium (Durand et al., 1992; Manning, 1978; Sugimoto et al., 2001). In mixed valence salt solutions, divalent and monovalent cations display anti-cooperative effects, with high sodium or potassium concentrations reducing triplex stability by destabilizing the C:G-C triplets, which speeds up the complex dissociation rate (Chen and Chen, 2011; James et al., 2003; Maher et al., 1990). This effect can be rescued by an

increase in magnesium concentration within the solution (Maher et al., 1990; Singleton and Dervan, 1993). Finally in buffers containing approximately physiological concentrations of sodium, magnesium and potassium ions, the addition of the tetravalent polyamines, spermine or spermidine, dramatically increases triplex stability at physiological pH (Hampel et al., 1991; Singleton and Dervan, 1993). This is noteworthy, since the polyamine concentration in the eukaryotic nucleus is in the millimolar range (Hampel et al., 1991), suggesting the formation of purine and pyrimidine triplexes at the appropriate sequences is readily possible *in vivo*.

2.2.3 The concept of the TFO antigene strategy

The full technical potential of the discovered triplex formation was unravelled by a series of publications, revealing the possibility to target specific DNA sequences with a short triplex forming oligonucleotide (TFO). The studies used these TFOs for the induction of site-specific cleavage events and revealed the possibility to inhibit transcription by the targeted formation of a triplex at the gene promoter (Cooney et al., 1988; Le Doan et al., 1987; Moser and Dervan, 1987).

Based on these observations the concept of the TFO antigene strategy was established, which aims at using their enormous potential as sequence specific high affinity binders for biochemical, biotechnological and therapeutic applications. Therefore extensive studies were undertaken, using the TFO technology for the inhibition of transcription (Besch et al., 2002; Carbone et al., 2004; Catapano et al., 2000; Postel et al., 1991), site directed recombination (Chin et al., 2007; Faruqi et al., 2000; Sandor and Bredberg, 1995), the targeted modification of genomic DNA and the introduction of specific mutations (Barre et al., 2000; Vasquez et al., 2001; Vasquez et al., 2000). These projects shed further light on the prerequisites necessary for triplex formation in physiological conditions and moreover clearly showed that the induced formation of these structures is possible in a chromatin context inside the nucleus, thereby making a point for their putative existence and function *in vivo*.

2.2.4 Evidence for the existence of triplex structures *in vivo*

Apart from all advancements made in the area of triplex research by using sophisticated *in vitro* methods and the TFO antigene strategy, to proof the natural existence of these structures *in vivo* and to gain conclusive insights on their biological functions still presents the biggest challenge in the field.

The existence of triple helical DNA in the nuclei of insects, nematodes and mammals has been suggested by immunostainings with triplex specific monoclonal antibodies (Agazie et al., 1994; Burkholder et al., 1988; Burkholder et al., 1991; Gorab et al., 2009; Ohno et al., 2002; Piergentili and Mencarelli, 2008). The results displayed a granular staining pattern throughout the nucleus, with abundant triplex signals. Since the initial antibodies were prepared against DNA/DNA triplex structures (DNA double helix/DNA third strand) the reports suggested that the observed signals could present H-DNA structures and the formation of intermolecular triplexes between adjacent chromatin loops, thereby playing a role in mediating genome organization at the nuclear matrix (Agazie et al., 1996; Lee et al., 1987; Ohno et al., 2002; Stollar and Raso, 1974).

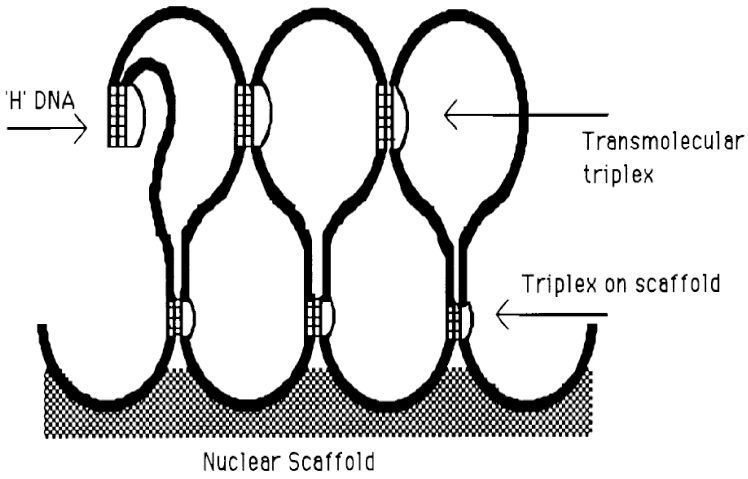


Figure 2.4: Model for triplex mediated genome organization
The formation of intermolecular triplexes and H-DNA structures mediates the organization of distinct chromatin loops at a nuclear scaffold or nuclear matrix. (model as suggested by Agazie and colleagues (Agazie et al., 1996))

In subsequent studies employing the same antibodies, it was revealed that they posses an even higher affinity to DNA/RNA triplex structures (DNA double

helix/RNA third strand), implicating that a large proportion of the detected signals may indeed reflect DNA/RNA triple helices (Bagasra et al., 2006; Kanak et al., 2010).

Further support for the existence of triplex structures in cells was put forward by the discovery of endogenous proteins, which specifically recognize and bind to triplexes (reviewed in (Buske et al., 2011)). Within the identified proteins one can find chromatin-associated proteins, such as high mobility box (HMG) proteins, as well as nuclear matrix binding proteins (Jain et al., 2005; Jimenez-Garcia et al., 1998; Li et al., 2002; Suda et al., 1996). However, it has to be mentioned that most of the observed interactions were established *in vitro* and may therefore in part only reflect the preference of some cationic proteins for the higher negative charge density of the multistranded triplexes. In summary, although the results provided substantial evidence for a possible existence of triplex structures *in vivo*, the ultimate proof is still missing.

2.2.5 Unraveling the abundant triplex forming potential of mammalian genomes

The overrepresentation of long oligopurine/oligopyrimidine sequences within eukaryotic gene regions and their potential as possible triplex forming sites has already been implicated around twenty years ago (Behe, 1995; Wells et al., 1988). However only the recent advancements in high throughput sequencing technologies in combination with bioinformatic analysis methods revealed the full triplex forming potential of mammalian genomes. The results not only unraveled a high abundance of putative triplex target sites (TTSs), but also illustrated a clear enrichment of those TTS sequences within the regulatory and repetitive regions of the genome (Buske et al., 2012; Goñi et al., 2004; Goñi et al., 2006; Wu et al., 2007). Strikingly, the highest concentration of TTSs is located upstream of transcription start sites and in promoter elements, with roughly 98% and 95% of all mouse and human promoters containing at least one possible TTS, respectively (Goñi et al., 2004; Goñi et al., 2006; Wu et al., 2007).

Of special interest for possible functional roles in gene regulation are those promoter polypurine/polypyrimidine tracts displaying characteristics of mirror repeats, which therefore may potentially participate in intra- and intermolecular triplex formation (Clark et al., 2006; Jain et al., 2008; Schroth and Ho, 1995). Both processes are depicted and described in (Figure 2.5). The transcribed RNA directly binds to the major groove of the mirror repeat and thereby forms an intermolecular triplex, which in turn exhibits some regulatory functions on the adjacent gene. In contrast for the formation of an intramolecular H-DNA structure, one DNA strand needs to fold back onto the duplex leaving an unpaired single strand behind. The presence of such H-DNA triplex structures in promoter regions has already been implicated to play a role in the development of certain diseases like the Friedreich ataxia syndrome (Sakamoto et al., 1999; Soragni et al., 2008; Wells, 2008).

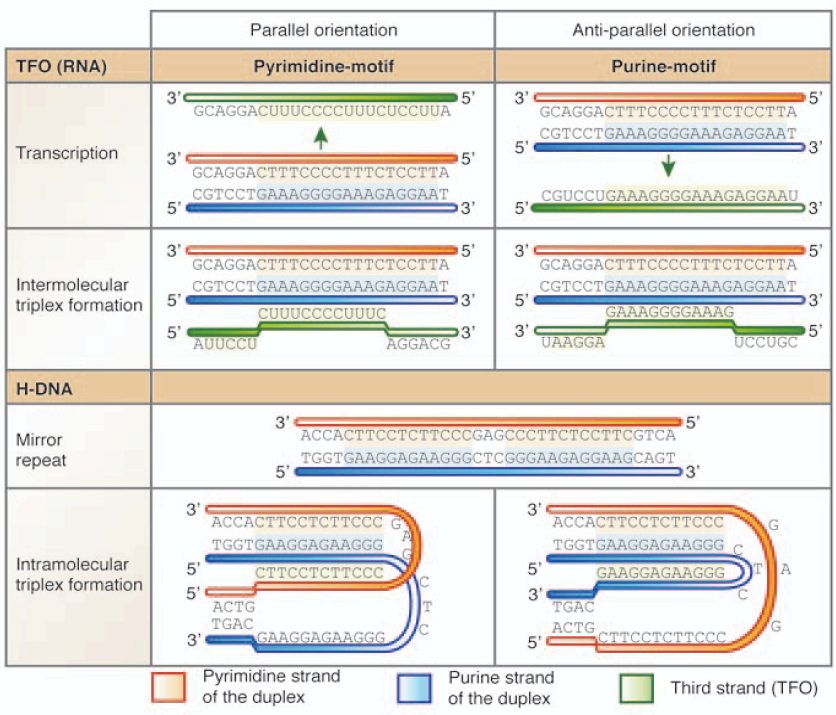


Figure 2.5: Triplex formation at mirror repeat sequences
Two different triplex structures can form on mirror repeat sequences in the genome. The upper panel illustrates the formation of intermolecular triplexes consisting of the DNA and a RNA third strand derived from the mirror repeat transcript. The lower panel shows the formation of an intramolecular H-DNA structure, where all three strands of the triplex are derived from the same duplex DNA. (from (Buske et al., 2012))

2.2.6 Possible functions of triplex structures *in vivo*

The data presented so far clearly suggest the actual existence of triplexes *in vivo*. This immediately leads to the question of their functional roles and regulation within the cell.

The intramolecular triplexes formed in RNA molecules represent the best-characterized example for the *in vivo* function of triplexes, so far. These RNA triplexes have been shown to contribute to folding and tertiary structure stability (Holland and Hoffman, 1996; Mitton-Fry et al., 2010; Nelson et al., 2012) and are furthermore crucial for proper enzymatic function by mediating catalytic activity and have been shown to work as translational enhancers for certain RNAs (Qiao and Cech, 2008; Wilusz et al., 2012). However, the major focus of *in vivo* triplex research lies on a different complex family involving the concept of regulatory non-coding RNAs (ncRNA).

The recent discoveries of an ever-growing number of different classes of functional ncRNAs, together with the apparent enrichment of putative triplex forming sites in regulatory genome regions, lead to the idea of a possible role for triplex structures as interface between functional ncRNAs and chromatin (Rinn and Chang, 2012). This mechanism suggests the formation of triplexes between ncRNAs and the DNA that serve as molecular beacon for the recruitment of chromatin modifying complexes to their specific target sites, as illustrated in (Figure 2.6).

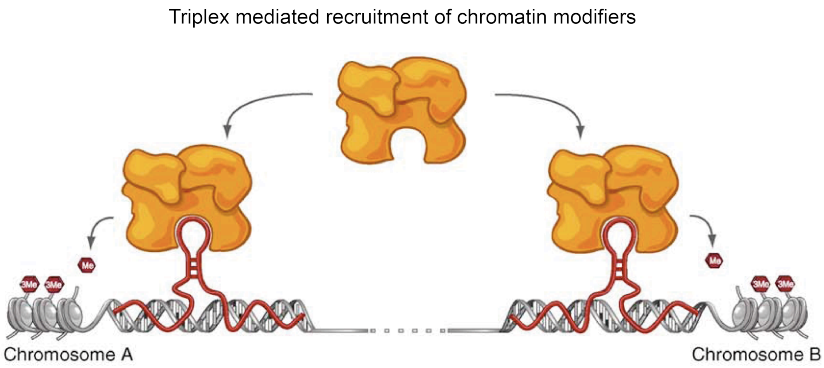


Figure 2.6: Triplexes between DNA and ncRNAs mediate the recruitment of chromatin modifiers

A recent model for triplex function in the context of ncRNA mediated chromatin regulation suggested a role as molecular beacon for the recruitment of chromatin modifying enzymes to their specific targets. The triplex would therefore provide the interface between functional ncRNAs and the chromatin. (modified from (Koziol and Rinn, 2010))

This mechanism implies the presence of TFOs within ncRNAs residing in the nucleus. Strikingly, a bioinformatical analysis of high throughput sequencing data of chromatin associated RNAs derived from ENCODE, revealed an enrichment of TFOs in this RNA fraction, underlining the possibility of triplex formation between DNA and ncRNAs (Buske et al., 2012). Accumulating evidence for the existence of such a mechanism is provided by a report on the regulation of the dehydrofolate reductase (DHFR) gene. The authors describe a ncRNA transcribed from an upstream promoter element, which supposedly forms triplex structures at the gene promoter inducing transcriptional repression of the genes (Figure 2.7) (Martianov et al., 2007). A similar mechanism was recently proposed for the transcriptional repression of the ribosomal RNA (rRNA) genes, however the nature and the formation of the triplex in this case are still highly controversial (Schmitz et al., 2010).

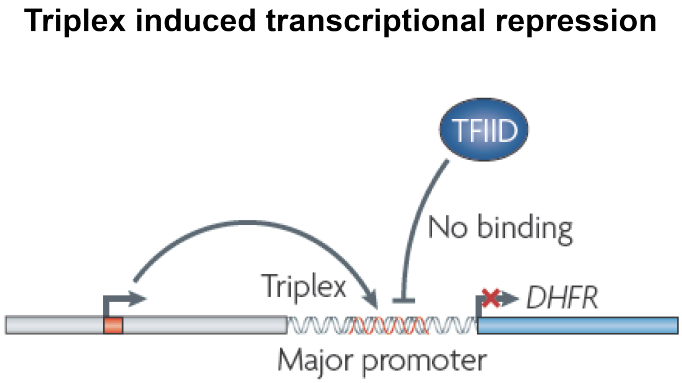


Figure 2.7: Transcriptional repression of the DHFR gene by triplex formation
A ncRNA is transcribed from an upstream promoter element and forms a triplex with a complementary sequence in the gene promoter. This inhibits the binding of TFIID and the formation of the initiation complex, which shuts down transcription of the gene. (from (Mercer et al., 2009))

Besides the functions in transcriptional regulation, triplex forming ncRNAs have also been implicated as factors playing a role in chromatin organization and architecture. A recently discovered GAA-repeat rich long non-coding RNA (lncRNA) that has been shown to interact with nuclear matrix binding proteins and locates to the nuclear matrix, was implicated in linking DNA repeat regions by triplex formation (Buske et al., 2011; Zheng et al., 2010). These triplexes could serve as anchor points or landmarks for the organization of specific chromatin domains to the nuclear matrix, similar to the model proposed by Agazie and colleagues (Agazie et al., 1996) (also illustrated in Figure 2.4).

2.3 The nuclear matrix

Since almost all DNA sequences supporting the formation of non-canonical DNA structures, like triplexes or quadruplex DNA have been shown to associate with the nuclear matrix, it presents a concept that brings together these various DNA elements and provides a consensus idea for their possible functions in genome organization and regulation.

2.3.1 Defining the nuclear matrix

Originally the nuclear matrix was defined as nuclear components resistant to extensive DNAse I digestion and high salt extractions (Berezney and Coffey, 1974). Although its exact nature is still under debate, it is supposed to represent a nuclear scaffold consisting of intermediate filament proteins like lamins, heterogeneous nuclear ribonucleoprotein particles, specific non-histone chromatin proteins, RNAs and associated DNA, which represents the matrix-attachment regions (MARs) of the genome (Elcock and Bridger, 2008; He et al., 1990; Ma et al., 1999). In addition it has been implicated to play a role in chromatin organization and regulation, by serving as an anchor point for chromatin loops, and providing a scaffold for the build up of regulatory protein complexes (reviewed in (Ottaviani et al., 2008)).

2.3.2 Nuclear matrix functions

The chromosomes in the interphase nucleus are supposed to occupy distinct areas termed chromosome territories (Cremer and Cremer, 2006; Cremer et al., 2006). Several studies suggested that the organization of these territories is mediated through the anchoring of chromatin loops of 50-200 kb in length to the nuclear matrix (Berezney et al., 1995; Ma et al., 1999). It was further hypothesized that this association is mediated by MAR sequences located within the chromatin domains (Heng et al., 2004; Laemmli et al., 1992). This tethering of DNA elements to the nuclear matrix is a dynamic process and not only shapes the nuclear architecture, but has also been shown to have specific effects on gene regulatory processes like transcription or replication (Eivazova et al., 2007; Heng et al., 2004; Radichev et al., 2005). Furthermore the anchor points on the nuclear matrix serve as platforms for transcription factory assembly and locate distal regulatory elements in close proximity to each other (Hart and Laemmli, 1998; Kimura et al., 1999).

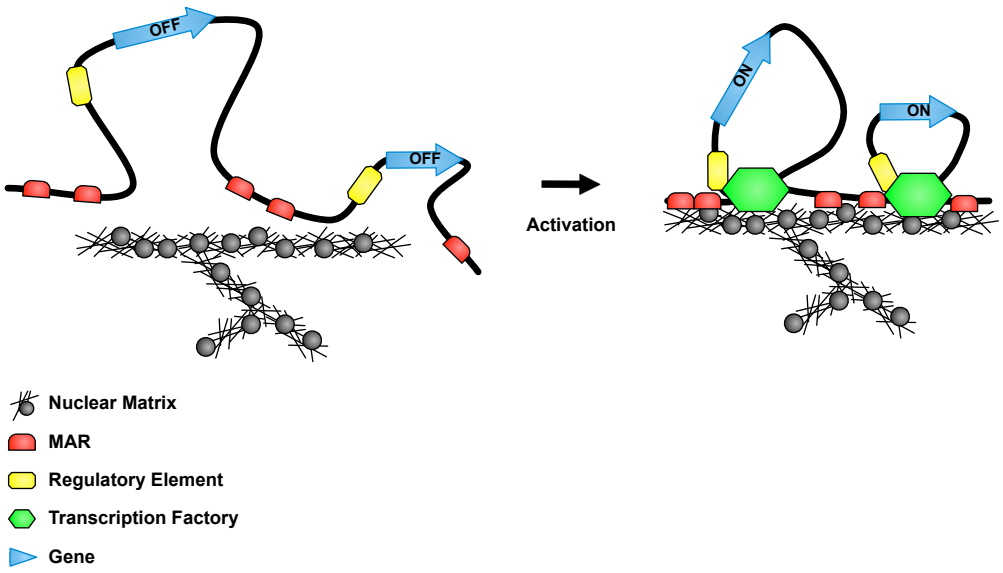


Figure 2.8: Concept of nuclear matrix function
Upon activation, chromatin is tethered to the nuclear matrix by its MARs. This results in the organization of chromatin loops that are anchored to the matrix. These anchor points on the matrix serve as scaffolds for the assembly of transcription factories and mediate the interactions of normally distal regulatory elements. Which in the end leads to transcriptional activation of the gene.

The key components in nuclear matrix function are the MAR elements, distributed throughout the genome. They are defined operational as motifs associated with the nuclear matrix and do not present any consensus sequences. MARs contain a wide array of DNA elements, like polypurine and polypyrimidine stretches, AT-rich elements, topoisomerase II binding sites, mirror repeat sequences and intrinsically kinked or curved DNA (Glazko et al., 2003). Importantly these MAR sites harbor most of the DNA elements also known to form non-canonical DNA structures. In general one can distinguish between two classes of MARs. Structural MARs are permanently associated with the nuclear matrix and mostly serve as molecular anchors to ensure proper chromatin organization. Functional MARs on the other hand are transiently tethered to the nuclear matrix, where they induce chromatin remodeling and regulate gene transcription (Elcock and Bridger, 2008; Glazko et al., 2003; Linnemann et al., 2007). In addition they interact with MAR binding proteins like CTCF at the nuclear matrix, playing further roles as insulator or enhancer elements (Allen et al., 2000; Bode et al., 1995; Forrester et al., 1994). It was hypothesized that the structure of a MAR is actually relevant for its function, and not its underlying sequence (Fukuda, 2000). Since most MARs contain DNA sequences that are prone to form non-canonical DNA structures, it is tempting to

speculate about possible roles for these structures in nuclear matrix organization, by working as molecular landmarks or tether for specific proteins or ncRNAs. Strikingly, this would also be in line with the recent findings concerning the GAA-repeat rich long non-coding RNA described previously (Zheng et al., 2010).
In summary the concept of the nuclear matrix functionally assembles the different DNA elements, prone to form non-canonical DNA structures, like triplexes or quadruplex DNA in the MAR sequences. It furthermore provides a mechanistic model for a possible role of these structured DNA elements in chromatin organization and regulation, which will be further investigated in course of the presented work.

2.4 The mammalian ribosomal RNA genes

2.4.1 Organization of the mammalian rRNA genes

The mammalian rRNA genes provided the model system of choice for further investigations of the possible functions of triplex structures and MARs in gene regulation and genome organization. These are multi-copy genes, with approximately 400 copies per cell. About 80% of the transcriptional activity within a eukaryotic cell is devoted to rRNA transcription, in order to keep up a ready supply of ribosomes for protein synthesis (McStay and Grummt, 2008; Nemeth and Langst, 2011). The genes are clustered in the nucleolar organizer regions (NORs) on the short arms of the acrocentric chromosomes (chromosomes 13, 14, 15, 21 and 22 in humans and 12, 15, 16, 17, 18 and 19 in mouse, respectively) were they are mostly, but not exclusively arranged in tandem repeated arrays in a head-to-tail fashion (Caburet et al., 2005; Dev et al., 1977; Henderson et al., 1972). The mammalian rDNA repeat units are quite large with roughly 43 kb in humans and about 45 kb in mice and consist of the rRNA gene and an intergenic spacer (IGS) (Gonzalez and Sylvester, 1995; Grozdanov et al., 2003).

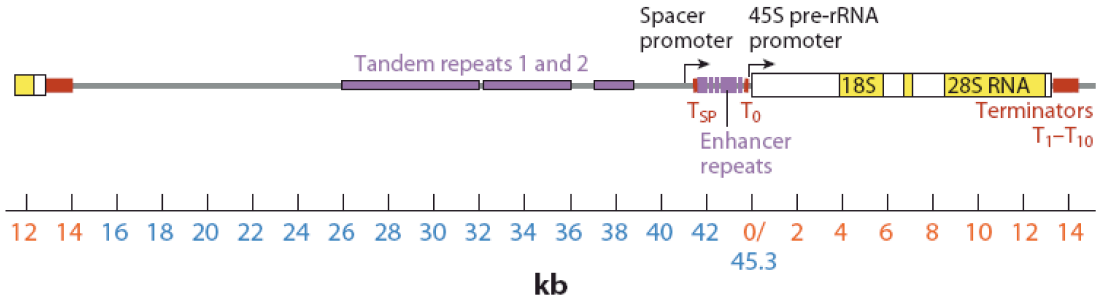


Figure 2.9: Structural organization of the mouse rDNA repeat unit
The structural organization of a mouse rDNA repeat unit is depicted. With the terminator elements illustrated in red, the repetitive enhancer repeats in purple and the locations of the spacer promoter and the promoter are indicated. The scale bar depicted is in kb and the zero point represents the transcription start site of the rRNA gene. (from (McStay and Grummt, 2008))

The rRNA coding region is transcribed by RNA polymerase I (RNA PolI) resulting in the 45S pre-rRNA precursor transcript, which is then processed into the 18S, 5.8S and 28S rRNAs needed for ribosome synthesis. The approximately 30 kb long IGS region contains regulatory elements, including the gene promoter, a spacer promoter, repetitive enhancer elements, and transcription terminators (McStay and Grummt, 2008). The gene promoter consists of two functional parts, a core promoter in close proximity to the transcription start site, and an upstream control element approximately 100 nucleotides further upstream (Haltiner et al., 1986; Learned et al., 1986). The gene is flanked by transcription terminators at its 5'-and 3'-ends, which are binding elements for the transcription termination factor I (TTF-I), that stops the elongation of RNA PolI and in addition plays a central role in the epigenetic regulation of rRNA genes (Grummt et al., 1986; Grummt et al., 1985; Henderson and Sollner-Webb, 1986; McStay and Reeder, 1986).

2.4.2 Epigenetic regulation of rRNA genes

About half of the 400 rRNA gene copies within a mammalian cell are silenced. The active and inactive repeats can be distinguished by different epigenetic states, which mirror their transcriptional status and result in specific chromatin structures that are maintained throughout the cell cycle.

The active rDNA repeats lack DNA methylation and are associated with RNA PolI and the PolI-specific transcription factors TTF-I and UBF. In general the

transcribed rRNA genes present euchromatic features and are permissive for transcription. They are associated with acetylated histones H3 and H4 and in addition carry H3K4me3 marks, which are all signs of active transcription (reviewed in (McStay and Grummt, 2008)). Furthermore, these genes are nucleosome depleted and have been suggested to form a loop structure between terminator and promoter elements, which is mediated by TTF-I and c-Myc and provides a structural basis for an efficient re-initiation of RNA PolI (Nemeth et al., 2008; Shiue et al., 2009).

The DNA of inactive rRNA genes on the other hand is methylated, especially in the enhancer and promoter regions and the repeats are packed in dense heterochromatin. The silenced rRNA genes are devoid of the PolI specific transcription factors and are associated with methylated H3K9, H3K20 and H3K27, which represent repressive histone marks, indicating inactive chromatin (reviewed in (McStay and Grummt, 2008)).

2.4.3 NoRC is a key player in rRNA repression

The epigenetic repression of rRNA genes is a complex process, which involves the interplay between different protein factors like remodeling enzymes and histone modifiers, resulting in large-scale alterations of the chromatin structure. The key player within rDNA silencing in human and mouse cells is the nucleolar remodeling complex NoRC. It consists of two subunits, the catalytic ATPase SNF2h and the larger, 205 kDa protein termed Tip5 (TTF-I interacting protein 5) (Strohner et al., 2001). Tip5 is the regulatory subunit of NoRC and physically interacts with TTF-I, which leads to the recruitment of the protein complex to the rRNA promoter. There, NoRC acts as a multifunctional chromatin-dependent regulator, coordinating histone deacetylation and methylation, ATP-dependent chromatin remodeling and DNA methylation, thereby inducing heterochromatin formation and transcriptional silencing (Santoro et al., 2002; Strohner et al., 2004; Zhou et al., 2002).

The exact mechanism of NoRC mediated rDNA repression has been subject for extensive research. The analysis revealed an exact hierarchical order of

interdependent events of chromatin remodeling, histone deacetylation and DNA methylation (Santoro and Grummt, 2005). Initially, a physical interaction between Tip5 and TTF-I bound to the rRNA promoter recruits NoRC to the gene. Interestingly the association of a small ncRNA, termed pRNA, is a prerequisite for proper NoRC function. The pRNA is derived from a longer PolI transcript, which originates from the spacer promoter element in the IGS roughly 2 kb upstream of the gene promoter. It is only present in low abundance and does not accumulate in the cell, since it is supposedly processed into ncRNAs of 150-300 nucleotides in length, immediately. These transcripts are either rapidly degraded or shielded from further degradation by interacting with NoRC (Mayer et al., 2006). Tip5 mediates the binding of the pRNA by recognizing specific secondary structures, which is implicated to result in a conformational change within the NoRC complex and the pRNA according to an induced fit mechanism (Mayer et al., 2008).

Subsequently, NoRC interacts with the Sin3 corepressor complex, which contains the histone deacetylases HDAC1 and HDAC2 and leads to deacetylation of histones H4 in nucleosomes at the rRNA promoter region (Santoro and Grummt, 2005; Zhou et al., 2002). This in turn is a signal for ATP-dependent chromatin remodeling, which slides the promoter bound nucleosome into a position unfavorable for transcription (Li et al., 2006). Thereby specific DNA elements become available for DNA methylation by Dnmt3b, inhibiting the binding of transcription factors to the promoter and thus preventing transcription (Santoro and Grummt, 2001; Santoro et al., 2002). Finally all these events trigger the recruitment of additional factors like HP1 and result in the formation of densely packed heterochromatin.

As already stated before, recently an alternative mechanism has been suggested, where a triplex formation between the pRNA and sequence elements within the rRNA promoter triggers the recruitment of Dnmt3b, thus leading to DNA methylation and chromatin compaction (Schmitz et al., 2010). The exact role of these pRNA transcripts is therefore still not completely understood and needs further investigations.

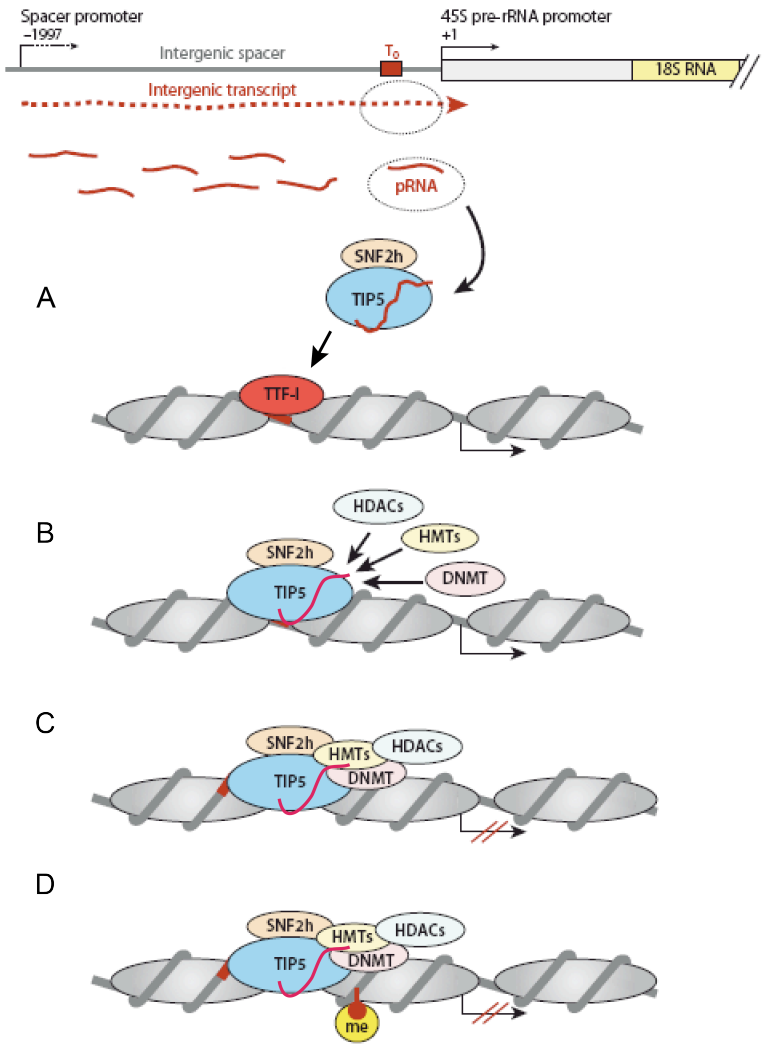


Figure 2.10: NoRC mediated silencing of the rDNA
A) An intergenic transcript is generated from the spacer promoter and immediately processed into 150-300 nucleotide long pRNAs, which are bound by NoRC through interactions with the TAM domain of Tip5. NoRC then gets recruited to the rRNA promoter through interactions with TTF-I. **B**) NoRC interacts with the Sin3 corepressor complex and HMTs, which leads to deacetylation of H3 and H4 and to methylation of H3K9, H3H20 and H3K27. **C**) These changes in histone modifications induce a chromatin remodeling by NoRC, which slides the promoter bound nucleosome into a position restrictive for transcription. **D**) In turn Dnmt3b methylates a no exposed CpG dinucleotide, thus inhibiting assembly of the pre-initiation complex. Finally the rDNA is packed into heterochromatin. (modified from (McStay and Grummt, 2008))

In summary, the regulatory key player NoRC has two different roles in the repression of rRNA genes. It induces chromatin remodeling, necessary for silencing and at the same time serves as scaffold for the assembly of the complex machinery that participates in this regulatory process.

2.4.4 Tip5 is a multifunctional protein with a modular structure

The various interactions and associations NoRC has to undergo in its role as functional scaffold in rDNA repression are largely mediated by its regulatory subunit Tip5. This protein illustrates a modular organization with a large number of different protein and nucleic acid interaction domains. The C-terminal located PHD (plant homeodomain) finger and bromodomain are required for the binding of Snf2h, HMTs (histone methyltransferases) and the association with HDAC1, HDAC2 and histone H4 acetylated at lysine 16 (H4K16ac), respectively. In addition Dnmts interact with the internal and also the C-terminal part of the protein as depicted in (Figure 2.11) (McStay and Grummt, 2008).

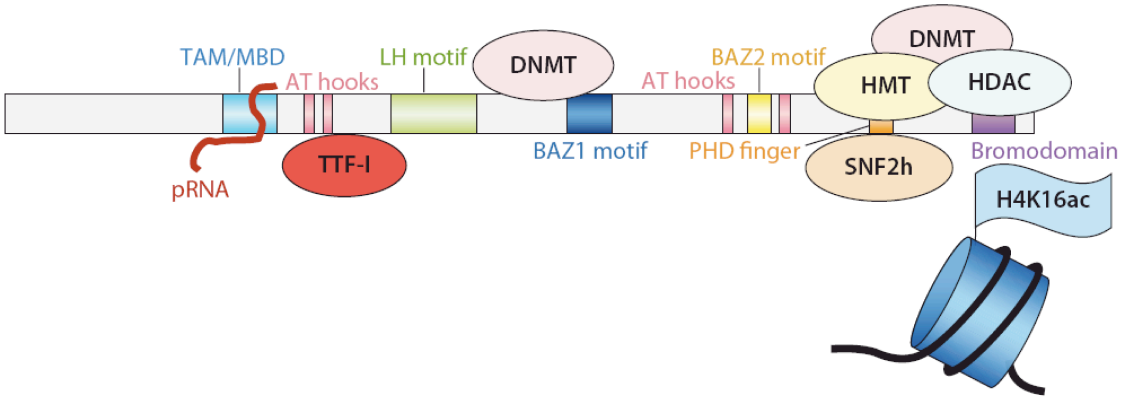


Figure 2.11: Modular organization of Tip5
The various functional domains of Tip 5 and their specific interaction partners are depicted. The C-terminal PHD finger and bromodomain are responsible for most protein-protein interactions mediated by Tip5. The TAM/MBD domain located N-terminal in the protein was shown to bind the pRNA. TTF-I and Dnmts bind in close to the core region of the protein. (from (McStay and Grummt, 2008))

More interesting in respect to this work however, are the depicted MBD (methyl-CpG binding domain)-like TAM (Tip5/ARBD/MBD) domain and AT-hook motifs located within the protein. Both domains have been assigned functions in pRNA recognition and binding to methylated DNA (Strohner et al., 2001; Zhou et al., 2009).

However, in general these motifs are supposed to represent minor groove DNA binders, which influence chromatin conformation, thereby enhancing accessibility

of functional DNA elements for regulatory factors (Baker et al., 2013). More interestingly both domains have also been shown to bind to MARs (Aravind and Landsman, 1998), which could suggest a role of Tip5 in nuclear matrix binding and association, which in turn could also implicate a function of the protein in rDNA organization to the nuclear matrix.

2.4.5 Large-scale organization of rDNA chromatin

The epigenetic regulation of rRNA transcription has been characterized extensively and the role of Tip5 in these processes is well established. However little is known about the effects on the large-scale organization of rDNA chromatin, that goes along with the silencing of the rDNA repeats and a possible function for Tip5 in these events.

The rRNA genes are organized within the nucleoli, which can be divided into three distinct structural and functional compartments: the fibrillar center, the dense fibrillar component and the granular component, where the early steps of ribosome biogenesis take place. The actively transcribed rRNA genes are located in the fibrillar center or at the fibrillar center-dense fibrillar component border. The inactive rRNA genes are placed outside this region (Nemeth and Langst, 2011). It has been shown that induced inhibition of rDNA transcription for example through serum starvation, leads to morphological changes in nucleolar architecture (Chan et al., 1985; O'Mahony et al., 1992). This may in part correlate with changes in the large-scale chromatin organization of rDNA repeats upon NoRC mediated silencing. Interestingly the sequencing of the mouse rDNA repeat revealed the presence of predicted MARs within the IGS (Grozdanov et al., 2003). Additionally, in matrix preparations the rDNA was found to be specifically enriched in the nuclear matrix fraction (Bolla et al., 1985; Keppel, 1986; Smith and Rothblum, 1987). The described MAR binding function of the TAM and the AT-hook domains in Tip5 made it tempting to speculate about a role of the protein in mediating rDNA recruitment to the nuclear matrix. This could present a mechanism to separate the active rRNA genes from the silenced ones by relocating them to the nuclear matrix.

3 Results

3.1 Characterization of putative triplex forming sites located in the rDNA enhancer

3.1.1 The mouse rRNA genes contain clusters of putative triplex forming sites in their regulatory elements

In recent years several studies emerged, investigating the number and distribution of possible triplex forming sequences throughout the mouse and human genomes (Buske et al., 2012; Goñi et al., 2004; Goñi et al., 2006; Wu et al., 2007). Therein it was illustrated, that the two regions mostly enriched for these motifs are regulatory and repetitive sequences within the genome. Interestingly, the IGS within the mouse rDNA repeat unit not only contains regulatory enhancer, promoter and terminator regions, but also harbors a large number of different repetitive elements. Strikingly, in this context the regulatory terminator and enhancer sequences only contain mononucleotide repeats, resembling classic low complexity triplex target sites (Grozdanov et al., 2003). In addition, the frequency of mononucleotide repeats within the regulatory rDNA regions displays a clear enrichment (8-fold) compared to the rest of the mouse genome (Grozdanov et al., 2003). Since these facts hinted at a large triplex forming potential of the regulatory elements within the mouse rDNA repeat, the sequence (GenBank AccNo.: BK000964.1) was investigated using a recently developed *in silico* triplex target site prediction tool (Buske et al., 2012). In that way the triplex forming potential of the low complexity regions located within the enhancer and terminator sequences was elucidated (Figure 3.1/Table 3.1).

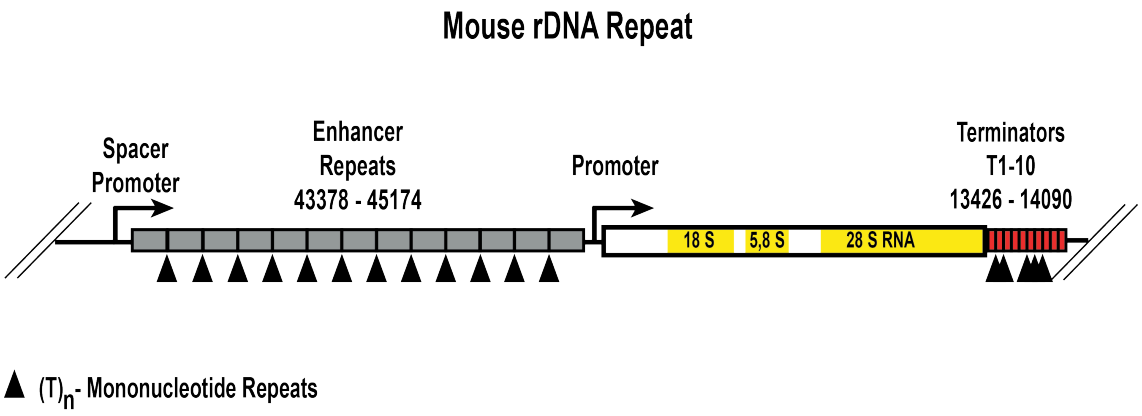


Figure 3.1: Localization and distribution of mononucleotide sequences within the mouse rDNA enhancer and terminator regions

The structural and functional organization of a mouse rDNA repeat unit is depicted here. The IGS contains the spacer promoter and promoter elements, as well as the regulatory terminator and enhancer regions harboring the $(T)_n$-mononucleotide repeats. The black triangles mark the location of the repeat stretches, which are part of putative triplex forming pyrimidine runs, as proposed by the performed *in silico* analysis.

The results presented in (Figure 3.1) and (Table 3.1) revealed a striking triplex forming potential of the pyrimidine runs located within the terminator and enhancer sequences of the rRNA genes. Especially the results for the enhancer were noticeable, revealing a clustering of partly overlapping, triplex forming sites in the $(T)_n$-stretch containing pyrimidine sequences, located between the 140-bp enhancer repeat units. Since these findings illustrated an enrichment of possible triplex forming structures within the regulatory rDNA regions, further investigations on the characteristics and putative functions of these elements with regard to rDNA organization and regulation were started.

rDNA Enhancer Triplex Target Sites

Start in Sequence	End in Sequence	TTS Sequence	Score
43410	43431	GAGGAAAAAAAAAAAAAAGGG	18
43541	43573	GGAGGAAAAAAAAAAAAAAAAAAAAAAAAGA	30
43685	43715	GGAGAAAAAAAAAAAAAAAAAAAAAAAAAA	30
43947	43969	GGAGGAAAAAAAAAAAAAAGA	20
44078	44094	GAGGAAAAAAAAAAA	16
44333	44355	GGAGGAAAAAAAAAAAAAAGA	20
44463	44491	GGAGGAAAAAAAAAAAAAAAAAAAAGA	28
44600	44623	GGAGGAAAAAAAAAAAAAAAGA	21
44732	44753	GGAAGAAAAAAAAAAAAAAGA	21
44995	45024	GGAGGAAAAAAGAAAAAAAAAAAAAAGA	27

rDNA Terminator Triplex Target Sites

Start in Sequence	End in Sequence	TTS Sequence	Score
13556	13584	AAAAAAAAAAAAAAAAAAAAAAAAGGGG	28
13742	13764	AAAAAAAAAAAAAAAAAGAGAA	20
13858	13875	GGAAAAAAAAAAAAAAA	17
13958	13978	AAAAAAAAAAAAAAAAAGAG	20

Table 3.1: Putative triplex target sites in the mouse rDNA enhancer and terminator region
The mouse rDNA repeat sequence (GenBank AccNo.: BK000964.1) was screened for putative triplex target sites using a recently developed software tool (Buske et al., 2012). The settings where adjusted for sequences between 15 to 30 bp in length, in order to determine motifs with the potential of forming stable complexes. Furthermore, the maximal allowed G content was set to 20 %. There were no non-pyrimidine mismatches tolerated and the low complexity filter was turned of, since the aim of the analysis was to get an insight on the triplex forming potential of the mononucleotide repeat containing elements within the enhancer and terminator regions. The table displays the putative triplex forming sites in these two elements determined by the analysis. The start and endpoints in respect to the annotation in the reference sequence are depicted on the left. The target motifs are displayed as – strand DNA sequences. The score on the right represents the triplex forming potential calculated by the algorithm. Since the sequences presented here a very similar the score is mainly determined by the length of the motif, with a longer motif providing more stabilizing Hoogsteen-interactions, therefore possessing a higher triplex forming potential.

3.1.2 The rDNA enhancer motif forms a stable triplex complex *in vitro*

The predicted triplex forming sites within the rDNA enhancer were verified by an initial *in vitro* characterization of various candidate sequences using EMSA and melting curve analysis (Lang, 2008). Thereby the triplex motif spanning the rDNA sequence from 44995 to 45024 (enhancer triplex) was identified as a candidate concerning a potential function in transcriptional regulation of the rRNA genes. This assumption was based on one hand on the observed relatively high triplex forming potential with a predicted score of 27. And on the other hand on the close spatial proximity to the gene promoter, which has been shown to be the target site of a complex regulatory network of protein complexes and ncRNAs, potentially also involving a DNA/RNA triplex (Schmitz et al., 2010). Furthermore the triplex forming potential could be validated in the EMSA and melting curve analysis. Therefore this motif was used as a model sequence for the further in depth *in vitro* characterization presented here.

To better detect the triplex formation of the enhancer motif, EMSA experiments with fluorescently labeled oligonucleotides were performed. With this assay it was not only possible to determine the triplex formation induced shift of the DNA, but

at the same time it allowed the detection of the TFO behavior. The results of the EMSA are shown in (Figure 3.2).

The DNA/DNA and the DNA/RNA triplex-forming reactions revealed a clear shift of both the DNA and the free TFO bands, representing the complex formation. Upon a four to one ratio the triplex formation was complete and almost all DNA was incorporated into the complex (lane 4 DNA/DNA triplex, lane 7 DNA/RNA triplex). The EMSA showed a reduced signal intensity of the Cy5-label (red) and the FAM-label (green), upon interaction of DNA and TFO. The effect on the Cy5-label was more profound for the reactions involving the RNA TFO pictured in lanes 5 to 7. For the DNA/RNA triplex there was also an additional faint Cy5 signal visible in between the free DNA and the triplex bands. In general the triplex formation was better visible for the DNA TFO.

In summary the results suggested that the formation of triplex complexes could be resolved by the color-coded EMSA approach, thereby validating the method as a useful tool for qualitative investigations. However, due to the quenching effects it was not possible to make any quantitative statements concerning the affinities of the TFOs with this assay.

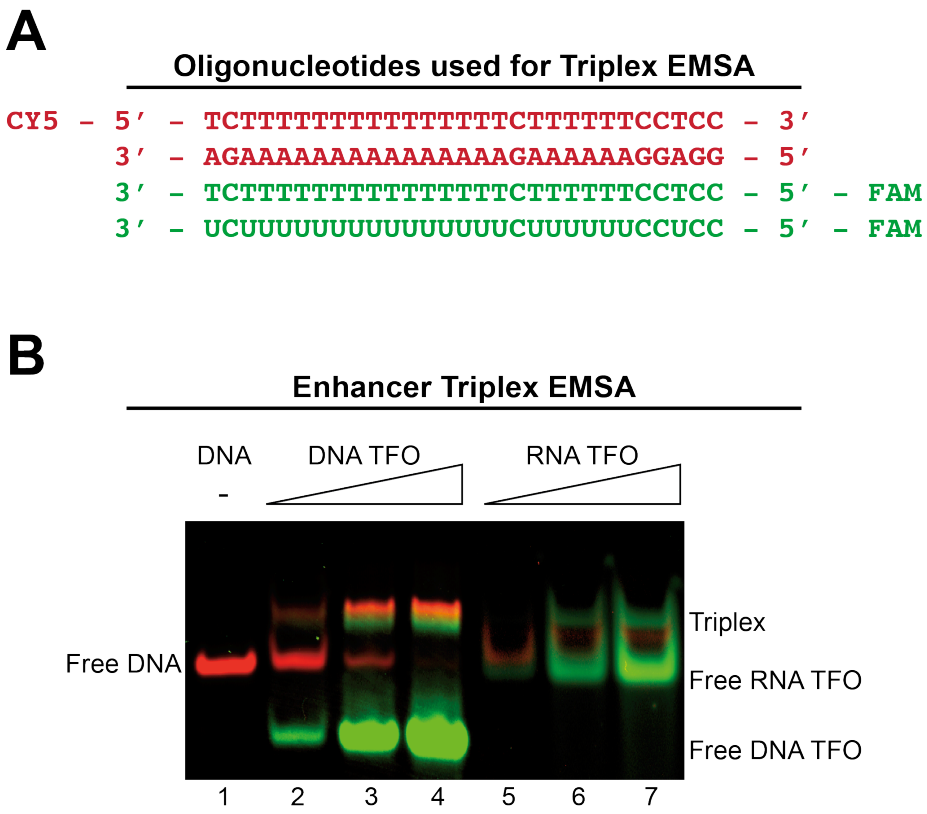


Figure 3.2: Two color triplex EMSA analysis
A) The principal setup of the experiment is illustrated. The DNA strand with the enhancer sequence was labeled with a Cy5-moiety on its 5'-end (shown in red). The two different TFOs, DNA and RNA, carried a FAM-label on their 5'-ends (shown in green). The TFOs bind to the DNA target in a parallel orientation to the purine strand, resulting in a triplex complex as depicted. Therefore the fluorescent labels are positioned on opposite sides of the complex.
B) The triplex formation was performed as described in (5.2.1.9). One pmol of DNA was used in each reaction and in the free DNA control. The triplex formation was carried out with rising amounts of TFO, starting at a 1:1 ratio, increasing to a 1:2 and a 1:4 ratio in respect to the target DNA. The annealing reactions were loaded on a native 15 % Tris-Acetat gel containing 10 mM MgOAc and electrophoresis was performed at 100 V for 2 h at room temperature.

3.1.3 Triplex formation can be monitored by Microscale Thermophoresis

While the triplex EMSA provided qualitative insights in whether the intended sequence motifs are capable of forming a triple helical complex, the quantitative analysis of complex formation under changing experimental conditions was not possible using this approach. Therefore an assay to monitor enhancer triplex formation via the MST technology was established. The goal of these experiments was to use this new technology for a fast and easy evaluation and characterization

of putative triplex forming sequences, thereby validating interesting target sites and their triplex forming potential *in vitro*.

3.1.3.1 Third strand binding leads to an increase of fluorescence in MST experiments

For this purpose, the same Cy5-labeled DNA sequence comprising the enhancer triplex motif, which was used for the EMSA analysis (Figure 3.2), was employed for the MST experiments. In contrast to the previous EMSA studies however, the RNA and DNA TFOs were not labeled with a fluorescent dye. In addition, two unspecific TFOs (RNA/DNA control TFO) comprising an unspecific sequence from the GFP coding region were used as negative controls. These oligonucleotides had the same length as the specific TFOs, but contained no continuous pyrimidine or purine stretches and should therefore not form a triplex. For the measurements a constant DNA concentration of 100 nM per capillary was used and a serial dilution of increasing concentrations, of either RNA or DNA TFO was titrated to the target sequence. The experimental settings were identical to the triplex EMSA analysis. The initial capillary scans shown in (Figure 3.3) revealed an increased fluorescence signal, correlating with the rising TFO concentration (Figure 3.3, upper panel). This observation suggested an interaction of the oligonucleotides to the DNA duplex, thereby directly influencing the Cy5-label. For the control TFOs no change in the fluorescence signal could be detected, indicating that the increase in fluorescence mirrors the specific binding event (Figure 3.3, lower panel).

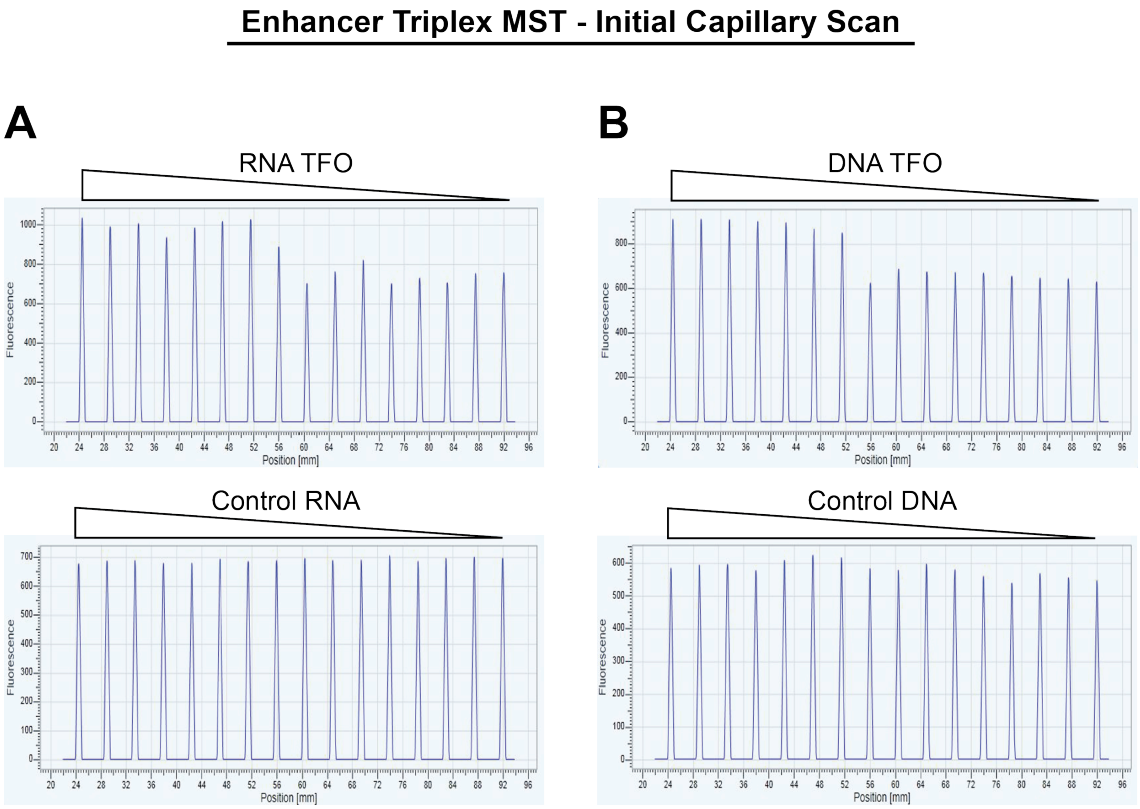


Figure 3.3: Triplex formation induces an increase in fluorescence
A) Triplex formation was performed as described in (5.2.8.2). The DNA concentration was set to 100 nM and a serial dilution of a 1:1 ratio of RNA TFO was titrated to the target DNA starting at 10 µM and finishing at 0.6 nM. Approximately 5 µl of each reaction was sucked into a standard treated glass capillary and incubated in the Monolith MST device at room temperature for 15 min. Afterwards the initial capillary scan was performed at an LED power of 20%. This was done once for the RNA TFO (upper graph) and the control RNA (lower graph). The black triangle resembles the increasing TFO concentration. The increase of fluorescence in correlation with the increasing TFO concentration is clearly visible.
B) Same as described above, only this time the DNA TFO and control TFO were used.

3.1.3.2 Triplex formation induces changes in the thermophoretic behavior of the DNA substrate

The thermophoresis curves generated for triplex formation by the Monolith MST device, illustrated in (Figure 3.4), clearly showed that the different states of triplex formation were reflected in a changing thermophoretic behavior. For the DNA/RNA triplex depicted on the left side in (Figure 3.4A), three different states could be observed correlating with the increase in TFO concentration within the samples, thereby reflecting a change from unbound to fully bound DNA duplex substrate. No change in thermpophoresis was detected for the RNA control TFO. The DNA/DNA triplex exhibited a similar thermophoresis pattern, as shown on the right side in (Figure 3.4B). In contrast to the profile of the DNA/RNA triplex, the

changes in the thermophoresis signals for DNA/DNA triplex were not as strictly separated in three states. Instead the transition from unbound to completely bound DNA duplex seems to involve more than one intermediate state. Additionally the thermophoresis curves corresponding to the completely bound DNA molecules were different from the DNA/RNA triplex, in that they crossed the thermophoresis curves of the unbound state at the 30 sec time point of the measurement cycle. Again, the control TFO showed no interaction with the enhancer sequence.

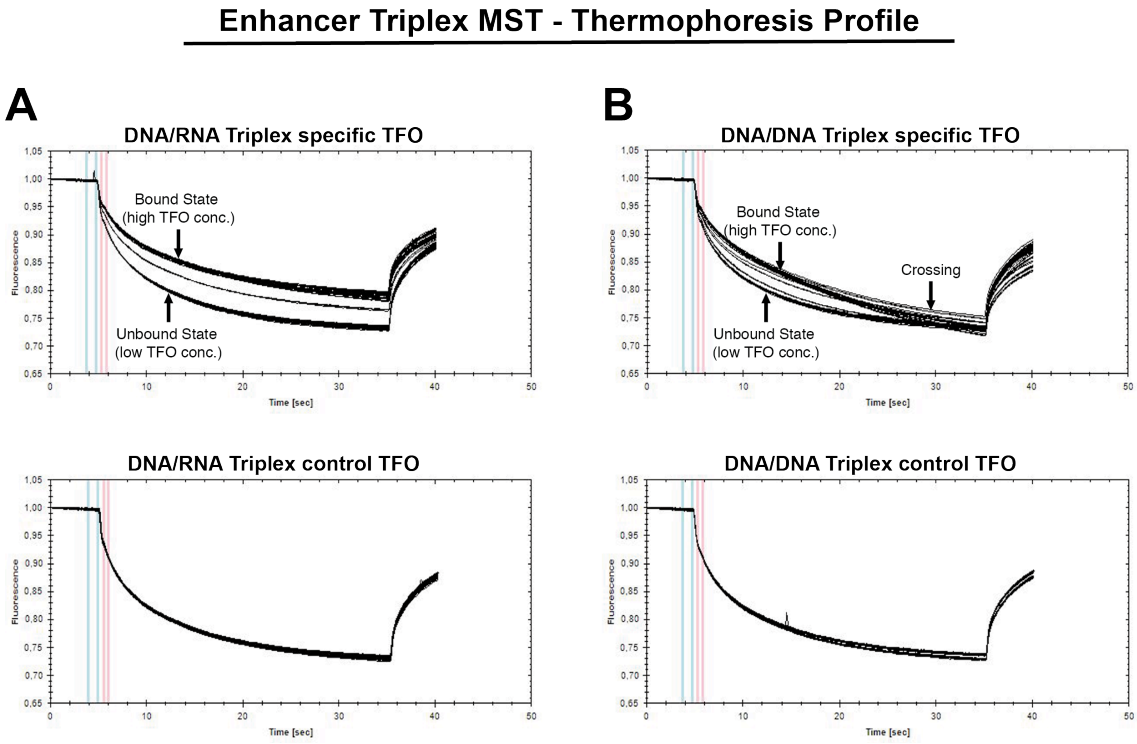


Figure 3.4: Triplex formation monitored by MST measurements
A) The thermophoresis curves for the DNA/RNA TFO triplex (upper figure) and the DNA/control RNA (lower figure) experiments already described in (Figure 3.3) are illustrated. MST reactions were performed at 30% MST power with a laser on time of 35 seconds at room temperature. On the y-axis the normalized fluorescence is shown and the x-axis represents the measurement time in seconds. A clear difference in the thermophoretic patterns between specific TFOs binding to the DNA and the control TFOs exhibiting no binding is shown. The arrows in the upper panel indicate the different states of the DNA duplex, with the shift from unbound to fully bound DNA.
B) As described for A), but this time the DNA/DNA TFO and DNA/control DNA reactions are shown. An arrow also marks the crossing of the thermophoresis curves at 30 sec described in the text.

3.1.3.3 Quantitative analysis of triplex formation by MST

The previous results in (Figure 3.3) showed a direct influence of the binding of the TFO on the fluorescence of the 5'-located Cy5 label. Furthermore, previous studies suggested that the binding of the third strand to the major groove of the DNA follows a highly cooperative nucleation zipper model, starting at the sequences 5'-ends (Alberti et al., 2002). Therefore the temperature jump signals derived for the different experiments (represented by the blue and red lines in the thermophoresis blots in (Figure 3.4)) were used for the quantitative analysis of the experiments, since they describe localized binding effects in close proximity of the fluorescent label (see 5.2.8.1). The generated temperature jump signals were plotted against the TFO concentrations and binding curves, and K_d-values for DNA/RNA and DNA/DNA triplex formations were calculated (5.2.8.2).

The results are illustrated in (Figure 3.5) and match the observations made with regard to the thermophoretic behavior (Figure 3.4). For the DNA/RNA triplex a K_d of 11.3 +/- 8.6 nM and for the DNA/DNA triplex a K_d of 16.3 +/- 4.6 nM was calculated. The control TFOs showed no interaction with the enhancer DNA. Interestingly the DNA/RNA triplex displayed a highly cooperative binding behavior with a calculated Hill coefficient of 7.7 (data not shown), which resulted in only one intermediate state observed in the MST analysis (Figure 3.5A). In case of the DNA/DNA triplex the calculated Hill coefficient of 2.1 (data not shown) was clearly lower indicating a weaker cooperativity in complex formation.

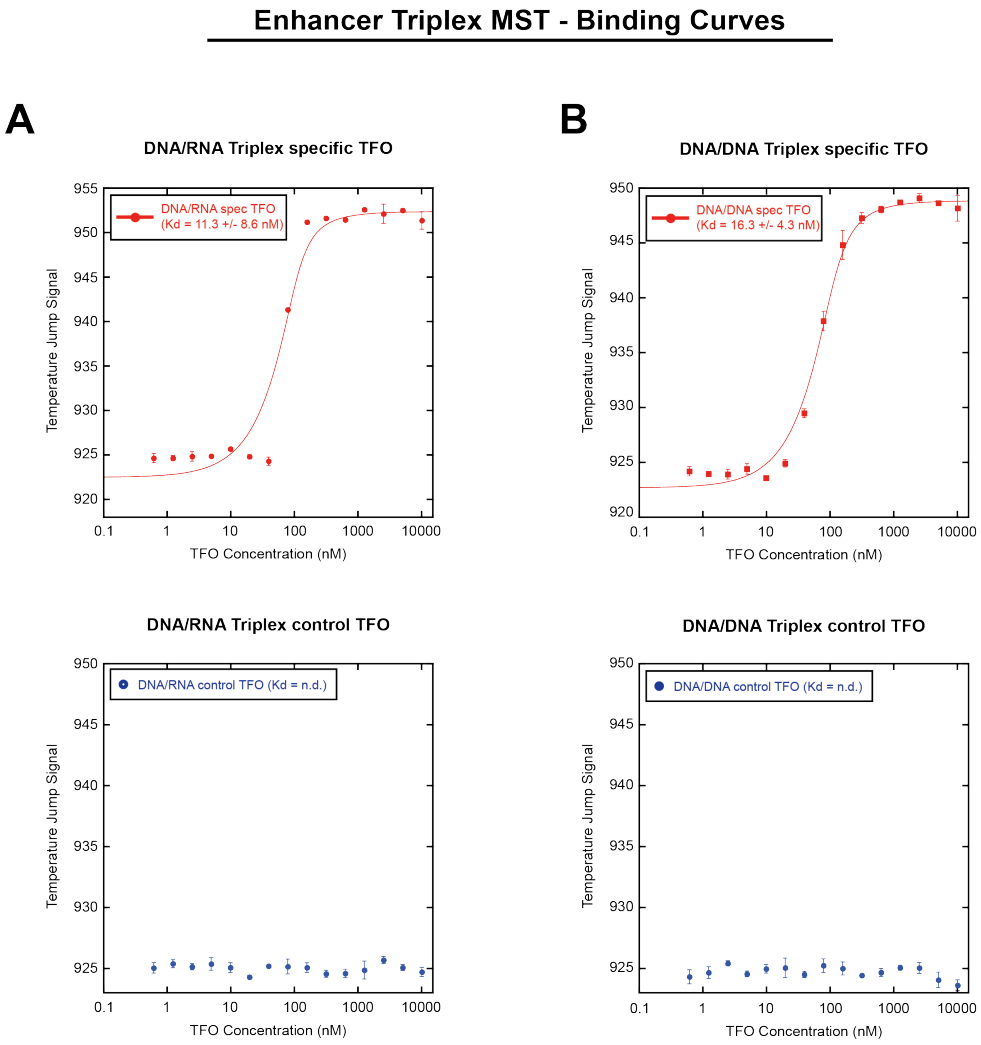


Figure 3.5: Quantitavtive analysis of triplex formation via MST
Binding curves and K_d-values were calculated from the thermophoresis experiments described in Figure 3.3 and Figure 3.4 for the DNA/RNA triplex (**A**) and DNA/DNA triplex (**B**) reactions. The temperature jump signals generated by the MST analysis are plotted against the TFO concentration in nM on a logarithmic scale. The upper figures illustrate the binding of the specific TFOs. In the lower figures the control reactions displaying no interaction are shown. The data points of the plots represent mean ± standard deviation values of four measurements.

3.1.4 The thermophoretic behavior of triplex complexes changes with stability

The previous experiments revealed a higher affinity and a stronger cooperativity in DNA binding for the RNA TFO compared to the DNA TFO. Strikingly, these differences in complex formation and stability seemed to translate into differences in the thermophoretic behavior between the DNA/RNA and DNA/DNA complexes. This was indicated by the different thermophoresis profiles illustrated in (Figure

3.4) with the loss of the clear 3 state profile and the crossing of the thermophoresis curves shown for the DNA/DNA triplex.

To get further insights on the effects of TFO binding affinity and triplex stability on the thermophoretic behavior, the same experimental setup as described for the initial triplex MST analysis was chosen. The incubation temperature was raised stepwise after each round of MST measurement, starting from 25°C up to 60°C, which represents the temperature limit of the Monolith MST device. For the analysis, thermophoretic profiles of each measurement were recorded and the resulting temperature jump, as well as the combined temperature jump and thermophoresis signals were plotted against the TFO concentration. Thereby the effects of the thermophoretic differences on the binding curves were visualized. The results are illustrated in (Figure 3.6).

On the left side the thermophoretic profiles of each measurement are depicted. The blue lines represent the cold status, before the induction of the laser. The green lines mark the time points used for the analysis of the temperature jump and finally the red lines indicate the data points utilized for the combined analysis of temperature jump and thermophoresis. On the right side the calculated binding curves for either the temperature jump data or the combined temperature jump and thermophoresis data are shown. As expected at 25°C the complex exhibited the same behavior that had been observed previously (Figure 3.4/Figure 3.5). There were no striking differences in the binding curves calculated from either the temperature jump data or the combined temperature jump and thermophoresis data. Interestingly, after an increase of the incubation temperature to 35°C a clear change in the thermophoretic behavior was visible. The curves now resembled the profile seen for the DNA/DNA complex in (Figure 3.4/Figure 3.5). They displayed a loss of the 3 state profile and crossing of the thermophoresis curves at the 25 sec time point of the measurement cycles (Figure 3.6, second row). This change in thermophoretic behavior led to a scattering of the thermophoresis curves over the whole amplitude of the measurement at the late time points of the measurement cycle. This scattering resulted in a v-shaped distribution of the thermophoresia data points when included in the binding curve. In contrast the temperature jump data analysis presented no problem resulting in a proper binding curve, comparable to the 25°C data. The change in thermophoretic behavior of the bound

samples continued and at 50°C all samples representing the bound DNA finally showed a negative thermophoretic trend in respect to the unbound samples and therefore could be used for the generation of a binding curve again, as displayed in (Figure 3.6, third row). Strikingly at this point the curves started to show an overlap at the early time points representing the temperature jump data. As a result the amplitude of the binding curve for these data points revealed a massive decrease compared to the previous measurements. At 55°C it was impossible to generate a binding curve from the temperature jump data, since the different curves showed a complete overlap. Nevertheless, an interaction between the DNA and RNA TFO could still be detected using the combined temperature jump and thermophoresis data. When increasing the incubation temperature further, differences in temperature jump signals were increasing again and comparable to the thermophoresis data, resulting in similar graphs.

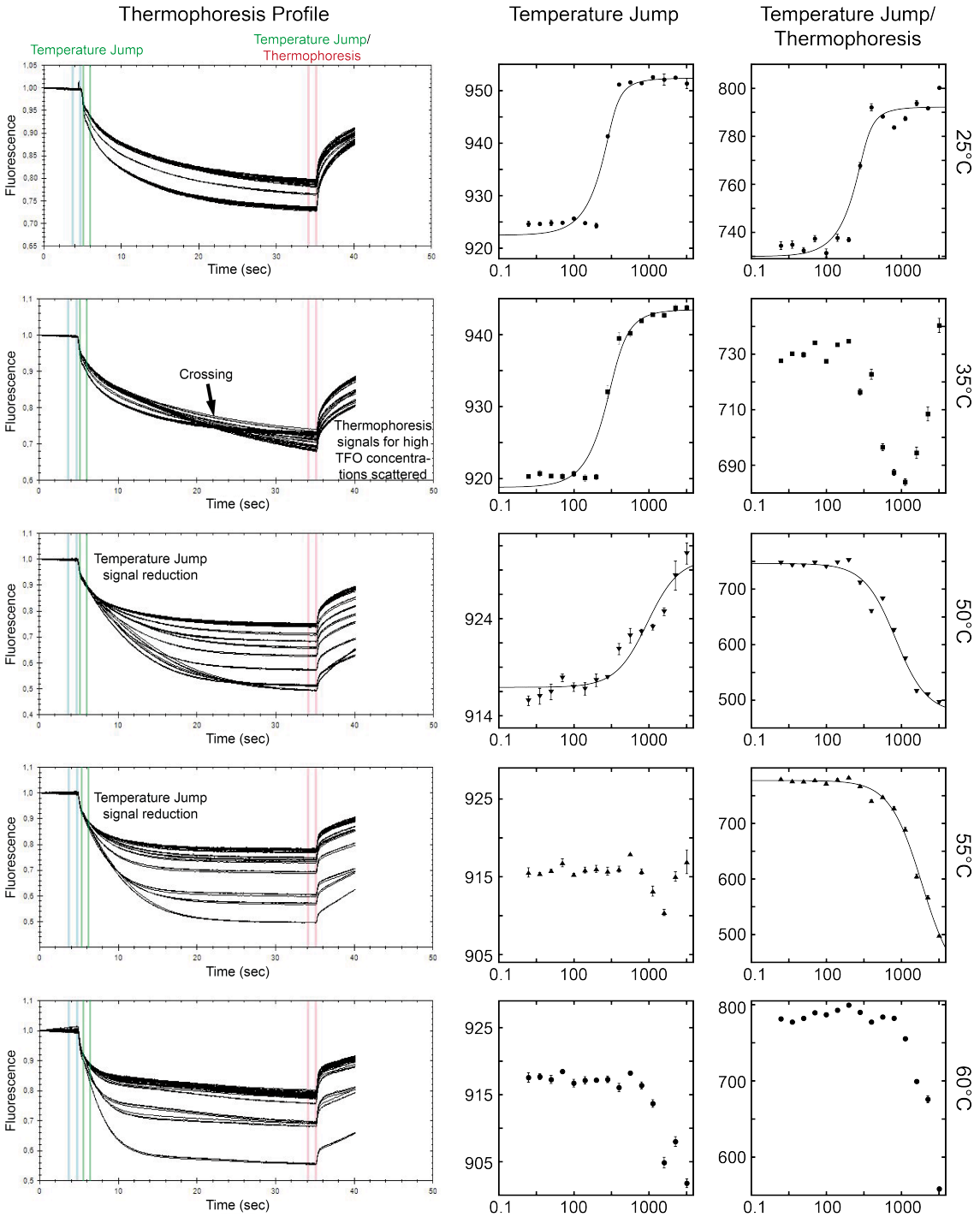


Figure 3.6: Changes in thermophoretic behavior during triplex MST analysis
DNA/RNA triplex formation was performed as described previously (Figure 3.3). MST experiments on the Monolith MST device were performed at 20% LED, 30% MST power and a laser on time of 35 sec. Before each measurement the samples were incubated for 15 min at the desired temperature in the MST device. On the left the thermophoresis curves recorded for selected temperatures (indicated on the right) are depicted. The normalized fluorescence is plotted against the measurement time in sec. The blue lines indicate the "cold" data points. The green lines depict the temperature jump data points and the red lines the thermophoresis data points. From these data points the binding curves for temperature jump, and combined temperature jump and thermophoresis data were generated, plotting them against the TFO concentrations in nM(x-axis, logarithmic). The error bars indicate the standard deviation of four technical replicates.

The observations presented here, lead to the conclusion that for the proper characterization of triplex formation based on MST experiments, the parameter used for the analysis of the complex formation had to be determined and adjusted for each experiment individually. The temperature jump data could be used in combination with the thermophoresis data of later measurement time points, when the thermophoretic trend of the individual samples was congruent, producing similar binding curves. If that was not the case, either the temperature jump signal or the thermophoresis data alone had to be examined individually for the analysis.

3.1.5 Triplex formation is dependent on multivalent cations and pH

Two of the major components influencing triplex formation and stability are the concentration of multivalent cations, especially magnesium and the pH of the media or buffer (Sugimoto et al., 2001; Wu et al., 2002). To gain further insights on the effects of these parameters, their influence on the enhancer triplex was monitored using the MST technology.

Therefore the triplex formation was measured with increasing Mg^{2+}-concentrations at rising temperatures to explore the effect on complex stability. In additional experiments the identical setup was used to determine the influence of a changing pH on the thermal integrity of the complex. Both assays were conducted for RNA and DNA TFOs and the MST data was analyzed as described in (5.2.8.2).

3.1.5.1 Influence of the Mg^{2+}-concentration on triplex formation measured by MST

Table 3.2 summarizes the calculated K_d-values obtained for the DNA/RNA triplex. The dependency of third strand binding on the Mg^{2+}-concentration was clearly visible. In the reaction buffer containing only 20 mM Tris-Acetat of pH 7.4 and no other additional cations than Mg^{2+}-ions, no complex formation could be detected without and with 1 mM Mg^{2+}. Starting at 5 mM a stable triplex could be detected at

25°C. Strikingly at this temperature there seemed to be no relevant increase in complex stability with increasing Mg^{2+}-concentration. The obtained K_d constants stayed in a comparable range between 15 and 20 nM. However, with an increase in incubation temperature a clear correlation of triplex stability and Mg^{2+}-concentration was revealed. Although no differences were observed between 20 mM and 50 mM Mg^{2+}-concentration, which resembled a saturation of the stabilizing effect exhibited by the cations. Above 55°C no triplex formation was visible, indicating the melting temperature for the DNA-third strand interactions.

Mg^{2+}-dependency of DNA/RNA Triplex Formation – MST Analysis

Mg^{2+}-conc.	**25°C**	**35°C**	**40°C**	**45°C**	**50°C**	**55°C**	**60°C**
0 mM	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
1 mM	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
5 mM	19.5 +/- 6.3	99.7 +/- 21.7	265.1 +/- 31.3	772.3 +/- 85.4	n.d.	n.d.	n.d.
10 mM	11.4 +/- 9.0	28.9 +/- 12.8	65.9 +/- 16.3	182.4 +/- 49.8	636.3 +/- 135.5	n.d.	n.d.
20 mM	16.6 +/- 9.4	17.6 +/- 9.7	34.6 +/- 12.9	69.6 +/- 16.3	153.2 +/- 30.5	784 +/- 69.9	n.d.
50 mM	15 +/- 5.1	22.4 +/- 5.8	28.9 +/- 7.4	79.2 +/- 33.0	147 +/- 19.9	716.3 +/- 58.4	n.d.
K_d (nM)							

Table 3.2: K_d-constants calculated for DNA/RNA triplex formations with increasing concentrations of magnesium

Triplex formation and MST experiments were performed as described previously (5.2.8.2/Figure 3.3) using the standard triplex formation buffer with increasing Mg^{2+}-concentrations as depicted on the left. For the analysis binding curves derived from temperature jump data and combined temperature jump and thermophoresis data were prepared and K_d constants were determined. The special circumstances observed for the analysis of triplex formation by MST, presented in 3.1.4 were taken into account. n.d. = not detectable.

The data generated for the DNA/DNA triplex formation showed a similar result as was revealed for the previously described DNA/RNA triplex experiments. Again 5 mM Mg^{2+} was the minimum concentration necessary for a stable triplex to occur. However, for the DNA TFO a clear difference in K_d could be observed between 5 mM and the higher Mg^{2+}-concentrations at 25°C, which was in contrast to the DNA/RNA data. In general the DNA/DNA triplex was less stable than its DNA/RNA counterpart at increasing temperatures. Still a saturation effect for 20 mM and 50 mM Mg^{2+}-concentrations could be observed. At higher temperatures the DNA/DNA

triplex showed a decreased thermal stability, which was in contrast to the DNA/RNA triplex.

Mg^{2+}-dependency of DNA/DNA Triplex Formation – MST Analysis

Mg^{2+}-conc.	25°C	35°C	40°C	45°C	50°C	55°C	60°C
0 mM	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
1 mM	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
5 mM	65.7 +/- 10.7	343.8 +/- 23.9	1588.3 +/- 134.5	n.d.	n.d.	n.d.	n.d.
10 mM	12.3 +/- 4.0	54.5 +/- 13.5	185.8 +/- 22.8	1078.1 +/- 100	n.d.	n.d.	n.d.
20 mM	8.8 +/- 7.4	8.6 +/- 7	9.8 +/- 9.4	306.4 +/- 87.3	1224.1 +/- 209.1	n.d.	n.d.
50 mM	11 +/- 4.7	9.6 +/- 5.7	8.4 +/- 10.7	69.7 +/- 13.8	309.8 +/- 65.1	1485.9 +/- 258	n.d.
K_d (nM)							

Table 3.3: K_d constants calculated for DNA/DNA triplex formations with increasing concentrations of magnesium
The MST experiments were performed and analyzed as described for the DNA/RNA triplex in (Table 3.2) For the analysis binding curves derived from temperature jump data and combined temperature jump and thermophoresis data were prepared and K_d constants were determined. n.d. = not detectable.

3.1.5.2 The pH determines triplex formation and stability

The second major factor influencing triplex formation and stability is the pH of the medium. This was confirmed by the MST-analysis of the enhancer triplex. The results generated for the DNA/RNA triplex are summarized in (Table 3.4). Here a clear decrease in triplex stability, in correlation with an increasing pH was demonstrated. At pH 6.4 the triplex was stable over the whole temperature range, up to 60°C with K_d values between 10 and 15 nM. Even more strikingly a slight decrease in the K_d constants with increasing temperature was observed. This clearly changed upon the increase of the pH to 7.4, which resulted in a complex dissociation above 40°C. Finally at pH 8.4 no triplex formation was detectable.

DNA/RNA Triplex – pH dependency

	pH 6.4	**pH 7.4**	**pH 8.4**
20°C	14.4 +/- 4.6	51.3 +/- 8.2	n.d.
25 °C	13.8 +/- 4.6	51 +/- 9	n.d.
30°C	11.3 +/- 4	68.1 +/- 11	n.d.
35°C	10.1 +/- 3	123 +/- 19.2	n.d.
40°C	8.8 +/- 3.3	294.6 +/- 45.5	n.d.
45°C	6.8 +/- 2.6	n.d.	n.d.
50°C	6.6 +/- 3	n.d.	n.d.
55°C	5.9 +/- 3.4	n.d.	n.d.
60°C	6.3 +/- 4.5	n.d.	n.d.
	K_d (nM)		

Table 3.4: Kd-constants calculated for DNA/RNA triplex formation with increasing pH
Triplex formation and MST experiments were performed as described for (Table 3.2) using the standard triplex formation buffer with 10 mM MgAcetat, supplemented with 25 mM NaCl. The pH for each buffer is depicted in the upper row. For the analysis binding curves derived from temperature jump data were prepared and K_d constants in nM were determined. n.d.=not detectable.

The DNA/DNA triplex exhibited a similar behavior. Although the complex showed even smaller K_d constants at pH 6.4 than the DNA/RNA triplex, however at 60°C a drastic increase in the K_d was observed. The behavior at pH 7.4 was also comparable to the DNA/RNA triplex, but the K_d values were again slightly lower for the DNA/DNA complex.

DNA/DNA Triplex – pH dependency

	pH 6.4	pH 7.4	pH 8.4
20°C	5 +/- 1.9	20.6 +/- 6.8	n.d.
25 °C	4 +/- 1.9	20.7 +/- 6.5	n.d.
30°C	3 +/- 1.8	33.1 +/- 10.3	n.d.
35°C	2.9 +/- 1.8	79.7 +/- 23.1	n.d.
40°C	2.3 +/- 2.2	229.2 +/- 83.7	n.d.
45°C	3.6 +/- 3.1	n.d.	n.d.
50°C	5.7 +/- 3	n.d.	n.d.
55°C	17.4 +/- 5.3	n.d.	n.d.
60°C	100 +/- 17.8	n.d.	n.d.
	K_d (nM)		

Table 3.5: Kd-constants calculated for DNA/DNA triplex formation with increasing pH
Triplex formation and MST experiments were performed as described previously using the standard triplex formation buffer with 10 mM MgAcetat, supplemented with 25 mM NaCl. The pH for each buffer is depicted in the upper row. For the analysis binding curves derived from temperature jump data were prepared and K_d constants in nM were determined. n.d.=not detectable.

In summary the MST experiments presented here provided further insights on the influence of pH and magnesium concentration on triplex formation.

3.1.6 Enhancer triplex formation occurs in near physiological buffer conditions

The analysis of triplex formation on the DNA sequences derived from the rDNA enhancer region, demonstrated a high triplex forming potential of this DNA motifs under *in vitro* conditions. However, the buffer conditions used in theses analysis hardly represented the situation inside a cell and therefore only provided limited information on a potential occurrence of these complexes *in vivo*. In order to get further insights on the formation of triplex structures under physiological conditions, the MST method was used in combination with a buffer system more closely matching the situation inside a cell.

Thus a simplified cell buffer was prepared, resembling the basic ionic concentrations typically found in a mammalian cell (Alberts, 2002). Special

attention was given to the concentration of multivalent cations, since previous studies showed their role as major regulator in triplex formation, which was also underlined by the results presented in this work. However, for the cell buffer system only Mg^{2+} was considered, since the concentration of other cations of this class, like Ca^{2+}, is very small inside a cell. The experiments were carried out with three different concentrations of Mg^{2+}, starting at 0.5 mM, representing the typical amount of free magnesium inside a cell. In addition 5 mM and 10 mM of Mg^{2+} were used, more closely matching the overall magnesium concentration found in complex with proteins and nucleic acids inside a cell (Alberts, 2002). The reactions were set up as described previously and all measurements were done at an incubation temperature of 37°C, resembling physiological conditions. For the analysis temperature jump and thermophoresis signals were normalized to the fraction bound and plotted against the TFO concentration. Again binding curves and K_d-values for each experiment were calculated. As shown in (Figure 3.7) no binding was detected at 0.5 mM Mg^{2+}-concentrations, therefore the data was normalized to the 10 mM data.

The results illustrated in (Figure 3.7) again showed a dependency of the triplex formation on the concentration of Mg^{2+}-ions. An interaction between DNA and RNA TFO was already detectable at 5 mM magnesium. At a concentration of 10 mM magnesium a clear triplex formation could be shown, resulting in a K_d of 276 +/- 24 nM (Figure 3.7A). For the DNA/DNA triplex formation the experiments revealed a slightly different behavior. As depicted in (Figure 3.7B) at 10 mM magnesium a triplex formation could be detected and a K_d of 602 +/- 47 nM was calculated, representing a more than 2-fold increase compared to the DNA/RNA triplex. Also in contrast to the DNA/RNA complex, the DNA showed no interaction with a DNA TFO at 5 mM Mg^{2+}. Finally at 0.5 mM Mg^{2+} whether the DNA/RNA neither the DNA/DNA samples exhibited any sign of triplex formation in the MST analysis.

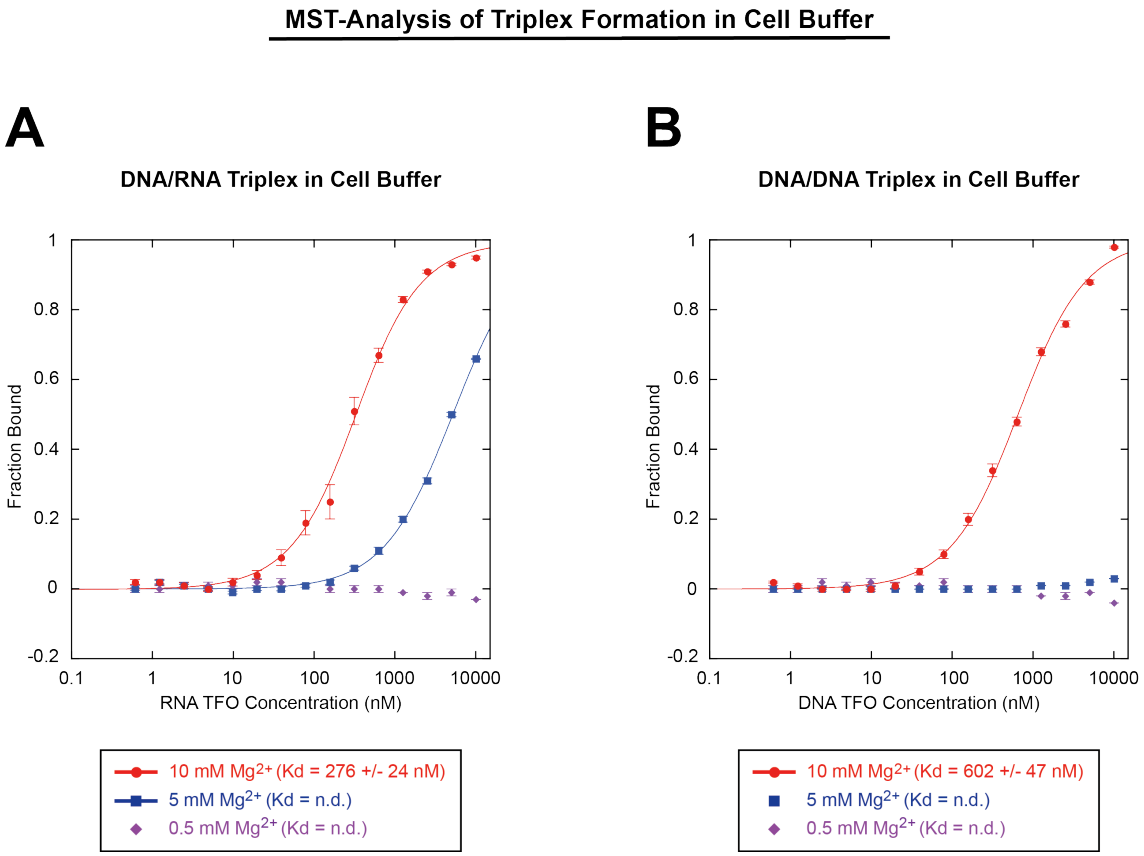


Figure 3.7: Triplex formation in near physiological buffer conditions
Triplex formation for DNA/RNA TFO (**A**) and DNA/DNA TFO (**B**) was measured by MST-analysis in near physiological buffer conditions (5.1.4/5.2.8.2). MST settings were chosen as described previously and the incubation temperature was set to 37°C for all reactions. The normalized results (fraction bound) are plotted against the TFO concentrations in nM on a logarithmic scale. Red data points represent 10 mM Mg^{2+}, blue 5 mM Mg^{2+} and purple 0.5 mM Mg^{2+}-concentration. The error bars indicate standard deviations of four technical replicates.

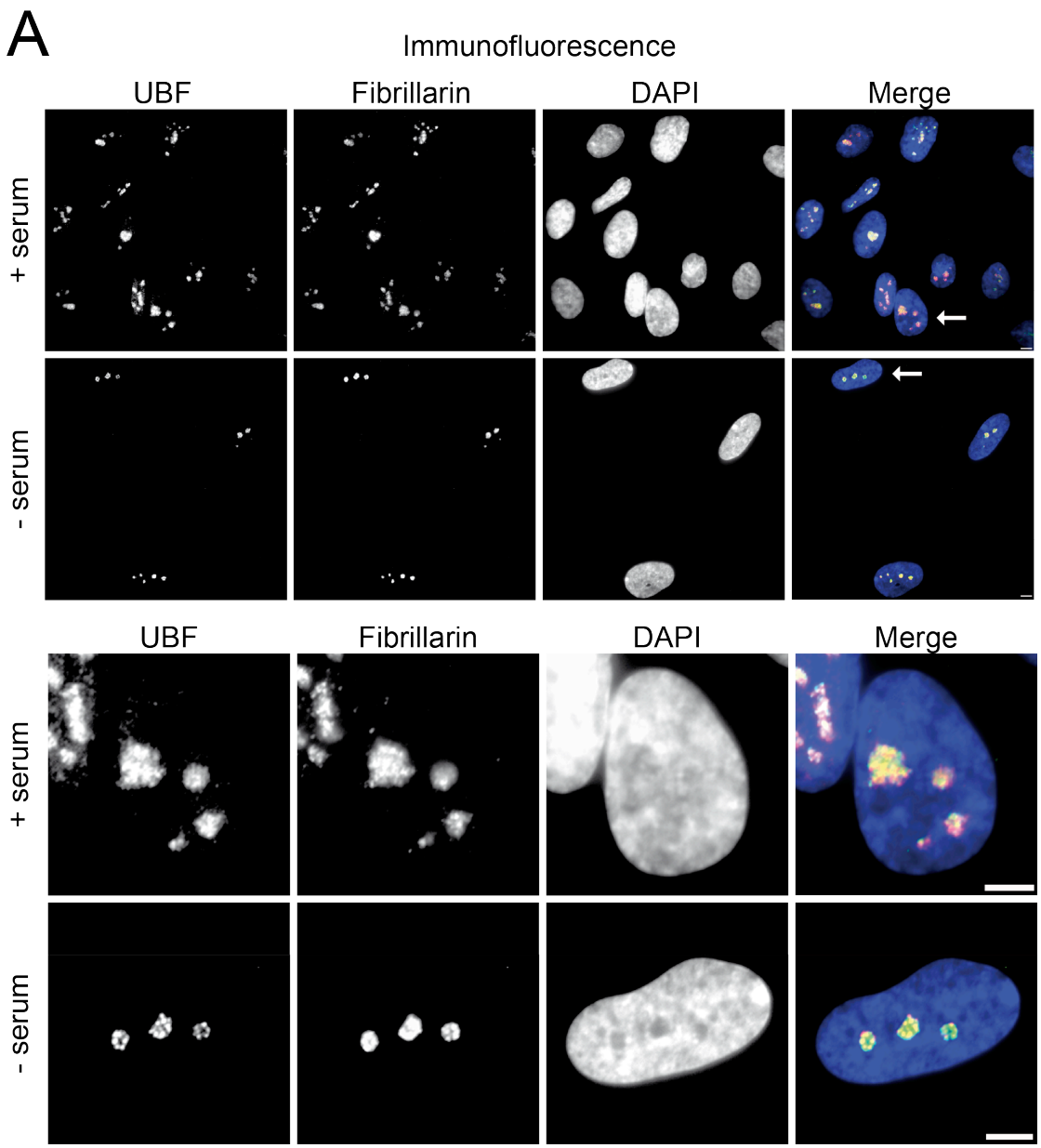
3.2 Large-scale organization of ribosomal DNA chromatin is regulated by Tip5

The results presented in this chapter have been published recently (Zillner et al., 2013). To ensure a good understanding of the complete story, additional data produced by the co-authors of the manuscript is presented briefly at some occasions. However, if that is the case the individual contributor is clearly highlighted and only a summary of the obtained results is given. The figures illustrating theses datasets are cited and can be found in the complete publication.

3.2.1 Serum starvation induces global changes in nucleolar architecture and an enrichment of rDNA in the nuclear matrix

To gain a better understanding of the spatial organization of silenced rDNA chromatin within the nucleolus and a possible role for Tip5 in these processes, a combination of serum starvation and immunofluorescence analysis was used to track changes in nucleolar structure, which correlate to the repression of rRNA synthesis (described in 2.4.5).

Therefore the distributions of the nucleolar marker proteins UBF, fibrillarin and RNA polymerase I (PolI) were compared in serum starved and normally growing IMR90 human embryonic lung fibroblast cells (Figure 3.8 A/B).


A
Immunofluorescence
UBF
Fibrillarin
DAPI
Merge
+ serum
- serum
UBF
Fibrillarin
DAPI
Merge
+ serum
- serum

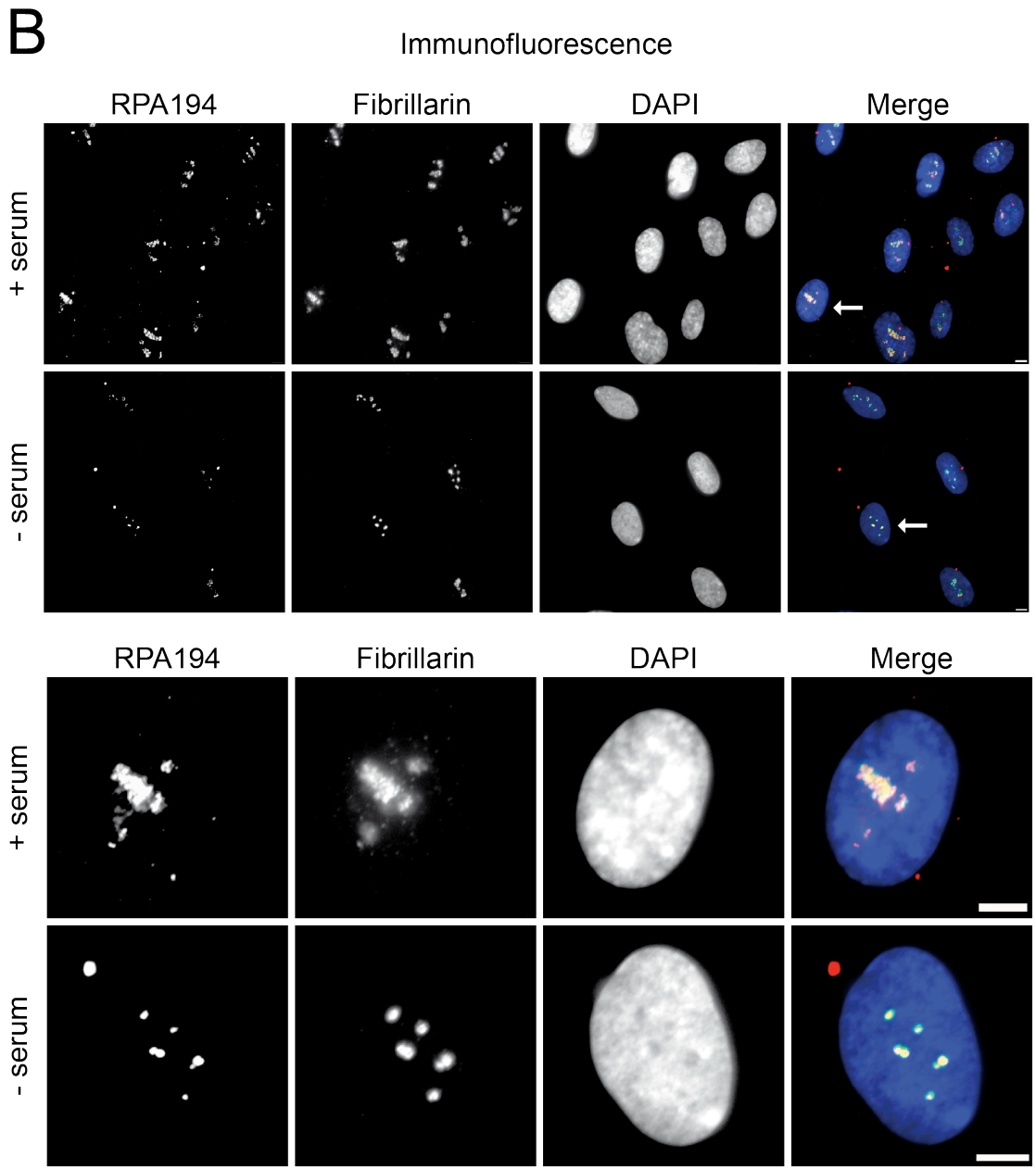


Figure 3.8: Distribution of UBF, fibrillarin and PolI (RPA194) in naturally growing and serum starved IMR90 human embryonic lung fibroblasts

A) Immunofluorescence detection of the nucleolar proteins UBF and fibrillarin in control and serum starved IMR90 cells. Nuclei were counterstained with DAPI. The lower panel shows a close up of the cells marked by an arrow in the upper panel. Bars represent 5 µm.

B) Immunofluorescence detection of the nucleolar proteins PolI (using a monoclonal mouse antibody against the RPA 194 subunit of PolI) and fibrillarin in control and serum starved IMR90 cells. Nuclei were counterstained with DAPI. The lower panel shows a close up of the cells marked by an arrow in the upper panel. Bars represent 5 µm.

Upon serum starvation a clear decrease in nucleolar size and a focal compaction of the UBF, PolI, and fibrillarin signals within the nucleoli could be detected. These observations together with similar results published in previous reports (O'Mahony et al., 1992; Seither et al., 1997; Yamamoto et al., 2004) lead to the

conclusion that the spatial organization of rDNA chromatin in the nucleoli changes upon repression of rRNA synthesis.

In order to get further insights in this reorganization process, the amount of various rDNA fragments in nuclear matrix preparations from control and serum starved IMR90 cells was measured using quantitative real-time PCR. This was compared to the level of INFβ promoter, which has been shown to be a *bona fida* MAR (Bode et al., 1992), stably associated with the nuclear matrix and in addition contains a well-characterized binding site for the AT-hook protein HMGA1 (Huth et al., 1997). This finding led to the assumption, that changes in the relative amount of rDNA content, compared to this specific MAR should represent alterations in the nuclear matrix association of these rDNA regions.

Therefore the putative MARs of the human rDNA were determined *in silico* using a published web-tool (Kramer et al., 1996). All of them localized to the IGS region of the rDNA repeat. In addition two other regions, which are no predicted MARs were also chosen for quantification. One of these sites is located within the rDNA promoter, which is a well-characterized binding site of Tip5. The protein comprises four AT-hooks and a TAM domain and thus potentially targets its binding sites to the nuclear matrix. The third region locates to the rDNA coding region (28S), where no Tip5 binding occurs (Santoro et al., 2002). qPCR primer pairs were designed for all regions of interest, to determine their amount in the different nuclear matrix fractions as illustrated in (Figure 3.9). This setup allowed a monitoring of MAR- and Tip5-dependent and –independent associations of rDNA sequences with the nuclear matrix.

Nuclear matrix template DNA from normal and serum starved IMR90 cells was purified and similar amounts were objected to analysis by quantitative real-time PCR reactions. Threshold cycle (C_t) differences between serum starved and control cells were determined at each of the three different regions of the human rDNA and normalized to the C_t differences of the IFNβ promoter. Figure 3.9 displays the results of three biological replicates and illustrates accumulation of the three tested rDNA regions in the nuclear matrix fraction upon serum starvation. Clear differences in enrichment for each region could be detected. The IGS sequence showed an increase of approximately 1.5-2-fold in the nuclear matrix compared to the INFβ promoter, for the coding region and the promoter an enrichment of 2-5-

fold and 5-10-fold could be determined, respectively. Therefore the experiments showed a clear effect of the serum starvation on the rDNA content in the nuclear matrix, which suggested a recruitment of silenced repeats to the nuclear matrix.

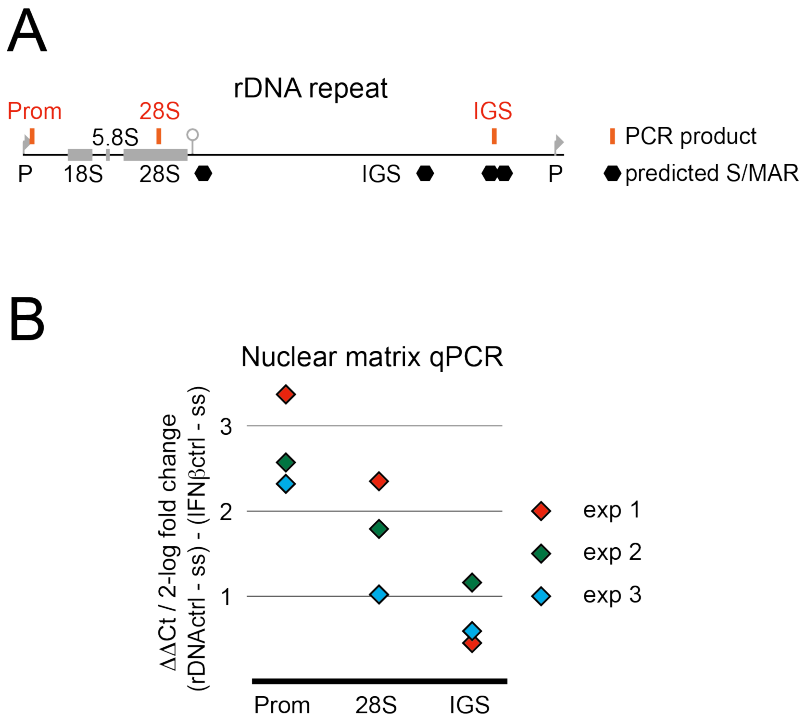


Figure 3.9: Real-time qPCR analysis of serum starvation-induced changes in the nuclear matrix association of rDNA

A) The scheme displays the organization of a human rDNA repeat with promoter, 18S, 5.8S and 28S coding region, and the IGS region (GenBank AccNo: U13369). In red are the locations of the sequences used in qPCR analysis. The black hexagons mark the predicted MARs.

B) Identical amounts of DNA from the nuclear matrix fractions obtained from normal and serum starved IMR90 cells were subjected to qPCR analysis. The results of three biological replicates are shown. Differences in C_t values between normal and serum starved cells were determined for the three different regions and normalized to the C_t values of the INFβ promoter. The different colors represent the biological replicates.

3.2.2 Tip5 is associated with the nuclear matrix

Due to the implications derived from the serum starvation experiments and the fact that Tip5 contains several predicted MAR binding domains (Aravind and Landsman, 1998), its potential to target rDNA to the nuclear matrix was examined next.

Therefore the sub-cellular localization of Tip5 in HeLa cervix carcinoma cells was analyzed using immunofluorescence. The results showed that the protein

predominantly, but not exclusively localizes to the nucleoli, which were marked by B23 immunostaining (Figure 3.10).

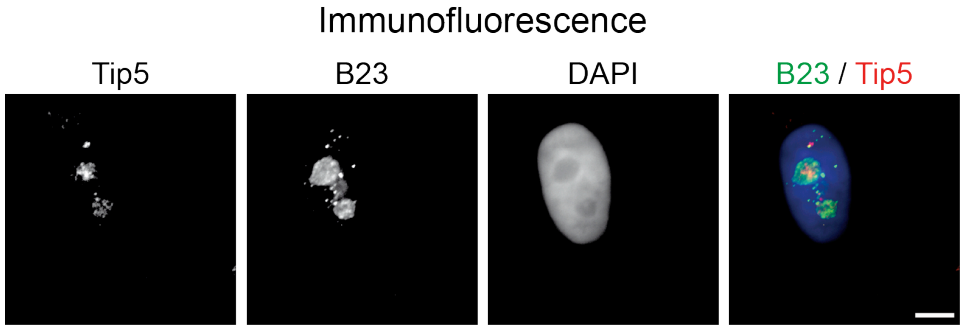


Figure 3.10: Sub-cellular localization of Tip5 in HeLa cells analyzed by immunofluorescence
HeLa cells were fixed in 4% PFA in PBS and immunostained with a rabbit monoclonal antibody against Tip5 and a mouse monoclonal antibody against B23. DNA was counterstained with DAPI. The image on the right shows the merge with Tip5 signals in red, the B23 localization in green and the DNA in blue. The bar represents 5 µm.

Furthermore, *in situ* preparations of nuclear matrix fractions combined with immunofluorecence analysis were used to determine a possible association of Tip5 with the nuclear matrix. These experiments were performed in HeLa cells and clearly displayed the presence of Tip5 within the matrix after extensive DNAse I digestion and chromatin extraction (Figure 3.11). This was similar to lamin A/C, which represents a key component of the nuclear matrix and functioned as a positive control for the preparation procedure.

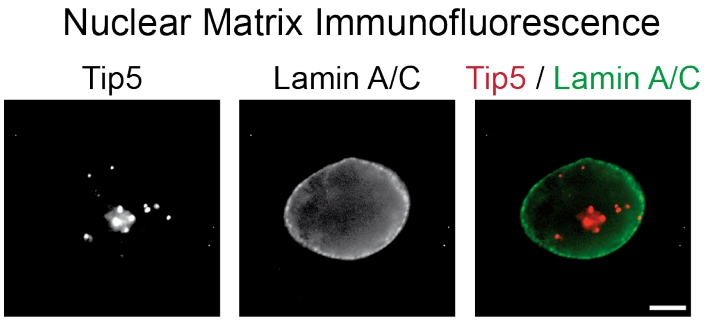


Figure 3.11: Association of Tip5 to the nuclear matrix analyzed by *in situ* matrix preparation and immunofluorescence
HeLa cells were subjected to matrix preparation *in situ*. The remaining nuclear matrix fraction was stained with a rabbit monoclonal antibody for Tip5 and a mouse monoclonal lamin A/C antibody. The picture on the right shows the merge with Tip5 in red and lamin A/C in green. The bar represents 5 µm.

In addition the presence of Tip5 in the nuclear matrix was also verified by immunblot analysis of whole cell lysates from HEK293 human embryonic kidney fibroblast cells. The samples were fractionated into cytoplasmic, soluble chromatin, high-salt wash and nuclear matrix fractions and submitted to immunoblot experiments ((Zillner et al., 2013) Fig.: 2B).

The results clearly depicted two pools of Tip5 co-existing in the cells. The protein could also be found in the soluble chromatin fraction. However the majority of Tip5 is associated with the nuclear matrix. This was in clear contrast to other remodeler subunits like Brg-1, Snf2h and Mi-2, which appeared preferentially in the chromatin fraction. Additionally the distribution of PolI in the different fractions clearly demonstrated that not all nucleolar transcription factors were concentrated within the nuclear matrix. The next point was to rule out an involvement of the recently discovered RNA binding activity of Tip5 (Mayer et al., 2006) in nuclear matrix association. Therefore matrix reparations in the presence of RNAse A were performed to test for RNA-dependent binding ((Zillner et al., 2013) Fig.: S2). The results showed no change in Tip5 localization to the matrix, however the association of the protein with the soluble chromatin fraction seemed to be sensitive to RNAse A treatment. This could indicate a possible role of regulatory RNAs in Tip5 recruitment to chromatin.

(Karina Zillner and Attila Nemeth performed the immunblot and RNAse A experiments described here.)

3.2.3 The nuclear matrix association of Tip5 shows no changes upon serum starvation

The next question was, whether the association of Tip5 also changes in a serum starvation dependent fashion, comparable to the rDNA content. Therefore immunoblot experiments were performed on the different fractions generated in nuclear matrix preparations from control and serum starved IMR90 embryonic lung fibroblast cells.

The results illustrated that there is no detectable change in the amount of Tip5 in the nuclear matrix upon serum starvation. In addition most of the protein seemed

to stay associated with this fraction (Figure 3.12). However the immunoblot analysis depicted represents only a partially quantitative analysis and can therefore not completely rule out an effect of the serum starvation on Tip5.

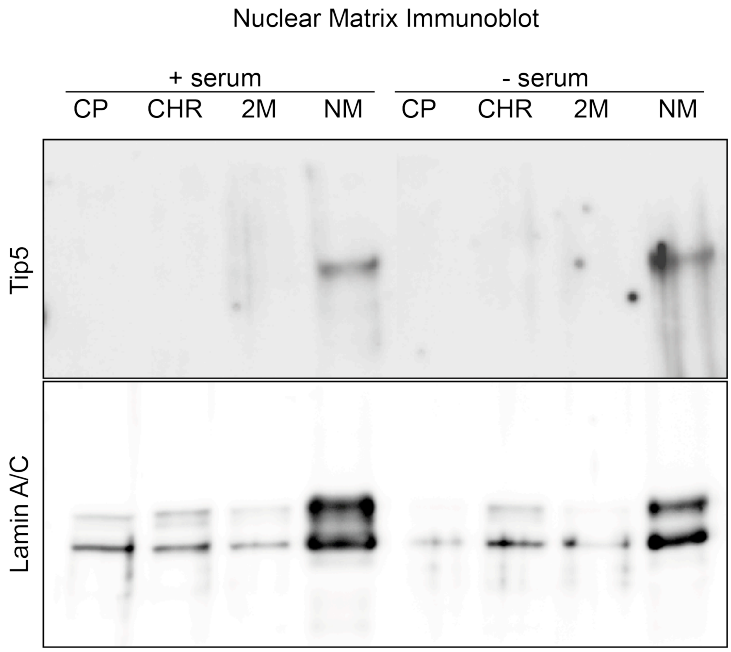


Figure 3.12: Immunoblot analysis of nuclear matrix fractions from control and serum starved IMR90 cells
The upper panel shows a Tip5 specific immunoblot of cell fractions from control (left) and serum starved cells (right). In the lower panel the distribution of lamin A/C throughout the fractions is depicted, which served as positive control for the preparations. (CP=cytoplasmic, CHR=soluble chromatin, 2M=high-salt wash, NM=nuclear matrix fraction)

3.2.4 Tip5 targets rDNA to the nuclear matrix

Taken together the facts that Tip5 not only contains several DNA binding motifs, which are also putative MAR binding domains, but also showed a predominant and stable association with the nuclear matrix, suggested a possible involvement of the protein in targeting of rDNA to the nuclear matrix.

To further verify this hypothesis, the rDNA content of nuclear matrix fractions from HEK293 cells, which were transfected with Tip5 or control constructs respectively, was investigated as described in 3.2.1. A strong overexpression of Tip5 was detected by immunoblot analysis 72 h post transfection ((Zillner et al., 2013) Fig.: 2D). The qPCR analysis revealed an enrichment of all three rDNA regions in the nuclear matrix following Tip5 overexpression, indicating a direct or

indirect role of the protein in rDNA recruitment to this fraction. The amount of IGS, coding region and promoter sequences increased approximately 2-8-fold in the matrix fraction compared to the IFNb MAR control. There was only a slight difference between the matrix association levels of different rDNA regions within the individual biological replicate experiments ((Zillner et al., 2013) Fig. 2E, Experiments were performed by Karina Zillner). The increase of all sites in the nuclear matrix could either suggest a binding of Tip5 to all investigated regions, or it could implicate that the recruitment of the Tip5 binding sites nucleates a further recruitment of the whole repeat to the nuclear matrix.

3.2.5 The potential MAR binding domains of Tip5 show a preference for AT-rich DNA

Besides the tandem PHD-bromodomain involved in protein-protein interactions, Tip5 contains a variety of nucleic acid binding motifs, like AT-hooks and a TAM domain, which were described as putative MAR binding elements (Aravind and Landsman, 1998). While the TAM domain has already been characterized previously (Strohner et al., 2001), the properties of the AT-hook domains have not been analyzed so far. Since the AT-hook domain has already been described as putative MAR binding motif, these domains present premier target motifs for being involved in a possible recruitment of rDNA to the nuclear matrix (Aravind and Landsman, 1998). Thus a profound *in vitro* analysis, using each individual AT-hook as well as a combination of the AT-hooks 1 and 2 in form of GST purified peptides, in EMSA and quantitative MST assays was performed ((Zillner et al., 2013) Fig.3A/3B/4A, performed by Karina Zillner, Michael Weinberger and Kathrin Rachow).

The experiments confirmed each AT-hook to represent a *bona fida* DNA binding motif, which is able to bind to different sequences with similar affinities. Furthermore it was possible to identify the combination of the AT-hooks 1+2 as the strongest binder, displaying a similar affinity like the HMGA1, which was used as a positive control. Additionally the AT-hooks showed a clear preference for AT-rich DNA sequences, derived from the rDNA enhancer region, over a GC-rich

control DNA (Figure 3.13). The EMSAs were performed under identical experimental conditions and displayed a preferred binding of the AT-rich target for the AT2 peptide, as well as for the double AT-hook 1+2 (Figure 3.13A). Looking at the data for the AT2 peptide a full binding of the AT-rich DNA could be detected starting in lane 8, whereas the GC-rich target is not completely shifted until lane 10. A similar observation could be made for the AT1+2 peptide, where the shift of the DNA started in lane 6 for the AT-rich DNA and in lane 8 for the GC-rich DNA, respectively. For the MST analysis a competitive binding assay was established, using equimolar amounts of Cy5-labeled AT-rich rDNA and Cy3-labeled GC-control DNA sequences in parallel. The EC50 values for AT2 and AT1+2 were again determined under these conditions, which lead to a similar result like seen in the EMSA experiments, confirming a preferential binding of the AT-rich DNA (Figure 3.13B).

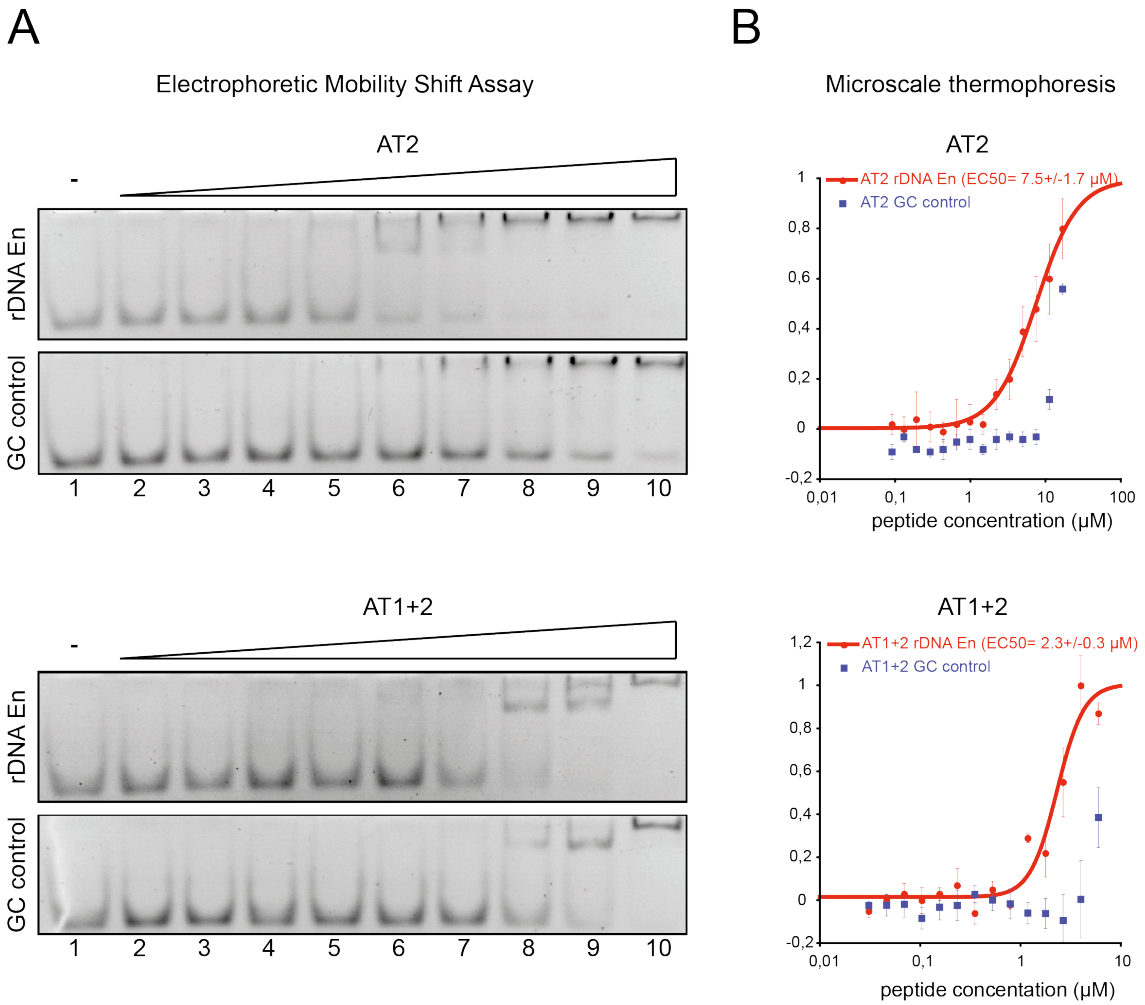


Figure 3.13: The AT-hooks AT2 and AT1+2 of Tip5 preferentially bind AT-rich DNA
A) EMSA showing the binding of the peptides AT2 (upper picture) and AT1+2 (lower picture) to an AT-rich rDNA sequence and a GC-rich control DNA in identical experimental conditions. 1.25 pmol of DNA were incubated with increasing amounts of protein (9 pmol, 14 pmol, 22 pmol, 33 pmol, 50 pmol, 76 pmol, 115 pmol, 174 pmol, 264 pmol) and loaded on a 7.5 % native PAA gel. Gels were stained with EtBr.
B) Competitive MST assays with the same peptides and DNA sequences used in A). The signals of the weaker bound GC-rich DNA were normalized to the fraction bound data of the AT-rich DNA.

3.2.6 The double AT-hook domain is a nucleolar targeting module

Since the previous experiments identified the double AT-hook as the strongest putative MAR binder the next step was to test the nuclear matrix association of this protein domain in transient transfections. GFP fusions of a wild type and a mutant version of the double AT-hook domain were generated. In the latter, the RGR core motif of both AT-hooks was changed to a DGD tripeptide, which has been reported

to cause a loss in DNA binding activity of this domain (Bourachot et al., 1999). The resulting GFP-AT1+2-WT and GFP-AT1+2-Mut constructs were transfected into HeLa cells and the sub-cellular localization of the proteins was determined by immunofluorescence.

The results in (Figure 3.14) illustrate the predominant localization of the GFP-AT1+2-WT proteins to the nucleoli. In contrast, the GFP-AT1+2-Mut transfections displayed an even distribution throughout the nucleus. Thus the experiments clearly identified the AT-hooks as a nucleolar targeting module for the GFP proteins.

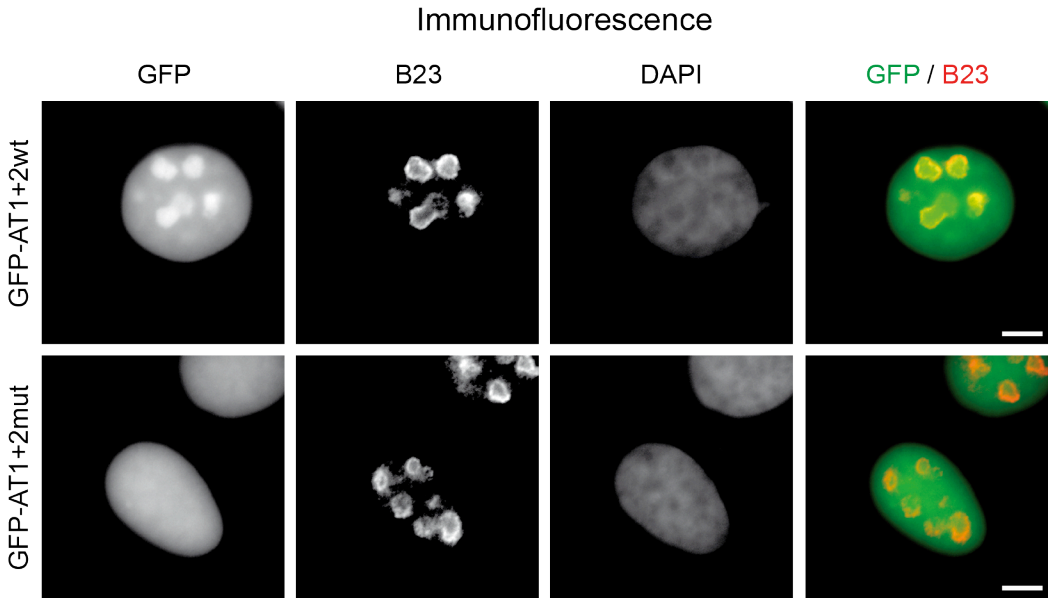


Figure 3.14: Sub-cellular localization of GFP-AT1+2-WT and GFP-AT1+2-Mut fusion proteins
HeLa cells were transfected and fixed for immunfluorescence 24 h post transfection. GFP signals were detected using a monoclonal rabbit GFP antibody and nucleoli were marked by B23 staining with a mouse monoclonal antibody. DNA was counterstained with DAPI. The upper panel displays the results from the GFP-AT1+2-WT transfections in the lower panel the GFP-AT1+2-Mut experiments are shown. The bars represent 5 µm.

3.2.7 The double AT-hook domain is not sufficient for nuclear matrix targeting

Surprisingly, when performing nuclear matrix preparations on transfected cells and analyzing the sub-cellular localization of GFP-AT1+2-WT and GFP-AT1+2-Mut by immunoblot (Figure 3.15A) and by immunofluorescence (Figure 3.15B) it occurred, that despite the *in vitro* MAR binding activity and nucleolar targeting, the

double AT-hook domain was not sufficient to mediate an association with the nuclear matrix. In both cases, no GFP-Tip5 proteins could be detected in the nuclear matrix fractions.

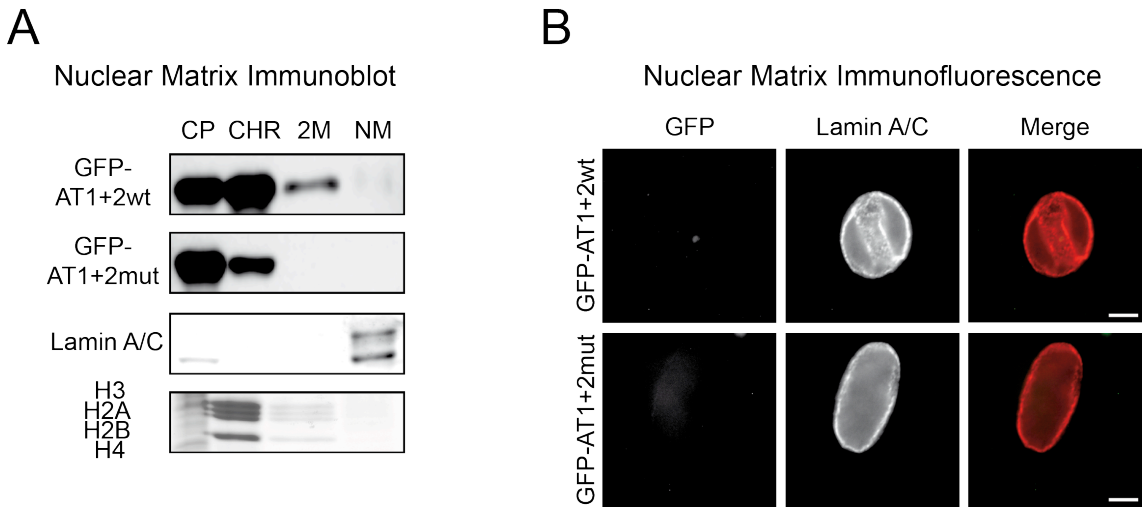


Figure 3.15: The double AT-hook domain is not sufficient for nuclear matrix targeting
A) Immunoblot and coomassie gel pictures of nuclear matrix preparations. HeLa cells were transfected with the indicated GFP-Tip5 proteins for 72 h and nuclear matrices were prepared. CP, CHR, 2M and NM indicate cytoplasmic, soluble chromatin, high-salt wash, and nuclear matrix fractions, respectively. Core histones and lamin A/C served as controls for the CHR/2M and NM fractions, respectively.
B) Cells were transfected with the indicated GFP constructs. After 48 h nuclear matrix preparations were performed *in situ* and analyzed by immunofluorescence. Lamin A/C served as positive control for the preparations. Bars indicate 5 µm.

3.2.8 The TAM domain mediates nuclear matrix association and nucleolar targeting

In order to test whether the other putative MAR binding domain present in Tip5, the TAM domain, is required to mediate the association to the nuclear matrix, new constructs including this motif were prepared. Thus the GFP fusions were extended with the TAM motif resulting in GFP-TAM-AT1+2-WT and GFP-TAM-AT1+2-Mut proteins. Again transfection experiments in HeLa cells were performed and the sub-cellular localization of the GFP-Tip5 proteins in fixed cells, as well as in *in situ* matrix preparations was determined using immunofluorescence. These experiments were complemented by immunoblots of nuclear matrix preparations derived from whole cell extracts of the transfected cells (Figure 3.16/Figure 3.17).

The immunofluorescence experiments showed a different sub-nuclear localization of the two proteins in fixed cells (Figure 3.16). The GFP-TAM-AT1+2-Mut proteins predominantly accumulated in the nucleoli, illustrating that the TAM domain is sufficient for nucleolar targeting, similar to the AT1+2 domain. However, the situation for the GFP-TAM-AT1+2-WT proteins was a bit different. Here the protein was rather enriched in perinucleolar chromatin and only a slight proportion entered the nucleolus. Additionally higher protein levels could be detected in the nucleoplasma.

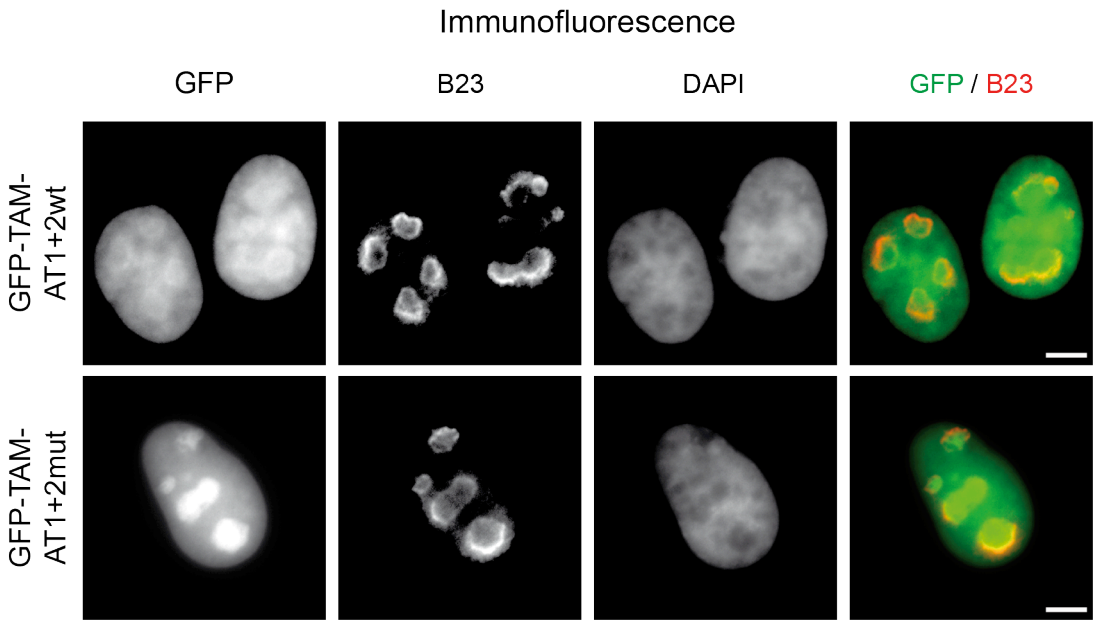


Figure 3.16: Immunofluorescence analysis of the sub-cellular localization of the GFP-TAM-AT1+2 proteins

HeLa cells were transfected with the proteins indicated on the left. 48 h post transfection, cells were fixed and the localization of the GFP-Tip5 proteins was determined by immunostaining with a commercial GFP-booster. The nucleoli were marked by immunostaining with a mouse monoclonal B23 antibody. DNA was visualized with DAPI. The bars indicate 5 µm.

Interestingly, the results derived from the *in situ* nuclear matrix preparations showed that the TAM domain seems to be necessary and sufficient for association with the nuclear matrix fraction. The experiments clearly demonstrated the presence of GFP-TAM-AT1+2 proteins in the nuclear matrix no matter if they were combined with a functional AT1+2 domain or not (Figure 3.17A). These observations could also be verified by the immunoblot analysis (Figure 3.17B), where GFP was found in the nuclear matrix fractions of both transfections.

Overexpression of the GFP-AT1+2-WT and GFP-AT1+2-Mut and the GFP-TAM-AT1+2-WT and GFP-TAM-AT1+2-Mut proteins had no effect on the rDNA content in the nuclear matrix fraction. This was observed again by nuclear matrix qPCR analysis performed by Karina Zillner ((Zillner et al., 2013) Fig.: S5).

The results suggested a combined role for the AT-hooks and the TAM domain in nucleolar and nuclear matrix targeting of Tip5. Furthermore they revealed that additional elements of Tip5 are needed for rDNA targeting to the nuclear matrix since overexpression of the individual domains had no effect on the amount of rDNA in the nuclear matrix.

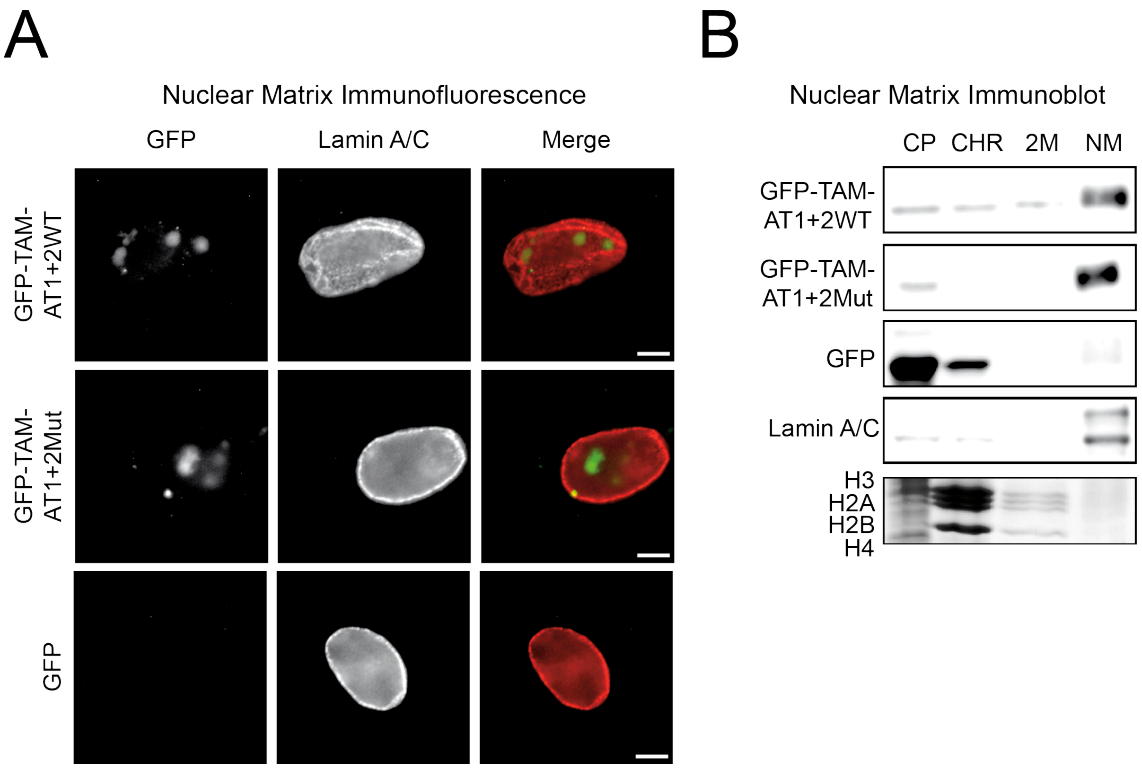


Figure 3.17: Nuclear matrix binding of GFP-TAM-AT1+2 proteins
A) HeLa cells were transfected with the GFP-Tip5 proteins or a GFP control as labeled on the left. After 48 h *in situ* nuclear matrix preparations were performed and the cells were again immunostained with the specific antibodies for lamin A/C and a GFP-booster as previously stated. Bars indicate 5 µm.
B) HeLa cells were treated as described for A) except the analysis was performed 72 h post transfection. Again immunoblot and coomassie gel pictures are shown. With CP, CHR, 2M and NM, representing the cytoplasmic, soluble chromatin, high-salt wash, and nuclear matrix fractions. Distributions of lamin A/C and the histones served as positive controls for the preparation protocol.

3.3 The extended AT-hook (eAT) is a novel nucleic acid binding domain

3.3.1 Identification of a new nucleic acid binding domain in human and mouse proteomes

In addition to the canonical AT-hooks found in Tip5 the protein contains another GRP tripeptide motif, which was not described as AT-hook domain according to (Aravind and Landsman, 1998). However patches of basic amino acids in the vicinity of the core motif are present in the protein sequence, which lead to the speculation of this GRP tripeptide being the core motif of a new kind of AT-hook like DNA binding domain. The putative motif was therefore termed extended AT-hook (eAT-hook) referring to the patches of basic amino acids located distant from the core tripeptide.

EMSA and MST experiments with different GST-eAT-hook proteins validated the domain as a new functional DNA binding motif. Furthermore, a bioinformatic search for $(K/R)_{>3}$-N_{6-20}-GRP-N_{6-20}-$(K/R)_{>3}$, describing the new domain, revealed 80 and 61 proteins containing an eAT-hook in the human and mouse proteome (Zillner, 2013). Therefore these results illustrated the widespread distribution of this new DNA binding domain suggesting a similar function to the canonical AT-hooks. To further validate these results and in order to get better insights into the biological functions of this motif, additional eAT-hook candidate domains were selected for further investigations.

3.3.2 The eAT-hook domains of PTOV1 and GPBP1 show DNA binding properties comparable to classic AT-hook motifs

After identifying the new eAT-hook domain as a functional DNA binding motif in Tip5 and showing its widespread distribution in terms of sequence motifs throughout the human and mouse proteome, the idea was to test the eAT-hook domains found in other proteins on their functionality in a similar fashion as performed for the Tip5 eAT-hook motif. Therefore candidate proteins were

selected regarding their putative involvement in nucleic acid metabolism, regulation, or interaction. At the same time the candidates should not contain any described nucleic acid interaction domains. Thus the proteins PTOV1 (Prostate Tumor Overexpressed 1) and GPBP1 (GC-rich Promoter Binding Protein 1) were selected for further investigations. GPBP1 was described as GC-rich promoter binding protein, which is able to induce transcription on TATA-box deficient promoters (Hsu et al., 2003). PTOV1 was initially reported as factor overexpressed in prostate carcinoma cells. In addition, it has been suggested to have some regulatory function in cell cycle progression and was observed to shuttle between the cytoplasm and the nucleus (Benedit et al., 2001). However the exact molecular function of PTOV1 and GPBP1 was so far not revealed. Since both proteins were rather uncharacterized, yet implicated to have potentially interesting functions and fulfilled the criteria of having no described nuclear interaction domain they were chosen for further investigations of eAT-hook function.

A set of 3 different GST-tagged proteins for each eAT-hook domain of the two different candidate proteins was generated, resulting in the proteins termed eAT, eAT Mut Core and eAT Core. Thereby the eAT protein comprises the full-length motif. In the eAT Mut Core constructs the RGRP and GRP core motifs of the two different domains were mutated to DGDP and GDP respectively, which again should interfere with the DNA binding, as described previously (Bourachot et al., 1999). The eAT Core finally had N- and C-terminal deletions, removing the basic amino acids patches outside the core motif. An overview of the amino acid sequences of each eAT-hook variant is illustrated in (Figure 3.18).

PTOV1 eAT-hook Peptides Amino Acid Sequences

```
eAT           MVRPRRAPYRSGAGGPLGGRGRPPRPLVVRAVRSRSWPASPRG
eAT Mut Core  MVRPRRAPYRSGAGGPLGGDGDPPRPLVVRAVRSRSWPASPRG
eAT Core                    LGGRGRPPRPLVVR
```

GPBP1 eAT-hook Peptides Amino Acid Sequences

```
eAT           NRYDVNRRRHNSSDGFDSAIGRPNGGNFGRKEKNGWRTHGRNG
eAT Mut Core  NRYDVNRRRHNSSDGFDSAIGDPNGGNFGRKEKNGWRTHGRNG
eAT Core                    DSAIGRPNGGNFGR
```

Figure 3.18: PTOV1 and GPBP1 eAT-hook protein variants
The amino acid sequences of the eAT-hook domains of PTOV1 (upper panel) and GPBP1 (lower panel), which were used to generate the GST-fusion proteins, are depicted. In red are the amino acids making up the eAT-hook domains with the GRP core motifs and the additional basic amino acids in the N-and C-terminal peripheries. The mutations introduced for the Mut Core variants are highlighted in green.

The proteins were expressed in *E.coli* and purified via the GST-tag. A coomassie gel showing the purified proteins is displayed in (Figure 3.19). The purified proteins were used for the subsequent investigations of this new nucleic acid binding motif.

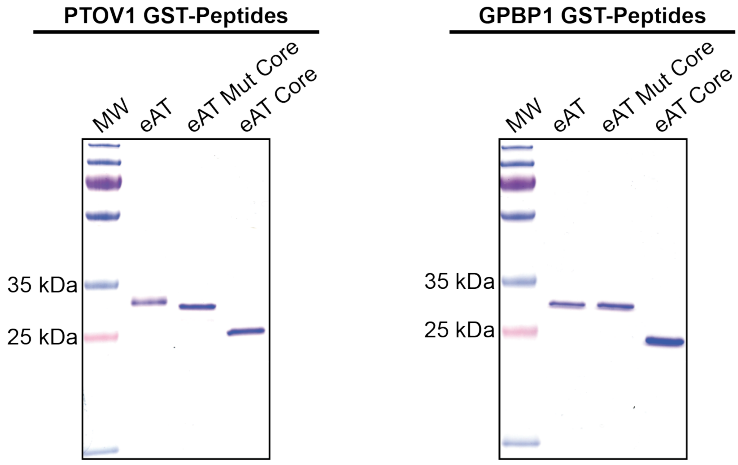


Figure 3.19: GST-purifications of the PTOV1 and GPBP1 eAT-hook domain variants
The different eAT-hook protein variants were expressed in *E.coli* and purified via the fused GST-tag. 0.5 µg of each protein purification was loaded on a 15% SDS gel and stained with coomassie. The left gel shows the different PTOV1 eAT-hook variants. On the right are the GPBP1 eAT-hook proteins. The nomenclature is as illustrated in (Figure 3.18).

Since a complete set of eAT-hook domain variants from the two different candidate proteins was now available, the first thing to test was their DNA binding activity. Hence the same AT-rich rDNA IGS sequence used for the *in vitro* analysis of the Tip5 AT-hook motifs was employed for the investigation of the different eAT-hook protein variants DNA binding behavior in EMSA (Figure 3.20). The different eAT-

hook proteins were incubated with the AT-rich rDNA sequence under identical experimental conditions and the DNA binding properties of the different GST-eAT proteins was monitored on native polyacrylamide gels.

The results for the PTOV1 proteins displayed in (Figure 3.20) clearly showed a binding of the target DNA by the PTOV1 GST-eAT protein (upper panel). This is illustrated by the appearance of two high molecular weight protein DNA complexes in lanes 5 and 6. Apparently, with an increase of the protein concentration from 30 pmol to 45 pmol all DNA was bound by the protein and only one band was visible in lane 7. This was indicating that the DNA was bound by more than one protein, resulting in a higher molecular weight complex. As expected the PTOV1 GST-eAT Mut Core protein showed only a very slight association to the DNA (Figure 3.20, middle panel), marked by the appearance of a faint high molecular weight signal in the lanes 6, 7, and 8 of the gel. In addition almost none of the free DNA was shifted with a protein concentration of 45 and 68 pmol in contrast to the wild type, were the complete DNA was already bound at these protein concentrations. The GST-eAT Core protein displayed no DNA binding activity in these experiments (Figure 3.20, lower panel).

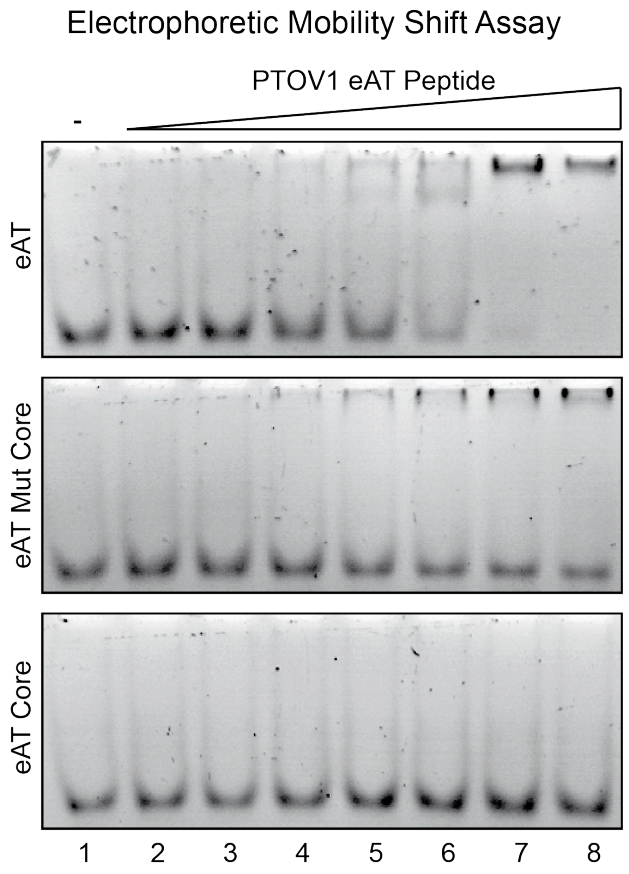


Figure 3.20: PTOV1 eAT-hook DNA binding activity
1.25 pmol of 34 bp AT-rich DNA were incubated with increasing amounts of the indicated protein (6 pmol, 9 pmol, 13 pmol, 20 pmol, 30 pmol, 45 pmol, 68 pmol) for 15 min on ice. The complexes were separated on 7.5 % native PAA gels. DNA was visualized with EtBr. Lane 1 represents the negative control containing only the DNA. Lanes 2-8 represent the samples with increasing protein concentrations.

The results obtained for the GPBP1 GST-eAT proteins were similar to PTOV1 (Figure 3.21). The GPBP1 GST-eAT protein also revealed a DNA binding activity, however in contrast to the PTOV1 motif, only one high molecular weight band appeared upon increasing protein concentrations (Figure 3.21, upper panel, lanes 4-9). The protein concentration was not high enough to shift the entire target DNA in the gel (lane 9). The observed binding pattern indicates the formation of complexes of DNA with multiple proteins, since there were no intermediate bands visible in the gel. This could also indicate a higher cooperativity in the GPBP1 binding compared to PTOV1. For the GPBP1 GST-eAT Mut Core protein no DNA binding activity could be detected (Figure 3.21, middle panel), the same was true for the GPBP1 GST-eAT Core variant (Figure 3.21, lower panel).

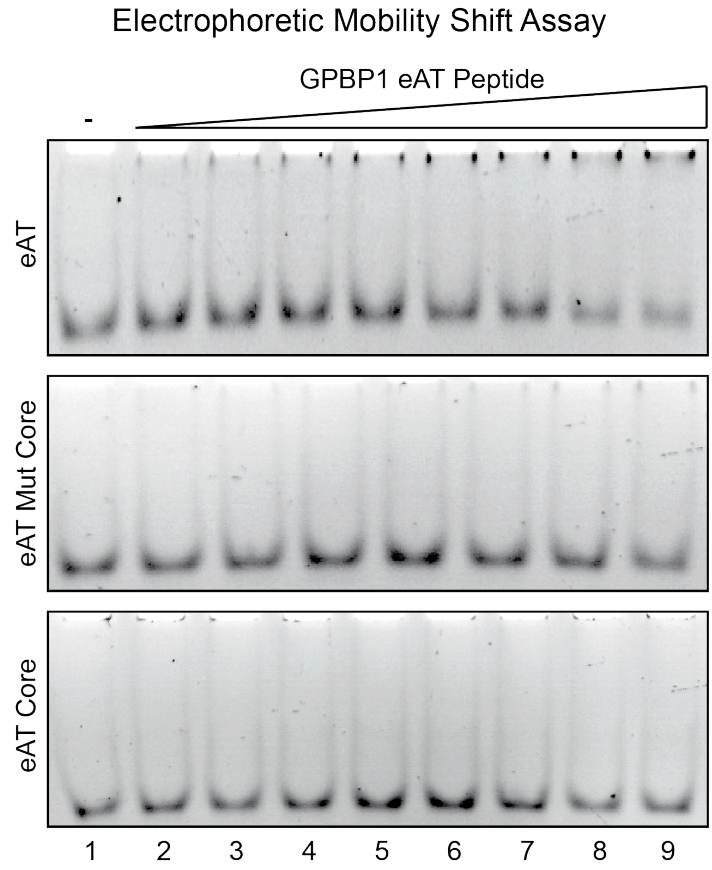


Figure 3.21: GBPB1 eAT-hook DNA binding activity
1.25 pmol of 34 bp AT-rich DNA were incubated with increasing amounts of the indicated protein (7 pmol, 10 pmol, 15 pmol, 23 pmol, 34 pmol, 52 pmol, 79 pmol, 120 pmol) for 15 min on ice. The complexes were separated on 7.5 % native PAA gels. DNA was visualized with EtBr. Lane 1 represents the negative control containing only the DNA. Lanes 2-9 represent the samples with increasing protein concentrations.

In summary the PTOV1 and GPBP1 eAT-hook proteins showed a DNA binding behavior similar to the results obtained for the canonical Tip5 AT-hook domains. To further dissect the characteristics of the eAT-hook domains and to get an idea on the DNA binding affinity of these domains, eAT-hook DNA interactions were quantified using MST.

For this purpose the AT-rich rDNA sequence used for the EMSA experiments was labeled with Cy5 and incubated with serial dilutions of the different PTOV1 and GPBP1 GST-eAT variant proteins. The obtained thermophoresis signals were normalized to the fraction bound as described in (5.2.8.3) and plotted against the protein concentration (Figure 3.22A/B). The results displayed corresponded to the data derived by the previous EMSA experiments. The PTOV1 GST-eAT and GPBP1 GST-eAT proteins bound the DNA with the highest affinity, displaying EC_{50} values

of 4.0 +/- 0.7 µM for PTOV1 (Figure 3.22A) and 7.0 +/- 1.7 µM for GPBP1 (Figure 3.22B), respectively. For the PTOV1 GST-eAT Mut Core a slight DNA binding activity could again be detected, similar to the results obtained by the EMSA. However, no EC_{50} value could be calculated from the data (Figure 3.22A, blue data points), since the reaction did not reach saturation and therefore no proper curve fit was possible. Since the PTOV1 GST-eAT Core protein showed no interaction with the DNA, the thermophoresis signals were normalized on the GST-eAT data (Figure 3.22A, purple data points). The same observation was made for the GPBP1 GST-eAT Mut Core and GST-eAT Core proteins, for both variants no binding could be measured, so again the data was normalized on the GPBP1 GST-eAT measurements Figure 3.22C, blue and purple data points).

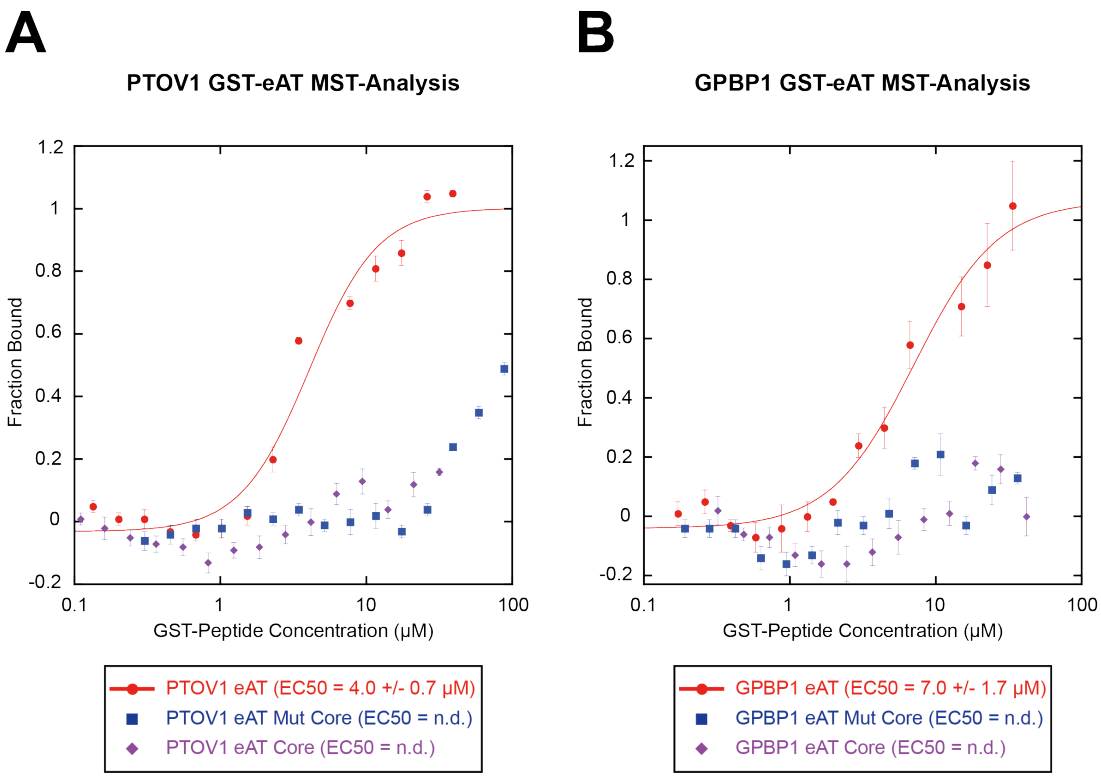


Figure 3.22: MST-Analysis of PTOV1 and GPBP1 GST-eAT variant proteins
For all measurements 50 nM of Cy5-labeled AT-rich DNA was incubated with a serial dilution of the indicated GST-proteins at room temperature. The concentration of the proteins in µM is depicted on a logarithmic scale. The LED power was adjusted to 40 %. Samples were subjected to MST analysis at 40 % MST-power, with a laser on time of 35 sec. Error bars indicate standard deviations of four technical replicates.
A) The MST results for all three PTOV1 GST-eAT variants are depicted. Color-coding relates to the three different proteins as labeled in the box. For the eAT Mut Core and eAT Core proteins no EC_{50} could be detected.
B) The MST results of the three GPBP1 eAT proteins are shown labeling and color-coding is analog to A).

3.3.3 The eAT-hook domains show a preference for RNA binding

Recent publications had already suggested a role for canonical AT-hooks in RNA interactions for the Tip5 and the HMGA1 proteins (Eilebrecht et al., 2011b; Zhou et al., 2009). These observations in combination with a very recent report implicating an RNA component in PTOV1 function (Marqués et al., 2013) made it tempting to speculate about a possible RNA binding activity for the eAT-hook domains. To test this hypothesis MST analysis with the already described GST eAT-variants of PTOV1, GPBP1 and Tip5 were performed.

For these experiments RNA oligonucleotides were selected, which had already been used for protein–RNA interaction studies in our lab (Schubert et al., 2012) (Figure 3.23A). The RNAs oligos represent two parts of the snoRNA: Me28S-U2134b from *Drosophila melanogaster*, which exhibited different secondary structures (Figure 3.23B), as predicted by the RNAfold web tool (Gruber et al., 2008). In a previous study it had been shown that a stem loop structure was necessary for a proper RNA binding in case of the HMGA1 AT-hook (Eilebrecht et al., 2011b). Thus the two RNA oligos with RNA1 producing a large stem loop structure with a long double stranded region and the RNA2, which only comprises a small hairpin structure and large single stranded region presented an interesting structural diversity for the intended assay.

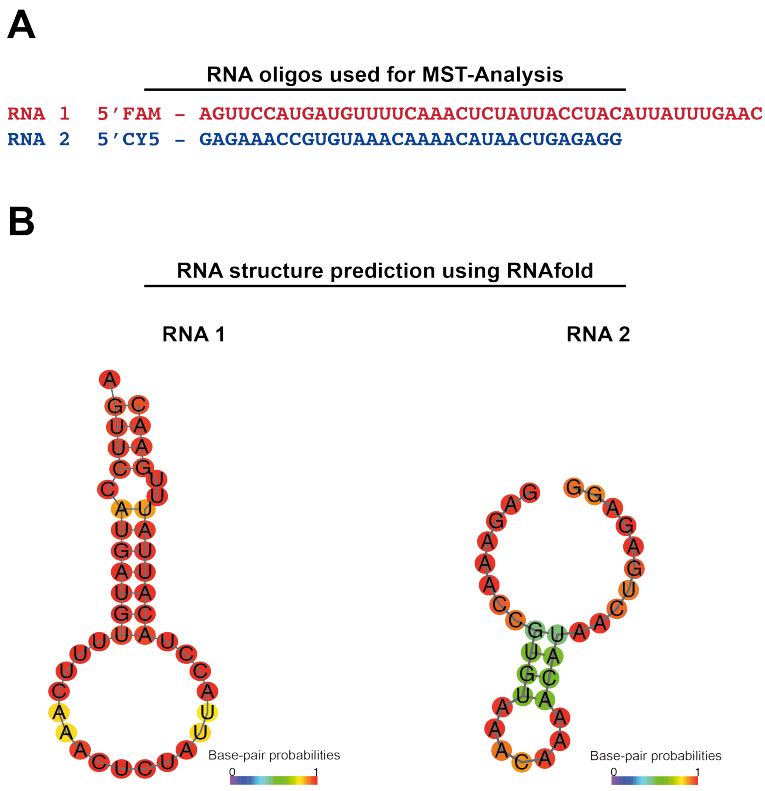


Figure 3.23: RNA oligonucleotides used in MST-analysis of eAT-hook RNA binding
A) The sequences of the RNAs used in the MST assay are depicted. RNA1 was labeled with a FAM moiety on the 5'-end (red sequence). The RNA2 oligonucelotide contained a Cy5-label on the 5'-end (blue sequence).
B) The RNAfold web tool was used for structure prediction of the different RNAs. The colors of the different nucleotides represent the calculated base-pair probabilities.

For the MST measurements, both RNAs were incubated with dilution series of the PTOV1, GPBP1, and Tip5 GST-eAT proteins under identical experimental conditions. The analysis was conducted as described previously, all data points were again normalized to the fraction bound of the RNA molecules and plotted to the protein concentration (Figure 3.24).

The results illustrated in (Figure 3.24) revealed a strong RNA binding activity for all three eAT-hook domains tested. Interestingly all proteins displayed a higher affinity for the RNA1 molecule, with EC_{50} values of 187 +/- 10 nM for Tip5 eAT, 203 +/- 6 nM for PTOV1 eAT and 378 +/- 18 nM for GPBP1 eAT. These values corresponded well to the observed DNA binding behavior, with the Tip5 and PTOV1 domains displaying similar affinities and the GPBP1 showing a decreased binding activity (see 3.3.2). All three proteins also bound the RNA2 molecule, however a two to three fold decrease in binding affinity was detected. Therefore

the EC_{50} values increased to 569 +/- 40 nM for Tip5 eAT, 600 +/- 30 nM for PTOV1 eAT and 938 +/- 92 nM for the GPBP1 eAT protein.

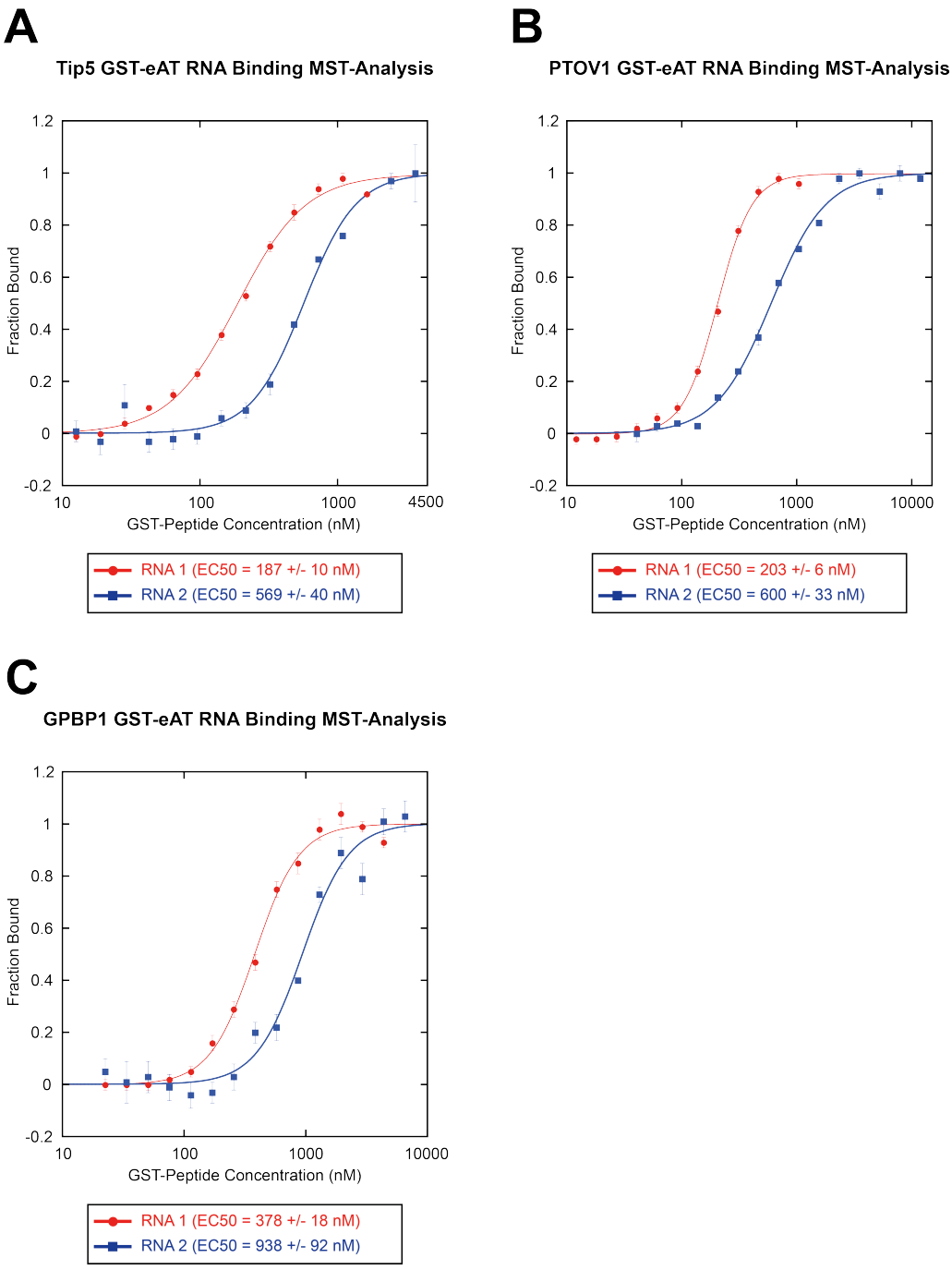


Figure 3.24: GST eAT-hook RNA binding analysis via MST
50 nM of Cy5 labeled RNA1 (red data points) or FAM labeled RNA2 (blue data points) oligonucleotides were incubated with serial dilutions of the three different proteins and MST measurements were carried out at room temperature with 40% led power, 40% MST power and a laser on time of 35 sec. The thermophoresis signals were again normalized to the fraction bound and plotted to the protein concentration of the individual proteins, depicted on a logarithmic scale in nM concentration. Error bars represent standard deviations of four technical replicates.
A) Tip5 GST-eAT-hook RNA binding, **B)** PTOV1 GST-eAT-hook RNA binding, **C)** GPBP1 GST-eAT-hook binding

In summary the results of the MST analysis showed a strong RNA binding behavior of all tested eAT-hook domains.

To further dissect the function of the eAT-hook domains in nucleic acid binding, competitive MST analysis for DNA versus RNA interaction was performed. For this purpose, the FAM-labeled RNA1 oligonucleotide and the Cy5-labeled AT-rich rDNA sequence used in previous assays were mixed in equimolar concentrations and MST measurements were carried out and analyzed as described in 3.2.5.

The results shown in (Figure 3.25) (red data points) underlined the observed preference for RNA of the three different eAT-hook domains. The Tip5 domain bound to the RNA with an EC_{50} of 277 +/- 8 nM, the PTOV1 displayed an EC_{50} of 252 +/- 10 nM and the GPBP1 domain revealed an EC_{50} of 373 +/- 19 nM. In all three cases no interaction with the DNA was detected, the proteins bound exclusively to RNA (Figure 3.25, blue data points).

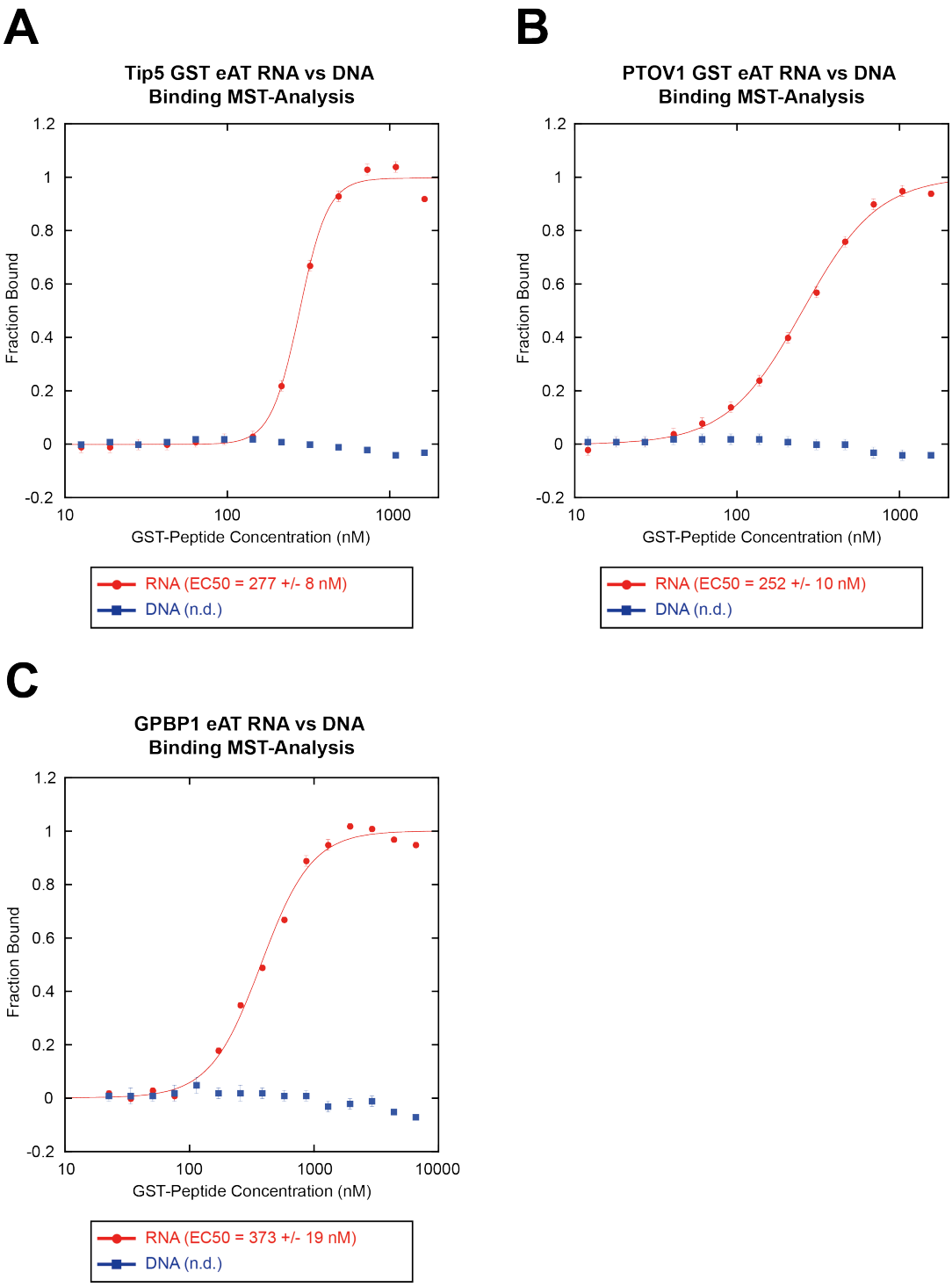


Figure 3.25: Competitive MST analysis of the eAT-hook domains DNA/RNA binding
MST analysis for Tip5 GST-eAT-hook (**A**), PTOV1 GST-eAT-hook (**B**) and GPBP1 GST eAT-hook (**C**) were performed and dissected as described previously (Figure 3.22/Figure 3.24). The RNA data is illustrated in red and the DNA measurements are labeled in blue. EC_{50} values for each protein are depicted in the boxes. Error bars represent standard deviations of four technical replicates.

3.3.4 Loss of the eAT-hook domain impairs shuttling of PTOV1 from the nucleus to the cytoplasm

In view of the fact that all three tested eAT-hook motifs seemed to be functional and not only revealed a DNA binding activity, but at the same time displayed a strong preference for RNA binding in the MST experiments, further investigation of these activities *in vivo* was the logical consequence. PTOV1 represented the prime candidate for these investigations, since a recently published study suggested a putative role of the protein in ribonucleoprotein shuttling to the cytoplasm and regulation of translation initiation. Hence it was tempting to speculate about a role of the eAT-hook domain as nucleic acid interaction module in these processes. To test this hypothesis, a GFP-PTOV1 full-length fusion protein (PTOV1 WT), as well as a GFP-PTOV1 ΔeAT construct with an N-terminal deletion of the first 48 amino acids was prepared. Since the eAT-hook motif presents the only characterized nucleic acid interaction domain of PTOV1, this deletion was likely to interfere with its functions in RNA or DNA mediated processes.

The two constructs were transfected in HeLa cells and the sub-cellular localization of the expressed GFP fusion proteins was examined by immunofluorescence. Strikingly, PTOV1 WT and PTOV1 ΔeAT revealed a clear difference in their localization patterns, which is illustrated in (Figure 3.26). In the upper panel the PTOV1 WT expressing cells are depicted. They predominantly showed 3 distinct patterns of GFP-signal distribution with cells demonstrating an exclusively nuclear localization (1), cells revealing evenly dispersed signals in the nucleus and the cytoplasm (2) and cells suggesting a higher concentration of GFP proteins in the cytoplasm (3). The numbered arrows mark examples of the different localization patterns in (Figure 3.26). These results were in good correlation with previously published reports, describing similar patterns, caused by the shuttling mechanism of PTOV1 (Marqués et al., 2013; Santamaría et al., 2005; Santamaría et al., 2003).

In contrast to the WT fusions, the PTOV1 ΔeAT proteins showed only one distinct localization pattern with the GFP signals predominantly located in the nucleus and only very faint cytoplasmic signals (Figure 3.26, middle panel). This suggested an impaired shuttling function for this protein variant. In addition a control transfection with GFP is also shown (Figure 3.26, lower panel).

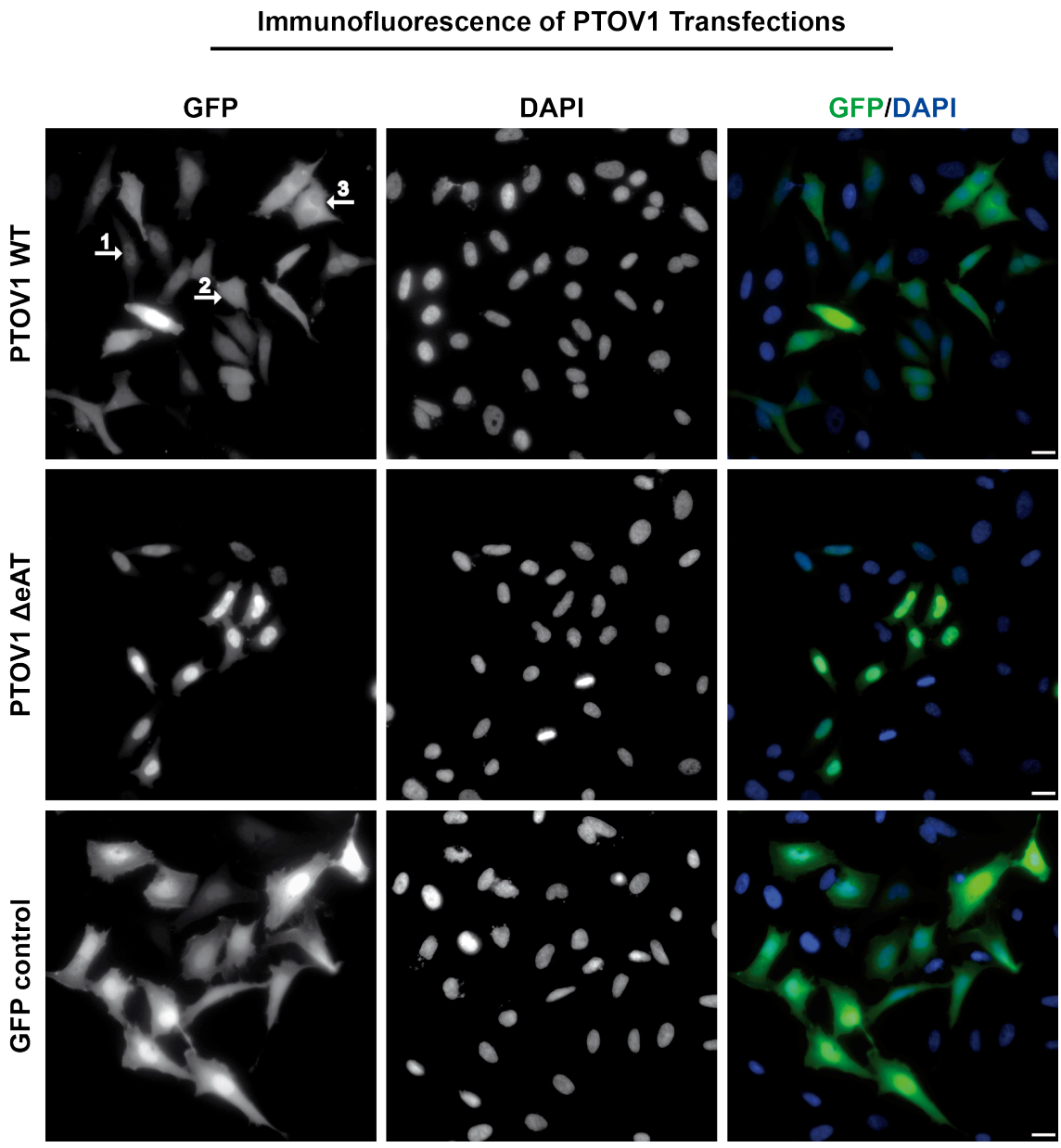
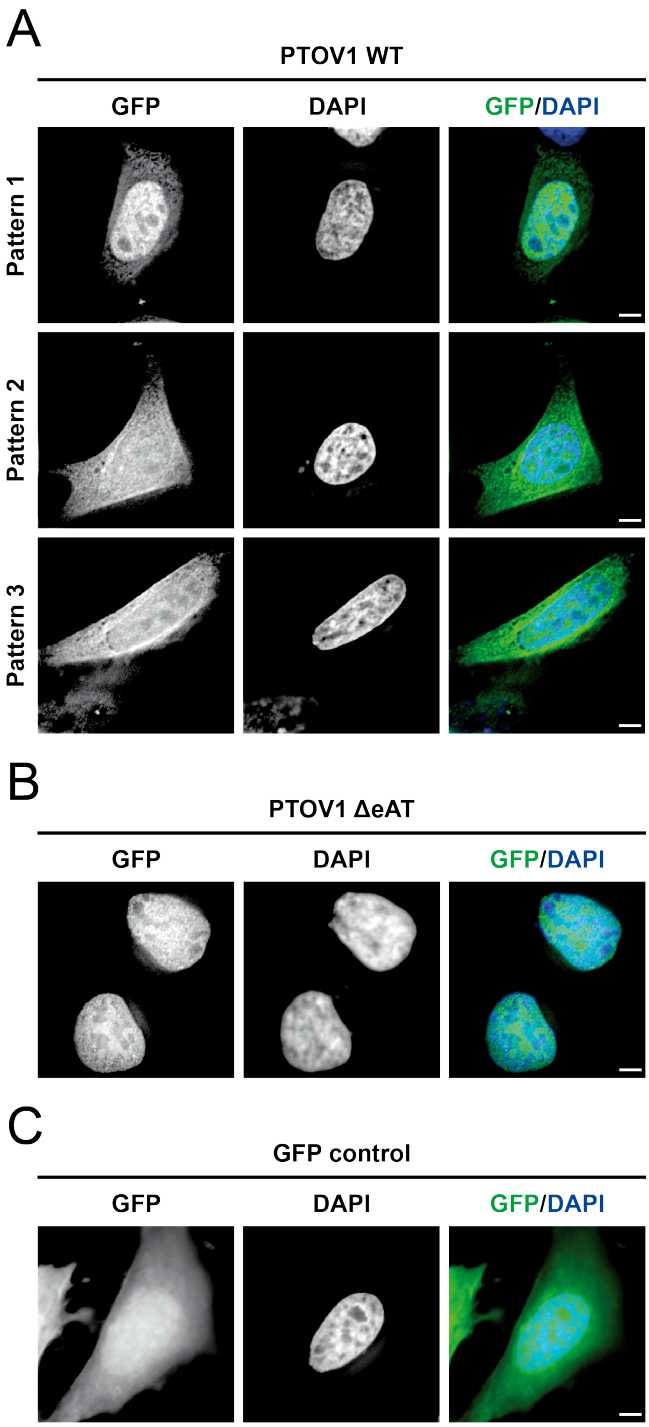


Figure 3.26: The eAT-hook domain influences the sub-cellular localization of PTOV1
HeLa cells were transfected with the GFP constructs as indicated on the left. 24 h post-transfection, cells were fixed in 4 % PFA PBS and immunostaining of GFP was done using a commercial GFP-booster. DNA was counterstained with DAPI. The numbered arrows in the PTOV1 WT GFP picture mark the different sub-cellular locations of the proteins as described in the text. The bar indicates 15 µm.

To further highlight distinct localization patterns, the transfected cells were examined with a higher magnification, in order to get a detailed view on the differential distribution of the GFP fusion proteins throughout the cells (Figure 3.27)



.

Figure 3.27: Loss of the eAT-hook domain leads to a distinct sub-cellular localization of PTOV1

HeLa cells were transfected with the different GFP-PTOV1 constructs and a GFP-control plasmid. 24 h post-transfection cells were fixed with 4 % PFA in PBS and immunostaining of GFP was done using a GFP-booster. The DNA was stained with DAPI. Bars indicate 5µm.

A) The three different predominant localization patterns of the GFP-PTOV1 WT protein are displayed, as described in the text.

B) The GFP-PTOV1 ΔeAT proteins revealed only a nuclear localization.

C) Transfection with a GFP-control plasmid is shown.

Again cells with pattern 1 presented a predominant nuclear localization of the GFP signals. In contrast cells resembling pattern type 2 exhibited a more even distribution of the GFP proteins throughout the nucleus and the cytoplasm. And finally in pattern type 3 the GFP signal intensity in the cytoplasm was slightly increased compared to the nucleus, also validating the previous results.

The PTOV1 ΔeAT expressing cells exhibited the exclusively nuclear localization as seen before. In (Figure 3.27C) a GFP control cell is shown.

These observations were clearly indicating an important role of the eAT-hook domain in regard to the shuttling mechanism of PTOV1 since the results presented here suggested an impairment of this function upon deletion of the eAT-hook domain.

4 Discussion

4.1 Characterization of putative triplex forming sites present in the rDNA enhancer and terminator regions

4.1.1 The mouse rRNA genes contain clusters of putative triplex forming sites in their regulatory regions

The overrepresentation of long oligopurine/oligopyrimydine sequences in the gene regions of eukaryotic genomes has already been proposed about 20 years ago (Behe, 1995). The notion that these unusual sequences could play a role in genome regulation by forming uncanonical DNA structures like triplexes was present right from the start (Wells et al., 1988). However, only recent advancements in high-throughput sequencing technologies unraveled the complete triplex forming potential present within the human and mouse genomes (Buske et al., 2012; Goñi et al., 2004; Goñi et al., 2006; Wu et al., 2007). Thereby highlighting the significant enrichment of putative triplex-forming sites in regulatory regions, like enhancer and promoter elements (Goñi et al., 2004; Goñi et al., 2006; Wu et al., 2007).

Strikingly the results presented in (3.1.1) illustrated that the mouse rDNA unit not only follows this observations, it also contains a cluster of putative TTSs throughout its regulatory enhancer and terminator regions. The analysis of these functional gene elements with the recently developed Triplexator software tool (Buske et al., 2012) clearly revealed those mononucleotide repeat containing oligopyrimidine motifs as possible target sites for triplex formation. Interestingly the partial sequence overlap caused by the mononucleotide repeat components of the putative triplex forming motifs, would enable a targeting of several TTSs with a single TFO. Remarkably, in this context the recently discovered non-coding pRNA (Mayer et al., 2006), could provide a source for such potential enhancer and terminator binding TFOs. Although non-specific binding in respect to the genome would be possible for such low complexity sequences, the close proximity of the

TFOs, their targets, and the conditions before and after transcription, could promote the specific binding in such a microenvironment (Buske et al., 2012).
Besides being putative TTSs the polyT-stretches present in the rDNA enhancer have already been implicated in the formation and positioning of the enhancosomes, which represent distinctly folded structures formed by UBF and the repetitive enhancer elements (Grozdanov et al., 2003; Moss and Stefanovsky, 1995). Furthermore they have been assigned special physical intrinsic properties (Nelson et al., 1987), which make them inhibitory for nucleosome formation (Segal and Widom, 2009; Suter et al., 2000), thereby playing a role in nucleosome positioning and as spacing blocks in the assembly of the transcriptional machinery (Goñi et al., 2006). However the exact function of these sequence motifs has not been unraveled so far. Thus it was tempting to follow up on their possible function as triplex target sites, by further validating the triplex forming potential *in vitro*.

4.1.2 The rDNA enhancer motif forms a stable triplex complex *in vitro*

As expected from previous results (Lang, 2008) and the *in-silico* analysis, the enhancer TTS readily showed a triplex formation with either a RNA or a DNA TFO in the double colored EMSA experiments (3.1.2). The complex formation induces changes in the microenvironment of the fluorophore, which lead to the observed decrease in fluorescence signal intensity. This was already reported by other groups, using these fluorescent changes to track triplex formation and dissociation kinetics (Ellouze et al., 1997). The results suggested an almost complete complex formation upon a duplex to TFO ratio of 1:2 for the DNA/RNA and 1:4 for the DNA/DNA triplexes in these experimental conditions. The apparent incomplete overlap of the two different fluorescent signals could be explained by a slight optical shift due to the imaging with two different filter sets. Since the EMSA experiments did not offer the possibility to gain quantitative data on triplex formation in different experimental conditions, an assay to measure the complex formation via the MST technology was established.

Strikingly, the results of these experiments provided insights into various aspects of triplex formation. First of all an enhancing effect of third strand binding on the fluorescent label located on the 5'-end of the duplex DNA was visible, correlating with an increased TFO concentration and triplex formation (Figure 3.3). Second, the MST profiles revealed a clear thermophoretic difference between TFO bound and unbound DNA duplex for both triplex pairs, illustrating the possibility to readily measure triplex formation with this assay (Figure 3.4).

As described in detail in (5.2.8.1) the MST data set provides information on different aspects of the binding over the time course of the experiments. The temperature jump represents fast binding events, which mostly affect the microenvironment around the fluorophore and therefore describe rather localized binding directly affecting the label. On the other hand, the thermophoresis behavior, which takes place on the later timescale of the MST experiments represents changes in the charge, size and the hydration shell and therefore indicates rather global changes of the fluorescently labeled molecules (Jerabek-Willemsen et al., 2011). The process of triplex formation has been characterized to follow a highly cooperative, so-called nucleation zipper mechanism starting at the 5'-end of the duplex DNA (Alberti et al., 2002; Maher et al., 1990; Rougée et al., 1992; Shindo et al., 1993). Thus the temperature jump signal of the MST measurements mirrors the binding effects of the TFO on the fluorescent label at the 5'-end of the duplex DNA. Therefore these signals should correlate with triplex formation and the corresponding time points were used for the analysis. The generated binding curves (Figure 3.5) depicted a clear triplex formation for both complexes, with K_d values comparable to the literature (Jain et al., 2008; Maher et al., 1990; McDonald and Maher, 1995). The DNA/RNA triplex showed a more cooperative binding behavior compared to the DNA/DNA triplex. This could represent a higher affinity of the RNA TFO, which leads to the better stabilization of a nucleation intermediate, favoring a rapid triplex formation. With the DNA TFO the binding process seemed to involve more than one intermediate state, suggesting a lower affinity for the DNA duplex, resulting in various intermediates. These findings correlate with previously published data displaying similar results for the differences in complex formation between DNA/RNA and DNA/DNA triplexes (Escudé et al., 1993; Roberts and Crothers, 1992). Overall the results not

only validate the expected triplex formation of the motif derived from the rDNA enhancer, but more importantly at the same time illustrated the potential use of the MST technology for triplex characterization.

4.1.3 The thermophoretic behavior of triplex complexes changes with stability

The initial MST analysis of triplex formation demonstrated differences in the thermophoretic profiles seen for DNA/RNA and DNA/DNA triplexes under identical experimental settings.

The experiments presented in (3.1.4) were designed to monitor the binding affinity at elevated temperatures to compare the behavior of DNA/DNA and DNA/RNA triplexes.

Upon an increase of the incubation temperature to 35°C the clear three state profile seen for the initial experiments at 25°C changed and the thermophoretic behavior of the samples with high TFO concentrations was reversed, showing a slower mobility in the temperature field compared to the lower TFO concentrations. This effect was not equal for all samples and correlated to the TFO concentrations, with the capillaries containing the highest TFO concentrations displaying a smaller decrease in mobility. This resulted in different slopes of the thermophoretic curves for the samples containing high concentrations of TFO and thus induced a scattering of the curves over the whole amplitude of the experiment at the later time points of the measurements, which is illustrated in (Figure 3.6). The effect could be explained by a previously observed stabilizing influence of high TFO concentrations (James et al., 2003), resulting in a slightly different reaction of the differently concentrated samples to the temperature increase. However, the exact nature of this effect still needs to be determined. Therefore it would be interesting to elucidate, if it is actually caused by a specific stabilization of the triplex via the TFO, or if it represents some indirect secondary effect. Thus further experiments using high concentrations of unspecific oligionucleotides could be performed to get a better insight in this mechanism.

In addition the thermophoretic shift led to a decrease in the temperature jump differences with rising temperature, which made it impossible to generate binding curves from these time points. However, the thermophoresis data still showed a clear interaction in these measurements. A possible explanation for this phenomenon would be that the decreased triplex stability leads to a loose binding at the 5'-ends of the DNA thereby decreasing the influence on the fluorescent label. This may in turn result in smaller temperature jump differences (Ellouze et al., 1997). However, the change in initial fluorescence observed with increasing TFO concentration could still be observed (data not shown), which rather suggested the loss of temperature jump differences to represent a technical limitation of the MST device. Nevertheless, the triplex formation could be detected throughout the whole temperature gradient in either the temperature jump or the combined temperature jump and thermophoresis data of the single experiments.

4.1.4 Triplex formation is dependent on cation concentration and pH

The formation of triplex structures displays a crucial dependency on the concentration of divalent cations and on the pH (Shindo et al., 1993; Singleton and Dervan, 1993; Sugimoto et al., 2001). The influence of these parameters on triplex formation and stability has already been investigated with a wide variety of methods. Thus this system was perfectly suited to further compare the triplex MST assay with already established methods.

Overall the results presented in (3.1.5.1) were in good correlation with previous reports investigating the stability of triplex structures in regard to magnesium concentrations. The data clearly highlighted the necessity of certain amounts of divalent cations for triplex formation. Without magnesium or with only 1 mM of magnesium and no other positively charged ions in the buffer no complex formation could be seen for the tested triplexes. At 5 mM magnesium the situation changed, and a triplex formation for both pairs could be detected. Interestingly, the observed minimal concentration of 5 mM magnesium for a stable triplex formation was almost identical with a previous report (Ellouze et al., 1997). Upon 20 mM

magnesium the DNA/RNA triplex seemed to be saturated, since there was no change in the measured parameters detectable anymore.

The exact comparison of the generated K_d values with published data proved to be rather difficult, due to the wide variety of different sequences, buffer systems and assay techniques used. In general the obtained values ranging from roughly 10 nM to 1.5 µM were matching previously published data (Jain et al., 2008; Maher et al., 1990; McDonald and Maher, 1995; Nakanishi et al., 1998; Shindo et al., 1993). Furthermore the melting temperatures for the triplex to duplex transitions were in good correlation to the literature as well (Hoyne et al., 2000; James et al., 2003). Strikingly, the DNA/RNA triplex seemed to be more stable at lower magnesium concentrations and at elevated temperature (up to 40°C), which was in line with previous results, suggesting that a triplex structure comprising a DNA duplex and an RNA third strand is slightly more stable than DNA/DNA complexes (Escudé et al., 1993; Roberts and Crothers, 1992).

Surprisingly, the K_d values for the DNA TFO at 20 mM and 50 mM magnesium and a temperature range up until 40°C were clearly lower than the corresponding RNA TFO constants, indicating a higher affinity (Table 3.2/Table 3.3). However, when the temperature increased above 40°C the K_d for the DNA TFO increased rapidly, suggesting a lower overall stability of the DNA/DNA triplex. This somewhat controversial behavior of a DNA/DNA triplex in highly stabilizing conditions, like buffers containing high amounts of magnesium has already been described in different reports (Durand et al., 1992; Hoyne et al., 2000). Therein it was shown that upon increase of magnesium, a DNA/DNA triplex lost its biphasic melting profile. The authors also state that although the complex itself still seemed to correspond well to that of a DNA/RNA triplex, the nature of complex formation and dissociation with the DNA TFO changes upon higher concentrations of magnesium, greatly differing from the characteristics of a DNA/RNA triplex. These effects could also influence the thermophoretic behavior of the DNA/DNA triplex under such conditions, leading to the observed decrease in the K_d values compared to the DNA/RNA triplexes. Interestingly, also the rapid increase of the K_d constants with increasing temperature clearly differs from the properties of the DNA/RNA triplex, underlining the reported differences in complex dissociation and formation

(Durand et al., 1992; Hoyne et al., 2000). However the exact mechanism of this behavior could not be elucidated so far and needs further investigations.

The results for the pH dependency presented in (3.1.5.2) painted a similar picture as described for the magnesium data. As expected in comparison to the triplex formation at physiological pH, the triplex complex is very stable at a lower pH of 6.4. Again in this strongly stabilizing environment, the DNA/DNA triplex displayed extremely small K_d values, which increased significantly faster in comparison to the DNA/RNA triplex upon a certain temperature. The DNA/RNA complex was stable over the whole temperature range with K_d values between approximately 6-15 nM, in contrast the DNA/DNA showed K_d constants as low as 2 nM at lower temperatures, which changed to 100 nM at 60°C. Interestingly at pH 7.4 the DNA/DNA triplex seemed to be significantly more stable than the DNA/RNA triplex. This could be explained by previous reports, showing that the higher stability of DNA/RNA triplexes compared to DNA/DNA triplexes is not always the case and largely depended on the sequence and buffer conditions (Han and Dervan, 1993). In the pH assay the annealing buffer contained 25 mM of sodium, which has been shown to compete with the positive magnesium effects in triplex formation (Chen and Chen, 2011; Maher et al., 1990). Therefore the slight decrease in the DNA/RNA triplex stability could be a consequence of a higher impact of the sodium effect on this complex. At pH 6.4 this effect was not detectable, correlating with previous studies suggesting the triplex formation to be independent, or less dependent on cation effects at lower pH (Shindo et al., 1993). Finally as expected upon a pH of 8.4 no triplex formation was visible.

In summary the results of the presented experiments confirmed previous observations on the differential effects of cation concentrations and pH on triplex formation and dissociation for DNA/RNA and DNA/DNA triplexes. Furthermore they provided new insights on the differences in triplex formation and dissociation of DNA/RNA and DNA/DNA triplexes.

4.1.5 Enhancer triplex formation occurs in near physiological buffer conditions

Even after 50 years of triplex research the *in vivo* existence and function of theses structures still presents the biggest question in the field. The cellular environment likely imposes specific constraints on the potential of triplex formation. Bivalent cations stabilizing the triple helix are present in complex with the DNA and proteins within the nucleus. However, physiological concentrations of potassium have been shown to interfere with triplex formation, against the stabilizing effect of bivalent cations (Olivas and Maher, 1995a; Olivas and Maher, 1995b). In addition polyamines like spermine and spermidine, which are present in the cell in millimolar concentrations, as well as cationic peptides, have been revealed as factors increasing triplex formation at physiological pH (Hampel et al., 1991; Potaman and Sinden, 1998). Finally the formation of such complexes *in vivo* is obviously highly dependent on the chromatin accessibility of the target motifs (Brown and Fox, 1999; Brown and Fox, 1998; Westin et al., 1995).

In chapter 3.1.6 triplex formation was analyzed using the MST technology in a simplified cell buffer system, approximately matching the pH and ionic concentrations found in eukaryotic cells (Alberts, 2008). Since the concentration of divalent cations has been shown to be a crucial parameter for triplex formation three different Mg^{2+} concentrations were used, resembling the free and complexed amounts of bivalent cations within the cell (Alberts, 2008). The results revealed an interaction between the RNA TFO and the DNA starting at 5 mM magnesium. At 10 mM a stable triplex formation could be detected (Figure 3.7). The observations for the DNA/DNA triplex were similar, although here no interaction was detectable at 5 mM magnesium concentration and the K_d of the complex formation at 10 mM magnesium was approximately 3-fold higher than determined for the DNA/RNA triplex. In general the obtained K_d values were considerably higher than observed for the optimal triplex buffer conditions, indicating the negative effects of the sodium and potassium ions on triplex formation and stabilization. On the other hand the cell buffer did not contain any polyamines, which have been shown to strongly stabilize the triplex structures and are present in cells at mM

concentrations (Hampel et al., 1991; Singleton and Dervan, 1993). Thus the complexes could probably be more stable in an *in vivo* situation.

These results were in line with previous reports, showing the formation of stable triplexes in physiological buffer conditions and *in vivo* systems (Besch et al., 2004; Hoyne et al., 2000; Singleton and Dervan, 1993; Sugimoto et al., 2001). This suggested that triplex formation between motifs in the rDNA enhancer and ncRNAs derived from the pRNA could possibly occur *in vivo*. Interestingly, a recent report described a putative triplex forming polypurine repeat containing lncRNA in mouse cells. The study implicated a possible role of this lncRNA in nuclear matrix organization and tethering, which could in part be mediated by triplex formation (Zheng et al., 2010). Thinking about a similar function for the pRNA is purely speculative at this point, however it builds the bridge to the following chapter dealing with the large-scale organization of repressed rDNA chromatin at the nuclear matrix.

4.1.6 Triplex formation – Summary, biological implications and further perspective

With the results presented here, the MST technology could be introduced as a valuable tool for the quantitative *in vitro* characterization of potential triplex forming DNA elements. The technique presents a very fast and easy possibility to screen putatively interesting DNA regions for their triplex forming potential, making it a valuable tool with regard to the therapeutic application of TFOs following the antigene strategy. Especially the possibility to easily measure interactions in solution under complex conditions (Seidel et al., 2013) presents a clear advantage to the alternative methods like for example, SPR. A further development of the assay in more complex buffer systems, with different templates, like *in vitro*-assembled nucleosomes, could help to gain a better insight into the processes influencing triplex formation *in vivo*.

It was furthermore possible to reveal the large triplex forming potential of the regulatory terminator and enhancer regions of the mouse rDNA. The experiments clearly showed that a formation of triplex structures between RNA and DNA would

be possible at these elements in conditions resembling those within the nucleus. Therefore with regard to recent discoveries (Mayer et al., 2006; Zheng et al., 2010) it is tempting to speculate about a functional role of such structures in the context of ncRNA-mediated processes. However the apparent transient nature of these complexes and the lack of reliable tools, like specific triplex antibodies, make the characterization of such mechanisms extremely challenging. Therefore the generation of a triplex specific binder and the development of protocols clearly detecting triplex formation and triplex induced effects on the rDNA, by transfections with psoralen linked TFOs could be the next steps for this project.

4.2 Large-scale organization of ribosomal DNA chromatin is regulated by Tip5

The nucleolar remodeling complex NoRC is a key player in epigenetic repression of rRNA gene expression, which functions through repositioning of the promoter bound nucleosome and the initiation of heterochromatin formation by its interactions with HDACs and Dnmts (reviewed in (McStay and Grummt, 2008)). Here the regulatory subunit of NoRC, Tip5 is revealed as nuclear matrix binding protein, mediating the association of rDNA chromatin to the nuclear matrix upon repression of rRNA synthesis, thereby playing a role in the higher order organization of rDNA cromatin. The results therefore imply an additional role of NoRC, besides regulation of the local chromatin structure at the promoter, in the large-scale chromatin domain organization of the rDNA locus.

4.2.1 Serum starvation induces global changes in nucleolar architecture and an enrichment of rDNA in the nuclear matrix

The results presented in chapter (3.2.1) showed that serum starvation lead to clear changes in nucleolar architecture, suggesting an alteration in the spatial

organization of rDNA chromatin upon the repression of rRNA synthesis. This was underlined by the fact, that a significant enrichment of rDNA within the nuclear matrix could be detected upon rRNA gene repression either by serum starvation (3.2.1) or Tip5 overexpression (3.2.4). The observations presented here, suggested the rDNA enrichment in the nuclear matrix to be Tip5 and MAR dependent for the promoter and IGS fragments, resembling a known binding site of the protein and a putative MAR respectively. At the same time the 28S fragment lacking MARs and Tip5 binding sites suggested a Tip5 and MAR independent enrichment. The apparent involvement of Tip5, the regulatory subunit of NoRC, which resembles a master regulator of rRNA gene repression (Li et al., 2006; Li et al., 2005; McStay and Grummt, 2008; Santoro et al., 2002; Strohner et al., 2001; Strohner et al., 2004; Zhou et al., 2002), made it tempting to speculate about a possible reorganization of the NoRC-repressed rDNA repeats to the nuclear matrix.

The association of rDNA with the nuclear matrix had already been described previously (Craig et al., 1997; Pardoll and Vogelstein, 1980). In these studies biochemical and cell biological experiments were used, to demonstrate the specific enrichment of rDNA in the nuclear matrix. Yet, none of these publications revealed any information concerning either the transcriptional status of the associated rRNA genes, or the sequences within the repeat unit mediating the matrix attachment. When dealing with the question on the transcriptional activity of matrix associated rRNA genes seemingly opposing models were proposed. On the one hand people suggested, that the active rRNA genes are located to nuclear matrix (Keppel, 1986; Smith and Rothblum, 1987), and on the other hand there were groups claiming that the inactive repeats are tethered to the nuclear matrix (Bolla et al., 1985), also representing sequences being replicated (Little et al., 1993). Regarding the question on sequence motifs mediating matrix attachment, within the rDNA gene unit a similar situation exists. The models proposed reach from the entire rDNA repeat is associated with the matrix (Keppel, 1986), to the coding sequence itself (Schwarzacher and Mosgoeller, 2000), or non-transcribed regions flanking the 47S rRNA coding sequence being predominantly enriched in the nuclear or nucleolar matrix (Bolla et al., 1985; Smith and Rothblum, 1987; Stephanova et al., 1993).

4.2.2 Defining the nuclear matrix

The discrepancies in defining the rDNA content associated to the nuclear matrix can be explained by significant differences in the use of terminology and largely varied experimental procedures. In general biochemical fractionation experiments, applying different endonuclease digestions, followed by high-salt or low-salt extractions, or fractionations at physiological salt concentration, are performed. Thereby, nuclear matrix (Berezney and Coffey, 1974), nuclear scaffold (Mirkovitch et al., 1984) or nuclear skeleton (Jackson et al., 1988) fractions are generated, which are merely operational definitions and not defined cell compartments. The key step in all these protocols is an extensive DNase I digest, largely responsible for the nature of the resulting fractions. Remarkably, the incubation time of the endonuclease digestion and the concentration of DNase I vary frequently between the protocols of different laboratories. An initial study clearly revealed that this might affect the association of the rDNA with the nuclear matrix, thereby introducing the observed differences (Pardoll and Vogelstein, 1980). To further complicate the situation there are also groups adding additional nucleolus isolation steps to their protocols (Bolla et al., 1985; Stephanova et al., 1993), making the comparison of published data regarding the nuclear matrix association of rDNA extremely difficult.

4.2.3 The whole rDNA repeat is recruited to the nuclear matrix

For the results presented here, the nuclear matrix fraction was prepared and named essentially as described in previous publications (Berezney and Coffey, 1974; He et al., 1990; Reyes et al., 1997). Therefore it is important to stress, that this nuclear matrix fraction does not represent a distinct sub-cellular compartment (Hancock, 2000; Pederson, 2000). However its protein components largely resemble that of the nukleoskeleton, which describes a well-defined, intermediate filament-based protein network of the nucleus (Simon and Wilson, 2011; Zillner et al., 2013). Additionally the content of the nuclear matrix mainly represents fractions of the genome, which are highly resistant to DNase I digestions, showing

an enrichment of specific sequences important for gene regulatory processes (Gavrilov et al., 2010; Ottaviani et al., 2008). Looking at the fact that open chromatin structures are highly accessible to nucleases (Reeves, 1984) and active rDNA is largely nucleosome-depleted (Hamperl et al., 2013) we supposed that predominantly inactive rDNA repeats are associated with the nuclear matrix.

The results generated by the serum starvation experiments and the overexpression of Tip5, lead to the assumption, that the whole rDNA repeat was recruited to the nuclear matrix. Interestingly the relatively small changes for the predicted MAR in the IGS region suggested the presence of this region in the nuclear matrix already prior to the serum starvation, therefore representing a nucleation site for the further association of the entire repeat.

4.2.4 Tip5 is a nuclear matrix protein, mediating the association of the rDNA to the nuclear matrix

Besides the DNase I inaccessible regions, the nuclear matrix fraction contains a large proportion of different proteins and RNA molecules. Interestingly, the immunofluoresence analysis of *in situ* nuclear matrix preparations clearly demonstrated a nuclear matrix association of Tip5. These results were further validated by immunoblot experiments ((Zillner et al., 2013), performed by Karina Zillner), confirming the observed interaction of a large proportion of the protein with the nuclear matrix.

Since an interaction with a recently discovered ncRNA, termed pRNA, had been described to be necessary for the targeting of NoRC to the rDNA promoter and a similar mechanism was suggested by the presence of the protein at the centromeres (Guetg et al., 2010; Mayer et al., 2006; Santoro et al., 2010; Zhou et al., 2009), a possible involvement of RNA in the proteins nuclear matrix association was considered. However, the results showed that the association was insensitive to RNase A treatment. Moreover, a loss of Tip5 from the soluble chromatin fraction was discovered. This opens up the possibility that the pRNA or other regulatory RNAs are involved in the tethering of NoRC to chromatin. At the same time these

observations suggested the presence of two functionally different pools of Tip5 in the cells.

The association of Tip5 with the nuclear matrix, together with the Tip5-dependent nuclear matrix targeting of rDNA ((Zillner et al., 2013)/performed by Karina Zillner) suggests that NoRC regulates rDNA repression at multiple levels. Strikingly, the predicted MAR site within the IGS, as well as a fragment from the 28S coding region and the Tip5 binding site in the promoter were enriched in the matrix fractions upon Tip5 overexpression. This suggests that besides its other functions in context of the NoRC complex as a master regulator in rDNA repression, Tip5 also regulates the DNase I accessibility of rDNA in the nucleus, i.e. nucleolar topology. These findings implicate an additional role of the NoRC mediated rDNA repression in enhancing the matrix association of the rRNA genes.

In summary these results implied a model in which Tip5 is a major regulator in rDNA recruitment to the nuclear matrix. The NoRC mediated chromatin compaction and heterochromatin formation leads to decreased DNase I accessibility and the tethering of large rDNA chromatin domains to the nuclear matrix. Altogether this represents a new role for Tip5 in the spatial organization of higher-order rDNA chromatin structures within the nucleolus.

4.2.5 TAM and AT-hook domains play a role in nucleolar targeting and association of Tip5 to the nuclear matrix

Since the results presented so far turned the spotlight on Tip5 as a major regulator in nuclear matrix association of rDNA chromatin and at the same time identified the protein as an integral part of the nuclear matrix, a further dissection of the protein domains responsible for this association was the next step.

Tip5 contains a TAM and four AT-hook domains, which are supposed MAR binders and mediators of nuclear matrix association (Aravind and Landsman, 1998) and were therefore the main subjects for further investigations. Tests of the DNA binding affinities of the different AT-hook motifs by EMSA and MST experiments identified a combination of the first two AT-hook domains of Tip5 as the most efficient MAR binder (performed by Karina Zillner/(Zillner et al., 2013)).

Additionally the TAM domain had already been described as slightly less competent in DNA binding than the AT-hooks (Strohner et al., 2001). Therefore the double AT-hook motif and its mutant version were also investigated in regard of a possible nuclear matrix targeting function (3.2.6/3.2.7).

Surprisingly, the double AT-hook motif failed to recruit GFP to the nuclear matrix. Hence the TAM domain was also included in the proteins and again the matrix association was tested. Strikingly, the TAM domain was necessary and sufficient to target the proteins to the nuclear matrix, thereby confirming its described role as mediator of nuclear matrix association. Moreover, the experiments showed that both, the TAM domain and the double AT-hook motif, act as nucleolar targeting sequences. Finally a function of these individual Tip5 domains in rDNA association to the nuclear matrix was investigated (performed by Karina Zillner/(Zillner et al., 2013)). But the results revealed no enrichment of rDNA chromatin in the matrix upon overexpression of the different GFP proteins. These findings implicated that additional parts of Tip5 are required for the specific enrichment of rDNA chromatin to the nuclear matrix. Furthermore the overexpression of the TAM domain could lead to genome wide unspecific MAR binding thereby preventing detectable effects on the rDNA. Nevertheless, the overexpression of Tip5 led to the expected enrichment of rDNA domains within the nuclear matrix. In summary, the findings presented suggested a dual role for the double AT-hook motif and the TAM domain in nucleolar targeting and anchoring of Tip5 to the nuclear matrix.

4.3 The extended AT-hook (eAT-hook) is a novel nucleic acid binding domain

In course of the investigations regarding the nuclear matrix binding potential of the Tip5 protein a so far unrecognized extended AT-hook (eAT-hook) domain was discovered. The initial characterization revealed DNA binding activities resembling those of the canonical Tip5 AT-hook motifs (Zillner, 2013). In addition further investigations hinted at a strong RNA binding activity for these new domains. Finally the importance of the motif for proper protein function was verified for a candidate protein called PTOV1 *in vivo*. The data presented here therefore implies

the eAT-hook as a new functional nucleic acid interaction motif, with possible roles as minor groove DNA binding module and nuclear matrix tether. Moreover the results gave a broad hint at a possible RNA interaction indicating additional functions in RNA mediated processes.

4.3.1 The eAT-hook domains of PTOV1 and GPBP1 show DNA binding properties comparable to classic AT-hook motifs

The characterization of the novel Tip5 eAT-hook domain as a functional non-canonical AT-hook (Zillner, 2013) set the stage for the identification of a whole group of novel putative DNA binding domains. The widespread distribution of the $(K/R)_{>3}$-N_{6-20}-GRP-N_{6-20}-$(K/R)_{>3}$ motif throughout the human and mouse proteomes with 81 and 60 different candidate proteins, resembling about twice the number of already mapped AT-hooks, further underlined the significance of this new motif. Interestingly, in addition to the high number, the candidate proteins showed similar molecular functions as canonical AT-hooks, with 29 mouse and 39 human candidate proteins having a suggested role in nucleic acid metabolism, further implicating a role comparable to the AT-hook motif (Zillner, 2013). This notion was backed up by the characterization of the PTOV1 and GPBP1 eAT-hook domains. Both proteins fell into the desired category of having a proposed function in nucleic acid metabolism and at the same time were lacking any described DNA or RNA binding motifs (Benedit et al., 2001; Hsu et al., 2003), making them prime candidates for further investigations in regard to eAT-hook function.

The results gained from EMSA and MST-Analysis (3.3.2) emphasized the observations made for the Tip5 eAT-hook motif (Zillner, 2013). Both domains displayed DNA binding affinities in a range comparable to canonical AT-hooks. Moreover, an unspecific DNA binding by electrostatic interactions with basic amino acids surrounding the GRP motif was ruled out. Although there were some slight interactions visible in the EMSA analysis performed with the GRP core mutants, especially for the PTOV1 eAT-hook, the quantitative MST data clearly

demonstrated a loss of efficient DNA binding by the introduced mutations (3.3.2). Thereby the function of the eAT-hook as new DNA interaction domain was further validated.

The identification of the eAT-hook DNA binding domain was in line with previous reports, suggesting a more divergent definition of the AT-hook domain than proposed in (Aravind and Landsman, 1998). These studies revealed the existence of different non-canonical AT-hooks in Drosophila (Metcalf and Wassarman, 2006) as well as human proteins (Baker et al., 2013), displaying alternative amino acid sequences in, or in close proximity of the GRP core motif. In summary, all these facts point to the existence of a large number of so far unrecognized DNA binding domains, performing similar functions as the AT-hook motifs in minor groove tethering an anchoring of proteins to specific DNA elements.

4.3.2 The eAT-hook domains show a preference for RNA binding

Besides the well-characterized role as variable DNA binding modules, recent reports gave a broad hint at a possible RNA binding activity of selected AT-hook domains. The reports revealed that interactions with functional ncRNAs are mediated by AT-hook motifs in the HMGA1 and Tip5 proteins (Eilebrecht et al., 2011b; Zhou et al., 2009).

Surprisingly, the tested eAT-hook domains of Tip5, PTOV1 and GPBP1 also illustrated a strong RNA binding activity in MST experiments. The affinities of the three different domains measured for RNA were in the nanomolar range and approximately 7-10-fold higher than observed for DNA molecules (3.3.3). Strikingly, the MST-analysis additionally highlighted a structural preference of the peptides for RNA oligonucleotides comprising a predicted long stem-loop structure, over RNAs featuring only a small hairpin structure (3.3.3). When compared, the putative double stranded RNA molecule was bound with an almost 3-fold increased affinity (3.3.3). Noteworthy these results exactly match previously published data, revealing a double stranded stem-loop structure as prerequisite for an efficient AT-hook binding to ncRNA (Eilebrecht et al., 2011b). Moreover, in

competitive MST assays the presence of DNA seemed not to interfere with the RNA interactions, further validating the preferred RNA binding of the eAT-hook proteins.

In recent years the discovery of an ever growing number of functional ncRNAs, working as guides, regulators, and scaffolds for the assembly and targeting of chromatin modifying complexes, unraveled a whole new layer of gene regulatory processes (reviewed in (Mercer and Mattick, 2013; Rinn and Chang, 2012)). Strikingly, the studies involving HMGA1 and Tip5 clearly illustrated the capability of AT-hook domains to mediate such RNA interactions in regulatory complex formation (Eilebrecht et al., 2011b; Zhou et al., 2009). For HMGA1 this results in a ribonucleoprotein complex consisting of the ncRNA 7SK, HMGA1 and P-TEFb playing a role in the transcriptional regulation of target genes (Eilebrecht et al., 2011a; Eilebrecht et al., 2011b). In Tip5 the deacytelation of a lysine residue within one of the AT-hooks leads to increased binding of the non-coding pRNA, which is a prerequisite for targeting of the protein to the rRNA promoter and the induction of epigenetic rRNA repression (Zhou et al., 2009). Additionally there are other AT-hook containing proteins, like the insulator protein CTCF, which are known to form regulatory ribonucleoprotein complexes with lncRNAs (Aravind and Landsman, 1998; Yao et al., 2010). This makes it tempting to speculate about a broader involvement of the AT-hook motif in such interactions.

In light of these facts, the high *in vitro* RNA binding activity of the eAT-hook proteins could suggest an analogue role in RNA protein complex formation within the cell. Therefore the AT-hook and eAT-hook motifs might not only play comparable roles in protein DNA interactions as tether and anchor of proteins to particular DNA structures and sequences (Aravind and Landsman, 1998), but at the same time may have a function as RNA binding modules in the formation of regulatory ribonucleoprotein complexes in a chromatin context.

4.3.3 Loss of the eAT-hook domain impairs shuttling of PTOV1 from the nucleus into the cytoplasm

PTOV1 was initially described as factor overexpressed in prostate carcinoma cells (Benedit et al., 2001). The exact function of the protein is so far not known, however several studies implicated a role in regulation of cell cycle progression and revealed a shuttling of the protein between cytoplasm and nucleus (Marqués et al., 2013; Santamaría et al., 2005; Santamaría et al., 2003). Since PTOV1 contains an N-terminal located eAT-hook domain and no other described DNA or RNA binding motifs and in addition supposedly possesses some interesting functions in the regulation of cell cycle progression, it was chosen for further *in vivo* investigations of eAT-hook functions.

The distinct effects of the missing eAT-hook domain on PTOV1 localization revealed in (3.3.4) clearly highlighted the importance of this motif for proper protein function. The results in (Figure 3.26) and (Figure 3.27) markedly demonstrate that without the eAT-hook domain GFP-PTOV1 presented an exclusively nuclear localization. The more diverse distribution shown for the full-length protein, which could roughly be divided in the three distinct states, nuclear, nuclear/cytoplasmic, and predominant perinuclear localization, was clearly lost upon the deletion of the eAT-hook motif.

This is noticeable since these findings clearly correlate with the previously published reports, illustrating a shuttling of PTOV1 between the cytoplasm and the nucleus. The studies revealed a similar sub-cellular localization pattern as presented here (Santamaría et al., 2005; Santamaría et al., 2003). Hence the solely nuclear localization of the PTOV1 ΔeAT protein suggested an impairment of the described shuttling mechanism caused by the lack of the eAT-hook domain. The possibility, that this effect was simply due to the loss of a nuclear export signal present in the eAT-hook motif, was excluded using a web based *in silico* prediction tool (la Cour et al., 2004).

Interestingly a very recent report depicted an association of PTOV1 to 40S ribosomes by its interaction with the receptor of protein C kinase 1 (RACK1), thereby mediating translation initiation of c-Jun (Marqués et al., 2013). The authors furthermore suggested a function for PTOV1 in ribonucleoprotein

shuttling, based in addition to their own data on a recently discovered structural similarity between the tandem PTOV blocks of PTOV1 and a so called SPOC domain (Bontems et al., 2011; Milbradt et al., 2011; Vojnic et al., 2011). This domain is found in the MINT/SHARP proteins, which are required for efficient mRNA nuclear export and contain N-terminal RNA binding motifs (RRM) in addition to the C-terminal SPOC domain (Zolotukhin et al., 2009). Strikingly, when taking into account the eAT-hook motif, the structural organization of PTOV1 would greatly mirror that of the MINT/SHARP protein. With the tandem PTOV blocks, and the N-terminal RNA binding eAT-hook. These facts propose a model were PTOV1 functions in the nuclear export and translational initiation of mRNAs, with the eAT-hook representing a RNA binding module necessary for a proper complex formation. This could in turn explain the impaired shuttling observed for the PTOV1 ΔeAT protein and would further underline the possible function of the eAT-hooks as novel nucleic acid interaction motifs.

4.3.4 eAT-hook – Summary, biological implications and further perspective

With the eAT-hook a new nucleic acid interaction domain could be described. It not only showed a widespread distribution throughout the human and mouse proteome, but the experiments also provided evidence for possible functions of this domain in DNA and RNA mediated processes. Therefore the domain could play a similar role as the AT-hook in chromatin organization, nuclear matrix binding and anchoring of proteins to structural DNA and also RNA elements. The importance of these features was highlighted just recently by a report linking mutations in the MeCP2 AT-hook 2 to the development of dysfunctional chromatin structures and Rett syndrome like symptoms in mice (Baker et al., 2013). Taken together these facts are elevating the importance of further characterizing the novel eAT-hook motif *in vitro* and *in vivo* to reveal putative functions in chromatin architecture and genome regulation of this domain.

Therefore additional experiments employing not only the isolated eAT-hook domains, but also the full-length versions of the candidate proteins are needed to

further verify the RNA and DNA binding characteristics in EMSA and MST assays. In addition the in vivo analysis of PTOV1 eAT-hook point mutations will be done, to validate the observations made so far. Also the *in vivo* analysis of additional candidate proteins could be considered to shed light on the functions of the eAT-hook domain.

5 Material and Methods

5.1 Materials

Unless otherwise stated, all common chemicals and materials were purchased from GE Healthcare (Freiburg), Merck (Darmstadt), Invitrogen (Karlsruhe), Fermentas (St. Leon-Rot), New England Biolab (Frankfurt am Main), Promega (Mannheim), Roche (Mannheim), Roth (Karlsruhe), Serva (Heidelberg), Bio-Rad (Munich), Stratagene/Agilent (Waldbronn), Sigma-Aldrich (Munich) and Qiagen (Hilden).

5.1.1 Technical devices

Description	**Supplier**
Agarosegel UV imaging system	GelMax, Intas
Sonifier 250	Branson
Fluorescence – Image Reader FLA-3000	Fujifilm
Chemiluminoscence – Image Reader LAS-3000	Fujifilm
Centrifuge Centrikon T-324	Kontron Instruments
PCR machine	Peqlab
PCR machine verity	Applied biosystems
PCR machine (old)	Perkin Elmer
Peristaltic– Pump LKB – P1	GE Healthcare
Real Time qPCR machine	Qiagen/Corbett Research, Rotor Gene, RG-3000
Table top centrifuge	Eppendorf
Trans – Blot® SD Semi-dry transfer cell	BioRad

Description	Supplier
Thermomixer Compact	Eppendorf
Ultracentrifuge Centrikon T-1170,	Kontron Instruments
Ultrospec 3100 pro	Amersham Biosciences
Uvikon Spectrophotometer 922	Kontron Instruments
Monolith NT.115	NanoTemper technologies
Monolith NT.015T	NanoTemper technologies
Purelab Ultra	ELGA
Axiovert 200M + ApoTome 2	Zeiss
Qubit® 2.0 Fluorometer	Invitrogen
Nanodrop® ND – 1000 Spectrophometer	peQLabBiotechnologie GmbH

5.1.2 Software tools

Software	Application	Supplier
Geneious	*in silico* cloning tool, organization software	Biomatters Ltd.
LabLife	lab organization tool	LabLife Software (www.lablife.org)
Multigauge V3.1	LAS reader viewer software	FujiFilm
Kaleidograph 4.1	graphical analysis software	Synergy software
RNAfold	RNA secondary prediction tool	(http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi)
RotorGene Analysis software	qPCR data analyis	Qiagen
Axiovision	Zeiss microscope software	Zeiss
NEB double digest finder	restriction digest organization tool	New England Biolabs (www.neb.com)
Triplexator	DNA triplex structure prediction software	http://acb.qfab.org/acb/triplexator

Software	Application	Supplier
SMART	modular protein domain prediction tool	http://smart.embl-heidelberg.de/

5.1.3 Chemicals and consumables

Description	Supplier
1.5 ml and 2 ml micro centrifuge tubes	Eppendorf/Sarstedt
15 ml and 50ml tubes	Sarstedt
Agarose (ME, LE GP and low melting)	Biozym
Ammonium acetate	Merck
Ammonium sulfate	Merck
Bacto Agar	BD
Bacto Peptone	BD
Bacto Tryptone	BD
Barrier food wrap	Saran
Boric acid	Merck
Bromphenolblue	Serva
BSA 98%	Sigma
BSA purified	NEB
ß-Mercaptoethanol	Sigma
Coomassie G250	Serva
Cover slips, 22x22 mm	Roth
Cryovials	Roth
Dialysis membrane	Roth
DAPI	Sigma
DMSO	Sigma
dNTP mix	NEB/Qiagen
DTT	Roth
EDTA	Sigma
EGTA	Sigma
Ethidium bromide	Sigma

Description	Supplier
ETOH tech., p.a.	Merck
FCS dialyzed	Sigma
Filter paper Whatman 3MM	Whatman
Filter tips	Roth
Filter unit	Nalgene, 0.2 µm filter holes
Fugene HD	Promega
Glass pipettes 5 ml and 10 ml	Hirschmann®
Glassware	Schott
Glycerin	Merck
Glycogen	Roche
Glutathione, reduced	Sigma
Glutathione Sepharose 4B	GE Healthcare
HEPES	Roth
IPTG	Roche
Isopropanol p.a.	Merck
Laboratory film	Parafilm®
Magnesium acetate	Merck
Magnesium chloride	Merck
Methanol p. a.	Merck
Microscope slides SuperFrost, white, cut edges	Roth
Milk powder	Sucofin
Monolith NT115 standard treated capillaries	NanoTemper Technologies
Nitrocellulose membrane (GSWP, 0,22µM)	Millipore
NP40	Sigma
Orange G	Sigma
Paraformaldehyde	Merck/Roth
Pasteur pipettes	Brand
PCR-reaction tubes 0.2 ml	Biozym
Petridishes and tissue culture plates	Greiner, Sarstedt

Description	Supplier
Pipes	Sigma
Pipette tips	Gilson, Brand
PMSF	Sigma
Potassium chloride	Merck
Protein gel cassettes (disposable)	Invitrogen
Rotiphorese Acrylamid-Bisacrylamidmix	Roth
SDS	Serva
ß-Mercaptoethanol	Sigma
Sodium chloride	VWR
Sodium dodecyl sulfate (SDS)	Roth
Sodium phosphate mono-sodium salt	Merck
Sodium phosphate di-sodium salt	Merck
Sucrose	Roth
Sybr Safe	Invitrogen
Sybr Green	Qiagen
Syringes and accessories	Roth
TCEP	Sigma
TEMED	Roth
Tris	Invitrogen
Triton X-100	Sigma
Trypsin/EDTA (TC)	PAA
Tween 20	Sigma
Ultracentrifugation tubes for SW40 rotor	Beckman Coulter
Urea	Merck
Vectashield mounting medium	Vector
Yeast extract	BD

5.1.4 Standard Solutions

Stock solutions and buffers were prepared according to standard protocols (Roche, 2011; Sambrook and Russell, 2001). Protease Inhibitor Cocktail (Roche) was freshly added. Common solutions are listed below.

Buffer	**Composition**
Annealing buffer	20 mM Tris-HCl pH 7.4 2 mM $MgCl_2$ 50 mM NaCl
Phosphate Buffered Saline (PBS)	140mM NaCl 2.7mM KCl 8.1mM Na2HPO4 1.5mM KH2PO4 pH adjusted to 7.4 with HCl
PBS 4% PFA solution	PBS with 4% (w/v) Paraformaldehyde
PBST buffer	PBS with 0.1% Tween 20
PBST milk buffer	PBS with 0.1% Tween 20 5% (w/v) milk powder
PBST 4% BSA buffer	PBS with 0.1% Tween 20 5% (w/v) BSA
TBE buffer	90mM Tris 90mM Boric acid 2mM EDTA
TE buffer	10mM Tris-HCl pH 7.6 1mM EDTA
DNA sample buffer (10x)	50% glycerol 50mM Tris-HCl pH 7.6 10mM EDTA 0.05% (w/v) bromophenol blue, xylene cyanol and Orange G

Buffer	Composition
Orange G loading dye (10x)	50% 102lycerin 10mM EDTA 0.05% (w/v) Orange G
SDS-protein sample buffer (5x)	300mM Tris-HCl pH 6.8 10% (w/v) SDS 50% glycerol 5% ß-Mercaptoethanol 0.2% (w/v) bromphenol blue
SDS-PAGE stacking buffer (4x)	0.5 M Tris-HCl 0.4% SDS, pH 6.8 with HCl
SDS-PAGE separating buffer (4x)	1.5M Tris-HCl 0.4% SDS, adjust to pH 8.8 with HCl
SDS-PAGE running buffer	192mM glycine 25mM Tris 0.1% (w/v) SDS
Transfer buffer (Towbin)	25mM Tris 192mM Glycin 20% Methanol 0.05% SDS
Coomassie staining solution	45% water 45% methanol 10% acetate acid
Lysis buffer GST purification	PBS 1mM TCEP 0.5% TritonX – 100 Protease Inhibitor Cocktail

Buffer	**Composition**
Wash buffer GST purification	PBS 1mM TCEP 0.5% Triton X – 100 Protease Inhibitor Cocktail
Elution buffer GST purification	20mM Tris, pH 8.0 20mM Glutathione Protease Inhibitor Cocktail
Transfer buffer (Towbin)	25mM Tris 192mM Glycin 20% Methanol 0.05% SDS
CSK buffer	10 mM Pipes, pH 6.8 100 mM NaCl 300 mM sucrose 3 mM $MgCl_2$ 1 mM EGTA 1mM TCEP 0.5% Triton X-100 Protease Inhibitor Cocktail
CSK buffer high salt	10 mM Pipes, pH 6.8 100 mM NaCl 300 mM sucrose 2 M NaCl 3 mM $MgCl_2$ 1 mM EGTA 1mM TCEP 0.5% Triton X-100 Protease Inhibitor Cocktail

Buffer	Composition
CSK extraction buffer	10 mM Pipes, pH 6.8
	250 mM $(NH_4)_2SO_4$
	300 mM sucrose
	3 mM $MgCl_2$
	1 mM EGTA
	1mM TCEP
	0.5% Triton X-100
	Protease Inhibitor Cocktail
CSK digestion buffer	10 mM Pipes, pH 6.8
	50 mM NaCl
	300 mM sucrose
	3 mM $MgCl_2$
	1 mM EGTA
	1mM TCEP
	0.5% Triton X-100
	Protease Inhibitor Cocktail
Urea buffer	10 mM Tris-HCl, pH 8.0
	100 mM NaH_2PO_4, adjusted to pH 8.0
MST protein buffer (protein analysis)	50 mM Hepes, pH 7.4
	5 mM $MgCl_2$
	100 mM NaCl
	0.05% (v/v) NP40
EMSA protein buffer (protein analysis)	10 mM Tris-HCl, pH 8.1
	180 mM NaCl
	1 mM $MgCl_2$
	0,01 % (w/v) BSA
	5 % glycerol
TA triplex buffer	40 mM Tris-Acetate, pH 7.4
	10 mM MgOAc

Buffer	Composition
EMSA triplex buffer (triplex analysis)	40 mM Tris-Acetate, pH 7.4 10 mM MgOAc 5% glycerol
MST triplex buffer (triplex analysis)	40 mM Tris-Acetate, pH 7.4 10 mM MgOAc 0.05% NP40
MST triplex buffer Mg^{2+}-dependency (triplex analysis)	40 mM Tris-Acetate, pH 7.4 0-50 mM MgOAc (as indicated) 0.05% NP40
MST triplex buffer pH dependency (triplex analysis)	10 mM Tris-HCl, pH 6.4/7.4/8.4 25 mM NaCl 10 mM $MgCl_2$ 0.05% NP40
MST triplex cell buffer (triplex analysis)	10 mM Hepes, pH 7.4 10 mM NaCl 140 mM KCl 0/5/10 mM $MgCl_2$ (as indicated) 0.05% NP40

5.1.5 Enzymes

Enyzme	Supplier
Antartic Phosphatase	New England Biolabs
DNase I (RNase free)	Roche
Herculase II Fusion Enzyme	Agilent
Hot Star Taq DNA Polymerase	Qiagen
Klenow enzyme	New England Biolabs
Proteinase K	Sigma
Restriction endonucleases	New England Biolabs
RNase A	Roche, Invitrogen

Enyzme	Supplier
T4 DNA Ligase	New England Biolabs

5.1.6 Kits

Kit	Supplier
Herculase II Fusion Enzyme dNTP Combo PCR Kit	Agilent
QIAquick PCR purification Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
Pure Link™Quick Gel Extraction Kit	Invitrogen
QIAprep Spin Miniprep Kit	Qiagen
Pure Link™HiPure Plasmid Maxiprep Kit	Invitrogen
Quanti Tec SYBR Green PCR Kit	Qiagen
Qbit dsDNA BR Kit	Invitrogen
Qbit dsDNA HS Kit	Invitrogen
Qbit protein Kit	Invitrogen
Super signal WEST Dura WB Kit	Pierce

5.1.7 Standard DNA and protein marker

Gene Ruler 1kb plus DNA ladder	Fermentas (SM1333)
Prestained Protein – Marker IV	Peqlab
Prestained protein marker page ruler	Fermentas

5.1.8 Protease inhibitors, RNase inhibitors and antibiotics

Substance	Supplier
Ampicillin	Sigma
Kanamycin	Sigma
Chloramphenicol	Roth
Rnasin	Promega
Penicillin/Streptomycin	Invitrogen/ Gibco
Protease Inhibitor Cocktail	Roche

5.1.9 Bacterial cell lines and media

DH5alpha	
description	general DNA plasmid propagation
resistance	none
genotype	F- φ80lacZΔM15 Δ(lacZYA-argF) U169 endA1 recA1 hsdR17 (rk-,mk+) supE44 thi-1 gyrA96 relA1 phoA
Rosetta 2 (DE3) pLysS	
description	protein expression
resistance	chloramphenicol
genotype	F-ompT hsdSB (rB-mB-) gal dcm (DE3) pLysSRARE2 (CamR)
XL1 Blue	
description	F´episome, general DNA plasmid propagation, blue/ white screening
resistance	tetracycline
genotype	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F'proAB lacIqZDM15

Luria-Bertani (LB) medium

1.0% (w/v) Bacto-Tryptone

1.0% (w/v) NaCl

0.5% (w/v) Bacto-Yeast extract

→ Adjust the pH to 7.0 with 10 M NaOH

The medium was autoclaved for 20 min at 120°C and the appropriate antibiotics were added in standard concentrations (Roche, 2011) prior to usage. For preparing plates the LB medium was mixed with 1.5% agar and the appropriate antibiotics were added after cooling down to 60°C.

SOB medium

2% (w/v) Bacto-Tryptone

10 mM NaCl

0.5% (w/v) Bacto-Yeast extract

2.5 mM KCl

10 mM MgCl2*

→ Adjust the pH to 7.0 with 10 M NaOH

* add before use

The medium was autoclaved for 20 min at 120°C.

5.1.10 Eukaryotic cell lines and media

Hek293	
cell type	human embryonic kidney
growth conditions	37°C, 5% CO_2, DMEM-GlutaMAX, 10% FCS
supplier	ATCC, cells were received second hand from A.Nemeth
HeLa	
cell type	Human cervix carcinoma
growth conditions	37°C, 5% CO_2, DMEM-GlutaMAX, 10% FCS
supplier	ATCC, cells were received second hand from A.Nemeth

Imr90	
cell type	Human fetal lung fibroblast
growth conditions	37°C, 5% CO_2, DMEM-GlutaMAX, 10% FCS
supplier	ATCC, cells were received second hand from A.Nemeth

Growth medium for human cell lines

450ml	DMEM
50ml	FBS (aliquoted in -20, 10% total)
5ml	1:100 Penicillin-Streptomycin (aliquoted in -20, 1% total)

Stored at 4°C

5.1.11 Antibodies

Antibodies	**Supplier**	**Description**	**Dilution**
Rb-α-B23	Santa Cruz	rabbit, polyclonal	IF 1:100
M-α-Fibrillarin	Novus Biologicals	mouse, monoclonal	IF 1:50
M-α-RPA194	Santa Cruz	mouse, polyclonal	IF 1:200
Rb-α-Lamin AC	Sanat Cruz	rabbit, polyclonal	IF 1:200, WB 1:500
M-α-Lamin AC	Sanat Cruz	mouse, monoclonal	IF 1:50
Rb-α-mTip5-N1-18	Grummt lab, DKfZ	rabbit, polyclonal	IF 1:25 WB 1:100
G-α-Rb-Cy3	Jackson	goat, polyclonal	IF 1:200
G-α-M-Cy3	Jackson	goat, polyclonal	IF 1:500
G-α-M-Alexa 488	Molecular Probes	goat, polyclonal	IF 1:500
GFP-Booster Atto488	Chromotek	specific GFP binder	IF 1:200

5.1.12 Oligonucleotides

All oligonucleotides used in this study were purchased from Eurofins MWG Operon and diluted in MilliQ-water to a final concentration of 100µM.

Unmodified DNA Oligonucleotides:

Name	Application	Sequence (5'-3')
Enhancer triplex for (en3 D_for)	triplex formation (EMSA/MST)	TCT TTT TTT TTT TTT TTC TTT TTT CCT CC
Enhancer triplex rev (en3 D_rev)	triplex formation (EMSA/MST)	GGA GGA AAA AAG AAA AAA AAA AAA AAA GA
Enhancer triplex TFO DNA (en3)	triplex formation (EMSA/MST)	CCT CCT TTT TTC TTT TTT TTT TTT TTT CT
GFP control TFO DNA	triplex MST control	AGT ACA ACT ACA ACA GCC ACA ACG TCT AT
PTOV1 EcoRI For	cloning	GTT TGA ATT CGG CAT GGT CCG TCC GCG CCG T
PTOV1 XbaI Rev	cloning	CAA ATC TAG ACT ACC CCC CCA TCC CTC GTT GC
PTOV1 Seq rev	sequencing	AGC AGC TTG TTG CTG AGC
pEGFPC2 seq for	sequencing	CAT GGT CCT GCT GGA GTT CGT G
pEGFPC2 seq rev	sequencing	TTT ATG TTT CAG GTT CAG GG
Mix control for	protein EMSA	CAC GCG CTC GCG CAC GCG CTC GCG CAC GCG CTC
Mix control rev	protein EMSA	GAG CGC GTG CGC GAG CGC GTG CGC GAG CGC GTG
rDNA IGS for	protein EMSA	TGG ATC TTT TTT TTT TTT TTT CTT TTT TCC TCC A

Name	Application	Sequence (5'-3')
rDNA IGS rev	protein EMSA	TGG AGG AAA AAA GAA AAA AAA AAA AAA AGA TCC A

Modified DNA Oligonucleotides:

Name	Application	Modification	Sequence (5'-3')
Enhancer triplex rev (en3 D_rev)	triplex formation MST	5'-Cy5	GGA GGA AAA AAG AAA AAA AAA AAA AAA GA
Enhancer triplex for (en3 D_for)	triplex formation EMSA	5'-Cy5	TCT TTT TTT TTT TTT TTC TTT TTT CCT CC
Enhancer triplex TFO DNA (en3)	triplex formation (EMSA/MST)	5'-FAM	CCT CCT TTT TTC TTT TTT TTT TTT TTT CT
Mix control	protein MST	5'-Cy3	CAC GCG CTC GCG CAC GCG CTC GCG CAC GCG CTC

Unmodified RNA Oligonucleotides:

Name	Application	Sequence (5'-3')
GFP control TFO RNA	triplex MST control	AGU ACA ACU ACA ACA GCC ACA ACG UCU AU
Enhancer triplex TFO RNA (en3)	triplex formation (MST)	CCU CCU UUU UUC UUU UUU UUU UUU UUU CU

Modified RNA Oligonucleotides:

Name	Application	Modification	Sequence (5'-3')
Enhancer triplex TFO RNA (en3)	triplex formation (EMSA)	5'-FAM	CCU CCU UUU UUC UUU UUU UUU UUU UUU CU
Stemloop RNA 1	protein MST	5'-FAM	AGU UCC AUG AUG UUU UCA AAC UCU AUU ACC UAC AUU AUU UGA AC
Stemloop RNA 2	protein MST	5'-Cy5	GAG AAA CCG UGU AAA CAA AAC AUA ACU GAG AGG

5.1.13 Plasmids

Plasmid	Insert	Created by
pEGFP-C2 Tip5 AT12	AT-hooks 1+2 of Tip5	Michael Filarsky
pEGFP-C2 Tip5 MUT AT12	Mutated AT-hooks 1+2 of Tip5	Michael Filarsky
pEGFP-C2 Tip5 TAM AT12	TAM domian and AT-hooks 1+2 of Tip5	Michael Filarsky
pEGFP-C2 Tip5 TAM MUT AT12	TAM domian and mutated AT-hooks 1+2 of Tip5	Michael Filarsky
pMK GPBP1/PIF1/PTOV1-eAT	eAT-hook domains of GPBP1, PIF1 and PTOV1	Geneart
pMK GPBP1/PIF1/PTOV1-eAT MUT	Mutated eAT-hook domains of GPBP1, PIF1 and PTOV1	Geneart
pGEX4T3 Tip5eAT	Tip5 eAT-hook domain	Michael Filarsky
pGEX4T3 GPBP1eAT	GPBP1 eAT-hook domain	Michael Filarsky
pGEX4T3 PTOV1eAT	PTOV1 eAT-hook domain	Michael Filarsky
pGEX4T3 Tip5eAT MUT	Mutated Tip5 eAT-hook domain	Michael Filarsky

Plasmid	Insert	Created by
pGEX4T3 GPBP1eAT MUT	Mutated GPBP1 eAT-hook domain	Michael Filarsky
pGEX4T3 PTOV1eAT MUT	Mutated PTOV1eAT-hook domain	Michael Filarsky
pGEX4T3 Tip5eAT core	Tip5 eAT-hook domain core motif	Regina Gröbner-Ferreira
pGEX4T3 GPBP1eAT core	GPBP1 eAT-hook domain core motif	Regina Gröbner-Ferreira
pGEX4T3 PTOV1eAT core	PTOV1 eAT-hook domain core motif	Regina Gröbner-Ferreira
pEGFP-C2 PTOV1	PTOV1 coding sequence	Michael Filarsky
pEGFP-C2 PTOV1ΔeAT	PTOV1 coding sequence missing the eAT-hook domian	Michael Filarsky

5.2 Methods

All molecular work like preparation and transformation of chemically-competent bacteria with DNA, amplification of plasmid DNA in *E. coli*, purification, concentration determination, restriction enzyme digestion, ligation of DNA fragments, analysis of DNA on agarose and polyacrylamide gels, and amplification of the DNA by the polymerase chain reaction (PCR) was performed following the standard protocols (Sambrook and Russell, 2001). Bacteria were cultured in LB medium and the appropriate antibiotics were added corresponding to the plasmid-encoded resistance. The plasmid DNA was isolated with plasmid purification kits (Invitrogen/Qiagen). Isolation of DNA fragments from agarose gels was performed using the Qiagen Gel Extraction kit.

5.2.1 Working with DNA

5.2.1.1 Determination DNA quality and quantity

The NanoDrop ND1000 spectrophotometer (Peqlab) was used to determine DNA concentration and purity by absorption measurements at 260 nm. For highly accurate determinations of small DNA concentrations prior of qPCR experiments the Qubit fluorometer (Invitrogen) was utilized. The technique is based on fluorescent dyes, which only emit a signal when bound to their specific partner, ensuring a high specificity and sensitivity.

5.2.1.2 Polyacrylamide and agarose gel electrophoresis

Agarose gel electrophoresis was generally performed with gels containing 0,8 - 1,2% agarose in 1X TBE buffer, 1:10000 SYBR Safe (Invitrogen), in 1x TBE running buffer at a constant voltage of 100-120V. DNA standard marker and samples were supplemented with 10x DNA loading dye. In contrast to agarose gel electrophoresis, DNA was separated by polyacrylamide gel electrophoresis (PAGE) in 0.4 X TBE at 100V. In order to remove unpolymerized acrylamide the gel was

pre-run for 1 h at 80V. For visualization the gel was stained after the gel run in 0,4x TBE containing ethidiumbromide (0,5mg/ml) for 15 min and washed twice with water for 10 min each.

5.2.1.3 Restriction digest

Restriction enzymes were used at reaction conditions according to the manufacturer's recommendations in regard to buffer, addition of BSA, incubation time and temperature (see www.neb.com). For an analytical digest 0.1 - 1 µg DNA was incubated with 5 units of the respective restriction endonuclease in a total volume of 20 µl. Preparative restriction digests were done with 10 µg DNA using 50 units of restriction endonuclease in a total volume of 50 µl. To check the completion of the digest, the DNA was separated using 0.8 - 2.0% TBE-agarose gels supplemented with SYBR Safe (Invitrogen). Digests for small fragments and low quantities of DNA were stained with ethidiumbromide (0,5mg/ml) for 15 min and washed twice with water for 10 min each.

5.2.1.4 DNA ligation

Ligation of sticky or blunt ended DNA fragments was performed using the T4 DNA ligase (NEB). The molar ratio of insert to vector was kept in a range between 3 and 5 to 1. The ATP containing 10x ligase buffer was stored in aliquots at -20°C. The ligation reaction was performed in a total volume of 10 µl at room temperature for 1 h or in a cold room at 16°C overnight.

5.2.1.5 Polymerase Chain Reaction

For PCR reactions used in cloning procedures the Herculase II fusion enzyme (Agilent) was used according to the manufacturers instructions. For reactions containing template DNA with a high GC content the reaction buffer was supplemented with DMSO. The PCR protocol was adjusted for each reaction concerning annealing temperature and time as well as elongation time. The general outline is depicted in the table below. It is recommended to include a water and a positive control.

Cycle step	Temperature	Time	Number of cycles
initial denaturation	95	5 min	1
denaturation annealing extension	95 50-60 (depending on the used primer) 72	30 sec 30 sec 1min/1kb	30
final extension	72	5min	1

5.2.1.6 Colony PCR

The colony PCR was employed as a fast and quick method to screen for positive clones after DNA ligation. Therefore a pipette tip was dipped into a bacterial colony on an agar plate that was to be tested for the presence of the insert. The adhering cells were smeared at the bottom of an empty 0,2 ml PCR tube and subsequently gridded on fresh LB plates containing the necessary antibiotics. Then, 50µl of colony PCR master mix were added to each PCR tube and the reactions were subjected to the colony PCR program. The PCR program contains a long initial denaturation step to open up the bacterial cells. For these experiments a lab purified Taq-Polymerase enzyme was used. Afterwards, the PCR reactions were analyzed on an agarose gel for presence of the amplicon.

Colony PCR master mix

50 µl	1 x (µl)
Primer for 10µM	0,5
Primer rev 10µM	0,5
dNTP´s 10 mM	1
Taq-Puffer 10x	5
Taq-Polymerase	1
H2O	42
Total volume	**50**

Colony PCR program

Cycle step	Temperature	Time	Number of cycles
initial denaturation	95	10 min	1
denaturation annealing extension	95 50-60 72	40 sec 40 sec 1 min/1kb	35
final extension	72	5 min	1

It is recommended to include a water and positive control.

5.2.1.7 Annealing of oligonucleotides

Equimolar amounts of sense and complementary antisense oligonucleotides were diluted to the desired concentration in 1x annealing buffer. The samples were incubated at 95°C for 10 min in a thermo block. Afterwards the block was switched of and the reactions were slowly cooled down to room temperature over 1-2 h. For annealings with fluorescently labeled oligonucleotides, the unlabeled strand was used in approximately 5% excess, to guaranty a complete annealing of all labeled oligonucleotides. This was done to avoid background signals from single stranded fluorescently labeled oligonucleotides in the MST assays.

5.2.1.8 Real-time quantitative PCR

Quantitative real-time PCR was used to determine the amount of specific DNA fragments within nuclear matrix preparations with high accuracy. The individual reactions were thereby performed in technical triplicates. The amount of DNA was determined after each cycle by measuring the fluorescent signal intensity of SYBR green. The dye intercalates in the DNA double strand, therefore the detected fluorescence is proportional to the amount of double stranded DNA within the sample. During a PCR reaction, the original amount of DNA Z(0) is amplified exponentially, until either the polymerase fails or one of the other reagents is

consumed. Therefore it is possible to calculate the total amount of DNA Z(n) according to the formula:

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

Z(0) represents the original amount of DNA, n resembles the number of cycles and E the efficiency of the reaction (1<E<2). The amplification efficiency can be evaluated from the slope of a semi-logarithmic plot of fluorescence against cycle number when measuring a standard curve for a primer pair. Thus when E is known for a primer pair, the relative amounts of specific fragments can be determined. The detected fluorescence is plotted against the cycle number 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 (Ct). This Ct 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 μL. The reaction contained 4 μL of DNA sample (concentration 10 ng/μL or less) and 16 μL master mix. This master mix contained 4 pmol of forward and reverse primer each, 0.25 μL of a 1:400,000 SYBRR-Green stock solution, 0.4 U HotStarTaq and premix. This premix consists of $MgCl_2$ (final concentration: 2,5mM), (final concentration 0.2mM), and 10xPCR buffer (final concentration 1x). qPCR was performed using the Rotor-Gene RG3000 system from CORBETT Research/Qiagen, SYBR-Green was excited at 480 nm, fluorescence was recorded at 510 nm.

The qPCR cycler program is listed below:

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

Standard curves were measured in serial 1:4 dilutions of a 10 ng (DNA)/μL stock solution till a dilution of 1:1,024. After evaluating the standard curves, the efficiencies of all used primer pairs were higher than 96%.

5.2.1.9 Triplex EMSA

For the fluorescently labeled triplex EMSA 15% polyacrylamide gels were prepared using 1x TA triplex buffer. Gels were run in 1x TA triplex buffer at room temperature. Before sample loading the gels were pre-run for 30 min at 100V.
The triplex annealing for the double colored EMSA was performed in sterile 0.2 ml PCR tubes. The total reaction volume was 20 µl. A master mix was prepared using a 5x stock of the EMSA triplex buffer, containing 1 pmol of Cy5-labeled Enhancer duplex DNA for each sample. To each reaction the appropriate TFO was added. Therefore rising amounts of 1 pmol, 2 pmol and 4 pmol of TFO were used and a negative control containing no TFO was also prepared. The samples were vortexed and incubated in the thermomixer at 37°C for 1 h. After this the block was switched off and the sample were slowly cooled down to room temperature and then immediately loaded onto the gel. The gel run was performed at 100V for 2-3 h. Afterwards the gels were imaged in the FLA3000 fluorescent imager with the matching filter sets.

5.2.2 Protein biochemical methods

Protein analysis was performed according to the standard protocols (Sambrook and Russell, 2001). Generally, proteins were kept on ice (4°C), in the presence of protease inhibitor cocktail (Roche).

5.2.2.1 Determination of protein concentrations

Protein concentrations were determined using the Qubit fluoromter (Invitrogen), which is based on fluorescently labeled dyes, emitting signals only when bound to

protein, providing highly accurate concentrations. In addition the concentration of purified proteins was estimated according to protein standards with a known concentration (e.g. BSA) in SDS-PAGE combined with Coomassie Blue staining.

5.2.2.2 Denaturing SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein separation according to the molecular weight using discontinuous SDS polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the method of Laemmli (1970). Separating and stacking gels were prepared following standard protocols with ready-to-use polyacrylamide solutions from Roth (Rotigel, 30 %, 49:1). For electrophoresis, protein samples were mixed with Laemmli SDS-PAGE sample buffer, heat-denatured for 5 min at 95°C and directly loaded onto the gel. Proteins were separated at 35mA, until the dye front reached the bottom of the gel. The molecular weight of proteins was estimated pre-stained protein markers (PAGE ruler, Fermentas). Following electrophoresis, proteins were stained with Coomassie Brilliant Blue.

5.2.2.3 Coomassie blue staining of protein gels

Polyacrylamide gels were stained for approximately 1 h on a slowly rocking platform with Coomassie staining solution (0.1% Coomassie Blue R in 10% acetic acid, 45% methanol). Gels were destained in water by repeated heating in a microwave, until the protein bands were clearly visible.

5.2.2.4 Semi dry Western Blot and immunodetection

Following separation by SDS-PAGE, proteins were transferred to PVDF membranes using the Bio-Rad 'Trans-Blot SD Apparatus' for 1h at 25V. In case more than one gel was blotted in parallel the transfer was prolonged (up to 2 h). For protein transfer, the PVDF membrane was activated in 100% methanol and then incubated together with the gel in Towbin transfer buffer for 5- 10min. The gel was placed between 3 gel-sized Whatman papers soaked in transfer buffer at the bottom and the PVDF membrane and 3 gel-sized Whatman papers soaked in transfer buffer on top (3MM Whatman pieces each). After transfer, the PVDF membrane was

incubated in PBST-milk for 15min -1h in order to reduce non-specific background. For immunodetection the PVDF membrane was incubated with the appropriate antibody in PBST-milk at 4°C over night, slightly shaking. Afterwards the membrane was washed three times in PBST (15 -20min each) and incubated for 1 h with a horseradish peroxidase-coupled secondary antibody in PBST-milk. After 3 additional washes (15 min each) in PBST, antigen-antibody complexes were detected using the Super signal WEST Dura WB Kit (Pierce) and imaged with the Image Reader LAS-3000 (Fujifilm). All steps were performed at room temperature.

5.2.2.5 Protein EMSA experiments

The protein DNA interactions were analyzed by EMSA experiments. Therefore 7.5% 0.4x TBE polyacrylamide gels were used. Gels were run in 0.4x TBE at 100V, prior to sample loading the gels were pre-run for 30 min.

For each protein used in the experiments a 2:1 dilution series was prepared in 20 mM Tris-HCl pH 8, 10% glycerol. The total reaction volume for each sample was 10 µl, 6 µl of a DNA master mix and 4 µl of the appropriate protein dilution. Therefore a master mix was prepared using a 2x stock of the EMSA protein buffer with 1.25 pmol of target DNA for each sample. All reactions were set up on ice and mixed by pipetting. The samples were incubated on ice for 15 min and then directly loaded onto the gels. Gels were run at room temperature for 30 min at 100V. Afterwards gels were stained in ethidiumbromide (0.5 mg/ml) for 30 min and then washed in ddH_2O twice, for 10 min.

5.2.3 *E. coli* culture and methods

5.2.3.1 Liquid culture

For plasmid preparations a single colony was picked with a sterile tip from an agar plate, inoculated into LB or SOB medium supplemented with the respective antibiotic and shaken overnight at 37°C at 180 rpm. For standard Mini-prep

(Qiagen) preparations, 5ml cultures were used. For expression cultures a small pre-culture was inoculated and incubated over night at 37°C to an OD600 of 3-5. This culture was then used to inoculate the expression culture to an OD600 of 0.05.

5.2.3.2 Glycerol stock

For long term storage of bacterial cultures and convenient handling of frequently used strains, 850µl of a stationary liquid culture are mixed with 400µl of 50% sterile glycerol and frozen at –80°C.

5.2.3.3 Transformation of chemically competent bacteria

For transformation 50µl of chemically competent bacteria were thawed on ice and either 10 ng of purified plasmid DNA or 5 µl of ligation reaction were added. The suspension was mixed by gently tapping the tube and incubated on ice for 15 to 30 min. Cells were transformed by a heat-shock at 42°C for 45 seconds, then cooled down on ice for 5 min. 450 µl of LB or SOB medium without antibiotics was added and the bacteria were incubated at 37°C shaking at 600 rpm for 30 to 60 min. For ampicillin, cells were incubated for 30 min, for chloramphenicol and kanamycin, one hour of incubation was performed. 50µl and 150µl of the mixture were plated on agar plates containing the appropriate antibiotics. Plates were incubated at 37°C overnight.

5.2.4 Expression and purification of recombinant proteins from *E. coli*

5.2.4.1 Protein expression for GST-purification

Rosetta 2 (DE3) pLysS cells were transformed with the respective expression plasmids coding for the proteins of interest.

For 50 ml expression volume a 5 ml pre-culture was inoculated and incubated at 37°C, shaking, over night. From this culture the 50 ml expression culture was

inoculated to an OD600 of 0.05 and incubated at 37°C, shaking, until an OD600 of 0.3-0.6 was reached. Protein expression was induced with 0.5 mM IPTG and the culture was incubated for 3-4 h at 37°C, shaking. After this expression time the culture was centrifuged for 10min at 4000rpm at 4°C, supernatant was removed and the pellet was frozen in liquid N_2 and stored at -80°C until purification.

5.2.4.2 GST-purification cell lysis

For protein purification via the GST-tag the bacterial pellets were resuspended in 1 ml of GST-lysis buffer supplemented with protease inhibitor cocktail (Roche) on ice. For efficient cell lysis the pellets were refrozen in liquid N_2 and thawed to room temperature, this was repeated two times. Subsequently, the cell lysate was sonified using a Branson digital sonifier 250D (big tip), 10 times for 30 seconds at 50% amplitude and 50% duty cycle, with 30 seconds pause on ice. 40 µl of the cell lysate was collected and mixed with 10 µl 5x Laemmli buffer for analysis with SDS-PAGE (CL). The remaining lysate was centrifuged at 13.000 rpm, 4°C for 1 h in a tabletop centrifuge. The remaining supernatant was used for further purification and the pellet was resuspended in 0.5 ml lysis buffer and 40 µl of the solution were collected for SDS – PAGE (P).

5.2.4.3 GST purification of recombinantly expressed protein

Per sample 150 µl of GST-Sepharose 4B beads were washed three times in 1 ml 1xPBS by resuspending and spinning down in a tabletop centrifuge at 500 g for 3 min. The sample was added to the washed beads and incubated in an overhead rotator for 1 h at room temperature. Again the beads were spun down and 40 µl of the supernatant was collected for SDS-PAGE (NB). The rest was discarded and the beads were washed three times with 1 ml GST-wash buffer at room temperature for 5 min. The protein was eluted at room temperature with 150 µl GST-elution buffer for 60 min in an overhead wheel. Again beads were spun down and 20µl of the solution was collected for SDS – PAGE (E1). To the remaining supernatant solution 10% (w/v) glycerol was added and the fraction was frozen in liquid N_2 and stored at -20°C. This was repeated once (E2). To test the successful

purification 5-10 µl of the samples collected at each step were subjected to SDS-PAGE.

5.2.5 Mammalian cell culture and methods

5.2.5.1 General

Work with mammalian cell lines was carried out according to standard protocols. All work was done under a sterile hood in laminar flow and all solutions and consumables were purchased sterile. The working space, gloves and devices were thoroughly wiped with 70% ethanol before use.

5.2.5.2 Propagation of mammalian cells

All cell lines used were cultivated in DMEM supplemented with 10% FCS and 1% penicillin/streptomycin. Cells were incubated at 37°C in a humified incubator with 5% CO_2. The medium was stored at 4°C and pre-warmed to 37°C in a water bath before usage. The medium of the cultures was changed every 2-3 days depending on the doubling time of the cells.

At an estimated confluency of 80% the cells were split. Therefore the medium was removed and a trypsin/EDTA solution was added to the cells. The cells were incubated at 37°C for 5 min and the cells still detached were removed from the surface by slightly tapping the culture dish. The trypsin reaction was stopped by adding culture medium at three times the volume of trypsin/EDTA. The appropriate volume of cells was then transferred to a new flask or culture disk with fresh growth medium and again incubated at 37°C, 5% CO_2.

5.2.5.3 Freezing

For cryopreservation, cells with a passage number as low as possible were expanded to a high number. At approximately 90% confluency, cells were trypsinized as described before and spun down at 500rpm for 5min. The

supernatant was removed and cells were gently resuspended in FCS containing 5% DMSO to a final cell density of $1x10^7$ cells/ml. The suspension was aliquoted into 1ml in sterile cryotubes pre-cooled to -20°C. The closed tubes were transferred to the -80°C freezer or for long-term preservation stored in liquid N_2.

5.2.5.4 Thawing

For thawing of mammlian cells 10 ml of pre-warmed culture medium were prepared in a 50 ml vial. An aliquot of frozen cells was thawed in hand and the cells were resuspended fast, by pipetting up and down with the pre-warmed culture medium. The cells were then transferred to a 10 cm culture dish and incubated at 37°C and 5% CO_2, until further usage.

5.2.5.5 Transfection of mammalian cells

Cells were transfected using FugeneHD (Promega) according to the manufacturers manual.

5.2.5.6 Growing cells on coverslips

For the immunofluorescence experiments mammalian cells were grown on glass coverslips. Therefore 22x22 mm coverslips were soaked in 100% ethanol p.a. in a petri dish, which was opened in the sterile bench only. 6-well culture plates were used and one coverslip was put in each well. The culture dish was left open in the sterile bench until the excess ethanol was evaporated. Then cells were treated as described for splitting and an appropriate number of cells were seeded in each well in a total volume of 3 ml growth medium. Cells were incubated at 37°C, 5% CO_2 until appropriate confluency for transfection or for microscopy analysis was reached.

5.2.6 Microscopy methods

5.2.6.1 Fixation of adherently growing cells

Adherent cells on cover slips were fixated in 4% paraformaldehyde (PFA) in PBS prior to immunfluorescence experiments.

Therefore the medium was aspirated and the cells were rinsed 2-3 times with pre-warmed PBS at 37°C.

Afterwards cells were fixed in 4% PFA in PBS for 10 min at room temperature, shaking. During the last minute a few drops of 0.5% Triton X-100 in PBS were added to the cells.

This was followed by washing the cells three times for 3 min in 0.01% Triton X-100 PBS at room temperature, shaking.

Finally cells were permeabilized with 0.5% Triton X-100 PBS for 5 min at room temperature, shaking.

The fixed cells were stored in PBS at 4°C until their usage for immunofluorescence.

5.2.6.2 Immunofluorescence

For immunofluorescence the fixed cells were washed in PBST three times for 3 min at room temperature, shaking.

Cells were then blocked with incubation in 4% BSA in PBST at 37°C for 30 min, shaking.

Afterwards the cells were incubated with the primary antibodies in 4% BSA in PBST for 1 h at 37°C in a humified chamber, shaking.

Cells were washed in PBST three times 5 min, shaking.

Then the secondary antibodies were added in 4% BSA in PBST and the cells were again incubated for 1 h at room temperature, shaking.

Cells were washed one time with PBST for 5min.

Afterward the cells were stained with DAPI (0.5 µg/ml) in PBST for 5 min at room temperature, shaking.

Finally cells were washed again one time with PBST for 5 min at room temperature, shaking.

The cells were rinsed in PBS and mounted on microscopy slides using the Vectashield mounting medium. Mounted samples were used immediately for microscopy analysis or stored at 4°C.

5.2.7 Nuclear matrix preparation

5.2.7.1 Nuclear matrix isolation

The isolation of nuclear matrix was carried out essentially as described (Berezney and Coffey, 1974; Cremer and Cremer, 2006; Reyes et al., 1997). 2.5-5x10^6 cells were washed in 1 ml PBS, followed by 5 min centrifugation at 1000g.

After this the cells were extracted in 200 µl CSK buffer by incubation for 5 min at 4°C. The soluble cytoplasmic proteins were separated by centrifugation at 5,000 g for 3 min (supernatant = 'CP' cytoplasmic fraction).

Chromatin was solubilized by DNA digestion with 400 U of RNase-free DNase I (Roche) in 110 µl CSK buffer plus protease inhibitors for 90 min at 37°C with shaking at 300 rpm. Then 50 µl 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*g* for 3 min. (supernatant = 'CHR' chromatin fraction).

The pellet was further extracted with 100 µl CSK buffer high salt for 10 min at 4°C on rotating wheel, and then centrifuged at 5,000*g* for 3 min. This treatment removes all accessible DNA and histones from the nucleus (supernatant = '2M' 2M salt wash fraction, which was diluted 1:2 with water before SDS-PAGE).

The remaining pellet was solubilized in 200 µl 8M urea buffer and was considered the nuclear matrix-containing fraction ('NM').

100 µl of each fraction was dialyzed against 1xTE buffer and subsequently digested with RNaseA (37°C, 1h) and proteinaseK (50°C over night), precipitated, dissolved in 50 µl ddH_2O, controlled on 1% agarose gel (10 µl) and the 'NM' fraction was subjected to qPCR analysis. 25 µl 5x Lämmli buffer was added to the remaining 100 µl of the fractions, and 20 µl each were subjected to Western blot

analysis or Coomassie staining after boiling the samples at 95°C, 10 min and separating on denaturing SDS-polyacrylamide gels.

5.2.7.2 Nuclear matrix preparations *in situ*

In situ preparations of nuclear matrix for immunofluorescence analysis was done essentially as described (Vecerova et al., 2004).

Cells were grown on coverslips in a 6-well culture plate until 50-75% confluency was reached.

Growth medium was removed and the cells were washed with ice cold PBS, twice.

After this cells were extracted in CSK buffer supplemented with 20 U/ml RNasin for 10 min at 4°C.

The CSK buffer was removed and the cells were incubated with CSK extraction buffer supplemented with 20 U/ml RNasin for 5 min at 4°C.

Afterwards the CSK extraction buffer was removed and the CSK digestion buffer supplemented with 20 U/ml RNasin and DNase I was added. This buffer contained 500 U/ml recombinant DNase I, RNase-free (Roche) and cells were incubated in this buffer at 30°C for 30 min.

Finally the CSK digestion buffer was removed and cells were incubated with CSK extraction buffer supplemented with 20 U/ml RNasin for 5 min at 4°C.

Following the removal of the CSK extraction buffer, the cells were fixed in 4% PFA in PBS and processed for immunofluorescence as described previously.

5.2.8 Microscale thermophoresis

5.2.8.1 The concept of microscale thermophoresis

The microscale thermophoresis (MST) technologie enables the quantitative measurement of molecular interactions in solution. The theoretical background of the method and its usage for a wide array of different assays was described in recent publications (Baaske et al., 2010; Jerabek-Willemsen et al., 2011).

In general the method is based on the thermophoretical movement of molecules along an induced temperature gradient. This movement is determined by

parameters like the size, charge and hydration shell of a molecule. Since the binding of another molecule influences all these parameters, the interaction leads to a changed thermophoretic behavior. These changes in thermophoretic behavior are basically recorded by the Monolith MST device (NanoTemper technologies) in the form of thermophoresis profiles. From these profiles binding curves for the molecules of interest can be generated, which are then analyzed regarding K_d constants or EC_{50} values.

The principal setup of a MST experiment is illustrated in figure 5.1A. The only prerequisite for the measurement is that one of the two binding partners contains a fluorescent label for detection by the optics of the MST device. The molecule bearing the fluorescent label is kept at a constant concentration and a serial dilution series of the putative binding partner is titrated to the labeled molecules. The sample mixtures are then sucked into glass capillaries for the measurement in the MST device.

The typical progress of an MST experiment is depicted in figure 5.1.B. Upon induction of the laser a temperature gradient is induced and the molecules start their movement. This is measured over a certain timeframe, then the laser is switched of and the back diffusion of the molecules sets in. For the analysis done in this work, two parts of the created curve are of interest. First the temperature jump signals, which describe the time points shortly after the laser is switched on. These signals represent fast binding events that directly influence the microenvironment of the fluorescent label and can therefore give some insights on the localized interaction of the two molecules. Second the thermophoresis signals of the later time points, which represent the overall thermophoretic behavior of the whole complex formed by the two binding partners and thus are more determined by changes in size, charge and the hydration shell.

Figure 5.1.C shows the differences in the thermophoretic curves, which are induced by the binding events. One can clearly distinguish the thermophoresis curves of the different states of unbound, partially bound and fully bound molecules in the reactions.

Finally these differences are translated into binding curves, by plotting the normalized temperature jump or thermopheresis data against the concentrations of the non-labeled binding partner. From these binding curves K_d constants and

EC_{50} values can be calculated, using the law of mass action or the Hill-equation respectively (Fig.: 5.1.D).

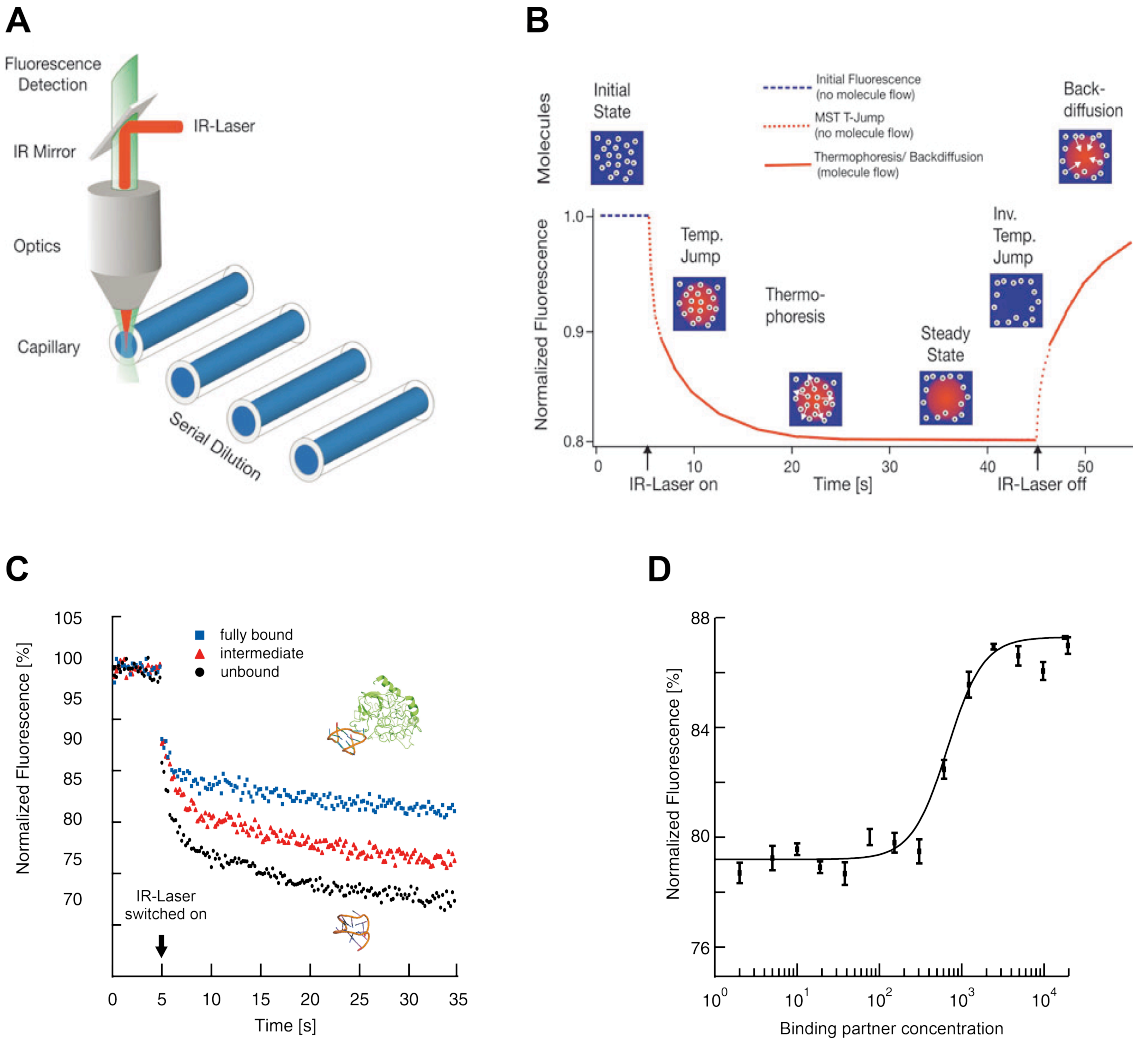


Figure 5.1: The concept of microscale thermophoresis
A) Setup of the Monolith MST device, with a fluorescence detector, an IR-laser for the temperature gradient induction and the glass capillaries containing the two reaction partners, the labeled one in a constant concentration and the non-labeled titrated in a serial dilution. **B**) The general outline of a thermophoresis experiment. Upon induction of the laser the molecules start to align themselves to the temperature gradient, these early time points describe the temperature jump behavior. At the later time points the molecules move along the temperature gradient according to their size, charge and hydration shell, this describes the thermophoresis data. Then a steady state is reached, the laser is switched off and the molecules start a back diffusion. **C**) Upon binding the thermophoretic movement of the molecules changes which can be seen in the curves, from unbound state (black), to partially bound (red) and finally fully bound state (blue). **D**) These different thermophresis and temperature jump signals can be plotted against the concentration of the unlabeled binding partner and thus binding curves are generated, from which K_d and EC_{50} values are calculated.

5.2.8.2 Triplex MST experiments

The triplex annealing for MST analysis was performed in a total reaction volume of 20 µl for each sample. The experiments were conducted in 0.2 ml sterile PCR tubes. Sample preparation was done on ice. A master mix containing 100 nM of the target DNA was prepared using a 5x stock of the appropriate MST triplex buffer as follows:

Master mix	1x
5x MST triplex buffer	4 µl
Enhancer DNA Cy5 (10 µM)	0.2 µl
ddH_2O	15.8 µl
Total volume	18 µl

In addition a serial dilution of 1:1 of the corresponding RNA or DNA TFOs was made in ddH_2O. Then 2 µl of each dilution step was added to 18 µl master mix. In addition a negative control containing no TFO was also prepared with 2 µl ddH_2O. The samples were mixed by vortexing and incubated at 37°C for 1 h in a thermomixer. After this the thermomixer was switched of and the samples were slowly cooled down to room temperature. Again samples were mixed by vortexing and then used for MST analysis.

Therefore 5-10 µl were sucked into standard treated glass capillaries and incubated at the appropriate temperature for 15 min inside the Monolith MST device. This was followed by and initial capillary scan with an LED power of 20%. Then the MST measurement was performed with an MST power of 30% and a laser on time of 35 sec. The experiments were performed in technical quadruplicates. The temperature jump and combined temperature jump and thermophoresis data from each round of measurement was exported to the KaleidaGraph software and binding curves were generated by plotting these datasets against the TFO concentrations. In this way the optimal data sets for evaluation of binding constants were determined. Finally K_d constants were obtained by fitting the binding curves according to the law of mass action, as described in (Baaske et al., 2010), following the formula:

$$x = \frac{1}{2c_A^0}\left(c_A^0 + c_T^0 + K_d - \sqrt{(c_A^0 + c_T^0 + K_d)^2 - 4c_A^0 c_T^0}\right)$$

with c_A representing the concentration of the DNA duplex and c_T the concentration of the TFOs. If direct comparisons between different binding reactions were illustrated, the temperature jump and thermophoresis signals were normalized to the fraction bound (X) by X= (Y(c)-Min)/(Max-Min) and again fitted as described above.

5.2.8.3 Analysis of protein-DNA interactions with MST

The analysis of eAT-hook binding to DNA with the MST device was performed as described in (Zillner et al., 2012).

The reactions were setup on ice in 0.2 ml PCR tubes. Total reaction volume was 15 µl, containing 5 µl of DNA master mix and 10 µl of protein dilution. For the master mix a 5x stock of the MST protein buffer was used, each sample contained 50 nM of Cy5 labeled DNA. For the proteins a dilution 2:1 dilution series in 20 mM Tris-HCl, pH 8 with 10% glycerol was prepared. The master mix and protein dilutions were mixed by pipetting and sucked into standard treated class capillaries. MST measurements were performed at 25°C with an LED power of 40% and a MST power of 40% and a laser on time of 35 sec in technical triplicates. For data evaluation the temperature jump signals were normalized to fraction bound (X) by X= (Y(c)-Min)/(Max-Min) and the EC_{50} values were obtained by fitting with the Hill equation corrected for the minimum (*Min*) and maximum (*Max*) values of the binding curve: $Y(c) = Min+(Max-Min)/(1+EC_{50} / c^0_{pep})^n$, with *Y* (c) being the thermophoresis signal, c^0_{pep} the variable concentration of the respective peptide. Again data analysis was done using the KaleidaGraph software.

5.2.8.4 Analysis of protein-RNA interactions with MST

The MST analysis of protein RNA interactions was done as described above for the protein DNA interactions. Only difference was the use of 50 nM fluorescently

labeled single stranded RNA instead of DNA. In the competitive assay 50 nM labeled DNA and 50 nM of labeled RNA were used per sample

6 Appendix

6.1 *Curriculum Vitae*

Name:	Filarsky
First name:	Michael
Nationality:	German
Date of birth:	30.11.1982
Place of birth:	Augsburg

Education:

Since May 2009	PhD, "Exploring the role of structured DNA elements and their specific binders in genome regulation" Lab of Chromatin Dynamics and Nuclear Architecture, Biochemistry III, University of Regensburg, Germany Supervised by Prof. Gernot Längst and Dr. Attila Nemeth
06/2008 – 04/2009	Diploma thesis: "Interactions between the Leptin and Adiponectin induced signaling pathways" Department of Internal Medicine I, University Clinic Regensburg Supervised by Prof. Christa Buechler
10/2003 – 04/2009	Studies (Diplom) of Biology, University of Regensburg, Germany Major subject: Biochemistry Minor subjects: Genetics and Organic Chemistry
09/1993 – 06/2003	Carl von Closen High School Eggenfelden, Germany

Internships:

1/2011 – 3/2011	Method training, Plückthun lab, Department of Biochemistry, University of Zürich, Switzerland, Generation of a triplex specific DARPin
9/2006 – 4/2007	Research internship, Initiative of Systems Biology, University of New South Wales, Sydney, Australia Supervision by Prof. Marc Wilkins

6.2 List of publications

Large-scale organization of ribosomal DNA chromatin is regulated by Tip5
Filarsky M[#], Zillner K[#], Rachow K, Weinberger M, Längst G, Nemeth A.
Nucleic Acids Res. 2013 Apr 10

Circulating levels of chemerin and adiponectin are higher in ulcerative colitis and chemerin is elevated in Crohn's disease.
Weigert J, Obermeier F, Neumeier M, Wanninger J, Filarsky M, Bauer S, Aslanidis C, Rogler G, Ott C, Schäffler A, Schölmerich J, Buechler C.
Inflamm Bowel Dis. 2010 Apr;16(4):630-7.

Systemic chemerin is related to inflammation rather than obesity in type 2 diabetes.
Weigert J, Neumeier M, Wanninger J, Filarsky M, Bauer S, Wiest R, Farkas S, Scherer MN, Schäffler A, Aslanidis C, Schölmerich J, Buechler C.
Clin Endocrinol (Oxf). 2010 Mar;72(3):342-8.

Micropreparative fractionation of the complexome by blue native continuous elution electrophoresis.
Huang KY, Filarsky M, Padula MP, Raftery MJ, Herbert BR, Wilkins MR.
Proteomics. 2009 May;9(9):2494-502.

Adiponectin upregulates monocytic activin A but systemic levels are not altered in obesity or type 2 diabetes.
Weigert J, Neumeier M, Wanninger J, Schober F, Sporrer D, Weber M, Schramm A, Wurm S, Stögbauer F, Filarsky M, Schäffler A, Aslanidis C, Schölmerich J, Buechler C.
Cytokine. 2009 Feb;45(2):86-91.

Reduced response to adiponectin and lower abundance of adiponectin receptor proteins in type 2 diabetic monocytes.
Weigert J, Neumeier M, Wanninger J, Wurm S, Kopp A, Schober F, Filarsky M, Schäffler A, Zeitoun M, Aslanidis C, Buechler C.
FEBS Lett. 2008 May 28;582(12):1777-82.

7 Acknowledgements

At the end I would like to take the chance and thank the people that made this work possible.
First of all, a big thank you to Prof. Gernot Längst, who gave me the opportunity to work in his lab and was a great supervisor. He even sacrificed his precious holiday time for giving me advice when needed, so thank you I really appreciate it.
Also a big thank you to PD Attila Nemeth for his supervision and all the help in pushing my project further, intellectually and experimentally ("magic hands").
A firm handshake goes out to all my colleagues at the House of the Ribosome, who made it easy to come to the lab, even in the more desperate times.
Furthermore I want to thank all my friends for making the last years the most incredible time, thanks for the ride, guys!
And finally I want to thank my family without you I would not be were I'm now, thank you so much for everything,

Cheers to you all!

8 Bibliography

Agazie, Y. M., Burkholder, G. D., and Lee, J. S. (1996). Triplex DNA in the nucleus: direct binding of triplex-specific antibodies and their effect on transcription, replication and cell growth. Biochem J *316 (Pt 2)*, 461-466.

Agazie, Y. M., Lee, J. S., and Burkholder, G. D. (1994). Characterization of a new monoclonal antibody to triplex DNA and immunofluorescent staining of mammalian chromosomes. J Biol Chem *269*, 7019-7023.

Alberti, P., Arimondo, P. B., Mergny, J. L., Garestier, T., Helene, C., and Sun, J. S. (2002). A directional nucleation-zipping mechanism for triple helix formation. Nucleic Acids Res *30*, 5407-5415.

Alberts, B. (2002). Molecular biology of the cell, 4th edn (New York: Garland Science).

Alberts, B. (2008). Molecular biology of the cell. Reference edition, 5th edn (New York: Garland Science).

Allen, G. C., Spiker, S., and Thompson, W. F. (2000). Use of matrix attachment regions (MARs) to minimize transgene silencing. Plant Mol Biol *43*, 361-376.

Aravind, L., and Landsman, D. (1998). AT-hook motifs identified in a wide variety of DNA-binding proteins. Nucleic Acids Res *26*, 4413-4421.

Baaske, P., Wienken, C. J., Reineck, P., Duhr, S., and Braun, D. (2010). Optical thermophoresis for quantifying the buffer dependence of aptamer binding. Angew Chem Int Ed Engl *49*, 2238-2241.

Bacolla, A., Jaworski, A., Larson, J. E., Jakupciak, J. P., Chuzhanova, N., Abeysinghe, S. S., O'Connell, C. D., Cooper, D. N., and Wells, R. D. (2004). Breakpoints of gross deletions coincide with non-B DNA conformations. Proc Natl Acad Sci U S A *101*, 14162-14167.

Bacolla, A., and Wells, R. D. (2004). Non-B DNA conformations, genomic rearrangements, and human disease. J Biol Chem *279*, 47411-47414.

Bagasra, O., Stir, A. E., Pirisi-Creek, L., Creek, K. E., Bagasra, A. U., Glenn, N., and Lee, J. S. (2006). Role of micro-RNAs in regulation of lentiviral latency and persistence. Appl Immunohistochem Mol Morphol *14*, 276-290.

Baker, S. A., Chen, L., Wilkins, A. D., Yu, P., Lichtarge, O., and Zoghbi, H. Y. (2013). An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. Cell *152*, 984-996.

Barre, F. X., Ait-Si-Ali, S., Giovannangeli, C., Luis, R., Robin, P., Pritchard, L. L., Helene, C., and Harel-Bellan, A. (2000). Unambiguous demonstration of triple-helix-directed gene modification. Proc Natl Acad Sci U S A *97*, 3084-3088.

Beal, P. A., and Dervan, P. B. (1991). Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. Science *251*, 1360-1363.

Behe, M. J. (1995). An overabundance of long oligopurine tracts occurs in the genome of simple and complex eukaryotes. Nucleic Acids Res *23*, 689-695.

Benedit, P., Paciucci, R., Thomson, T. M., Valeri, M., Nadal, M., Càceres, C., de Torres, I., Estivill, X., Lozano, J. J., Morote, J., and Reventós, J. (2001). PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. Oncogene *20*, 1455-1464.

Berezney, R., and Coffey, D. S. (1974). Identification of a nuclear protein matrix. Biochem Biophys Res Commun *60*, 1410-1417.

Berezney, R., Mortillaro, M. J., Ma, H., Wei, X., and Samarabandu, J. (1995). The nuclear matrix: a structural milieu for genomic function. Int Rev Cytol *162A*, 1-65.

Besch, Giovannangeli, Schuh, Kammerbauer, and Degitz (2004). Characterization and quantification of triple helix formation in chromosomal DNA. J Mol Biol *341*, 979-989.

Besch, R., Giovannangeli, C., Kammerbauer, C., and Degitz, K. (2002). Specific inhibition of ICAM-1 expression mediated by gene targeting with Triplex-forming oligonucleotides. J Biol Chem *277*, 32473-32479.

Bochman, M. L., Paeschke, K., and Zakian, V. A. (2012). DNA secondary structures: stability and function of G-quadruplex structures. Nat Rev Genet *13*, 770-780.

Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klehr, D., Mielke, C., and Kohwi-Shigematsu, T. (1992). Biological significance of unwinding capability of nuclear matrix-associating DNAs. Science *255*, 195-197.

Bode, J., Schlake, T., Rios-Ramirez, M., Mielke, C., Stengert, M., Kay, V., and Klehr-Wirth, D. (1995). Scaffold/matrix-attached regions: structural properties creating transcriptionally active loci. Int Rev Cytol *162A*, 389-454.

Bolla, R. I., Braaten, D. C., Shiomi, Y., Hebert, M. B., and Schlessinger, D. (1985). Localization of specific rDNA spacer sequences to the mouse L-cell nucleolar matrix. Mol Cell Biol *5*, 1287-1294.

Bontems, F., Verger, A., Dewitte, F., Lens, Z., Baert, J. L., Ferreira, E., de Launoit, Y., Sizun, C., Guittet, E., Villeret, V., and Monte, D. (2011). NMR structure of the human Mediator MED25 ACID domain. J Struct Biol *174*, 245-251.

Bourachot, B., Yaniv, M., and Muchardt, C. (1999). The activity of mammalian brm/SNF2alpha is dependent on a high-mobility-group protein I/Y-like DNA binding domain. Mol Cell Biol *19*, 3931-3939.

Brown, and Fox (1999). DNA triple-helix formation on nucleosome core particles. Effect of length of the oligopurine tract. Eur J Biochem *261*, 301-310.

Brown, P. M., and Fox, K. R. (1998). DNA triple-helix formation on nucleosome-bound poly(dA).poly(dT) tracts. Biochem J *333 (Pt 2)*, 259-267.

Burkholder, G. D., Latimer, L. J., and Lee, J. S. (1988). Immunofluorescent staining of mammalian nuclei and chromosomes with a monoclonal antibody to triplex DNA. Chromosoma *97*, 185-192.

Burkholder, G. D., Latimer, L. J., and Lee, J. S. (1991). Immunofluorescent localization of triplex DNA in polytene chromosomes of Chironomus and Drosophila. Chromosoma *101*, 11-18.

Buske, F. A., Bauer, D. C., Mattick, J. S., and Bailey, T. L. (2012). Triplexator: Detecting nucleic acid triple helices in genomic and transcriptomic data. Genome Res.

Buske, F. A., Mattick, J. S., and Bailey, T. L. (2011). Potential in vivo roles of nucleic acid triple-helices. RNA biology *8*.

Caburet, S., Conti, C., Schurra, C., Lebofsky, R., Edelstein, S. J., and Bensimon, A. (2005). Human ribosomal RNA gene arrays display a broad range of palindromic structures. Genome Res *15*, 1079-1085.

Carbone, G. M., Napoli, S., Valentini, A., Cavalli, F., Watson, D. K., and Catapano, C. V. (2004). Triplex DNA-mediated downregulation of Ets2 expression results in growth inhibition and apoptosis in human prostate cancer cells. Nucleic Acids Res *32*, 4358-4367.

Catapano, McGuffie, Pacheco, and Carbone (2000). Inhibition of gene expression and cell proliferation by triple helix-forming oligonucleotides directed to the c-myc gene. Biochemistry *39*, 5126-5138.

Chan, P. K., Aldrich, M., and Busch, H. (1985). Alterations in immunolocalization of the phosphoprotein B23 in HeLa cells during serum starvation. Exp Cell Res *161*, 101-110.

Chen, G., and Chen, S. J. (2011). Quantitative analysis of the ion-dependent folding stability of DNA triplexes. Phys Biol *8*, 066006.

Cheng, A. J., and Van Dyke, M. W. (1994). Oligodeoxyribonucleotide length and sequence effects on intermolecular purine-purine-pyrimidine triple-helix formation. Nucleic Acids Res *22*, 4742-4747.

Chin, J. Y., Schleifman, E. B., and Glazer, P. M. (2007). Repair and recombination induced by triple helix DNA. Front Biosci *12*, 4288-4297.

Clark, R. M., Bhaskar, S. S., Miyahara, M., Dalgliesh, G. L., and Bidichandani, S. I. (2006). Expansion of GAA trinucleotide repeats in mammals. Genomics *87*, 57-67.

Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., and Hogan, M. E. (1988). Site-specific oligonucleotide binding represses transcription of the human c-myc gene in vitro. Science *241*, 456-459.

Craig, J. M., Boyle, S., Perry, P., and Bickmore, W. A. (1997). Scaffold attachments within the human genome. J Cell Sci *110 (Pt 21)*, 2673-2682.

Cremer, T., and Cremer, C. (2006). Rise, fall and resurrection of chromosome territories: a historical perspective. Part II. Fall and resurrection of chromosome territories during the 1950s to 1980s. Part III. Chromosome territories and the functional nuclear architecture: experiments and models from the 1990s to the present. Eur J Histochem *50*, 223-272.

Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I., and Fakan, S. (2006). Chromosome territories--a functional nuclear landscape. Curr Opin Cell Biol *18*, 307-316.

Dev, V. G., Tantravahi, R., Miller, D. A., and Miller, O. J. (1977). Nucleolus organizers in Mus musculus subspecies and in the RAG mouse cell line. Genetics *86*, 389-398.

Durand, M., Peloille, S., Thuong, N. T., and Maurizot, J. C. (1992). Triple-helix formation by an oligonucleotide containing one (dA)12 and two (dT)12 sequences bridged by two hexaethylene glycol chains. Biochemistry *31*, 9197-9204.

Eilebrecht, S., Benecke, B. J., and Benecke, A. (2011a). 7SK snRNA-mediated, gene-specific cooperativity of HMGA1 and P-TEFb. RNA Biol *8*, 1084-1093.

Eilebrecht, S., Brysbaert, G., Wegert, T., Urlaub, H., Benecke, B. J., and Benecke, A. (2011b). 7SK small nuclear RNA directly affects HMGA1 function in transcription regulation. Nucleic Acids Res *39*, 2057-2072.

Eivazova, E. R., Vassetzky, Y. S., and Aune, T. M. (2007). Selective matrix attachment regions in T helper cell subsets support loop conformation in the Ifng gene. Genes Immun *8*, 35-43.

Elcock, L. S., and Bridger, J. M. (2008). Exploring the effects of a dysfunctional nuclear matrix. Biochem Soc Trans *36*, 1378-1383.

Ellouze, C., Piot, F., and Takahashi, M. (1997). Use of fluorescein-labeled oligonucleotide for analysis of formation and dissociation kinetics of T:A:T triple-stranded DNA: effect of divalent cations. J Biochem *121*, 521-526.

Escudé, C., François, J. C., Sun, J. S., Ott, G., Sprinzl, M., Garestier, and Hélène, C. (1993). Stability of triple helices containing RNA and DNA strands: experimental and molecular modeling studies. Nucleic Acids Res *21*, 5547-5553.

Faruqi, A. F., Datta, H. J., Carroll, D., Seidman, M. M., and Glazer, P. M. (2000). Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway. Mol Cell Biol *20*, 990-1000.

Felsenfeld, G., and Rich, A. (1957). Studies on the formation of two- and three-stranded polyribonucleotides. Biochim Biophys Acta *26*, 457-468.

Floris, R., Scaggiante, B., Manzini, G., Quadrifoglio, F., and Xodo, L. E. (1999). Effect of cations on purine.purine.pyrimidine triple helix formation in mixed-valence salt solutions. Eur J Biochem *260*, 801-809.

Forrester, W. C., van Genderen, C., Jenuwein, T., and Grosschedl, R. (1994). Dependence of enhancer-mediated transcription of the immunoglobulin mu gene on nuclear matrix attachment regions. Science *265*, 1221-1225.

Fukuda, Y. (2000). Interaction of nuclear proteins with intrinsically curved DNA in a matrix attachment region of a tobacco gene. Plant Mol Biol *44*, 91-98.

Gavrilov, A. A., Zukher, I. S., Philonenko, E. S., Razin, S. V., and Iarovaia, O. V. (2010). Mapping of the nuclear matrix-bound chromatin hubs by a new M3C experimental procedure. Nucleic Acids Res *38*, 8051-8060.

Glazko, G., Koonin, E., Rogozin, I., and Shabalina, S. (2003). A significant fraction of conserved noncoding DNA in human and mouse consists of predicted matrix attachment regions. Trends Genet *19*, 119-124.

Goñi, J. R., de la Cruz, X., and Orozco, M. (2004). Triplex-forming oligonucleotide target sequences in the human genome. Nucleic Acids Res *32*, 354-360.

Goñi, J. R., Vaquerizas, J. M., Dopazo, J., and Orozco, M. (2006). Exploring the reasons for the large density of triplex-forming oligonucleotide target sequences in the human regulatory regions. BMC Genomics *7*, 63.

Gonzalez, I. L., and Sylvester, J. E. (1995). Complete sequence of the 43-kb human ribosomal DNA repeat: analysis of the intergenic spacer. Genomics *27*, 320-328.

Gorab, E., Amabis, J. M., Stocker, A. J., Drummond, L., and Stollar, B. D. (2009). Potential sites of triple-helical nucleic acid formation in chromosomes of Rhynchosciara (Diptera: Sciaridae) and Drosophila melanogaster. Chromosome Res.

Grozdanov, P., Georgiev, O., and Karagyozov, L. (2003). Complete sequence of the 45-kb mouse ribosomal DNA repeat: analysis of the intergenic spacer. Genomics *82*, 637-643.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R., and Hofacker, I. L. (2008). The Vienna RNA websuite. Nucleic Acids Res *36*, W70-74.

Grummt, I., Kuhn, A., Bartsch, I., and Rosenbauer, H. (1986). A transcription terminator located upstream of the mouse rDNA initiation site affects rRNA synthesis. Cell *47*, 901-911.

Grummt, I., Maier, U., Ohrlein, A., Hassouna, N., and Bachellerie, J. P. (1985). Transcription of mouse rDNA terminates downstream of the 3' end of 28S RNA and involves interaction of factors with repeated sequences in the 3' spacer. Cell *43*, 801-810.

Guetg, C., Lienemann, P., Sirri, V., Grummt, I., Hernandez-Verdun, D., Hottiger, M. O., Fussenegger, M., and Santoro, R. (2010). The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. EMBO J *29*, 2135-2146.

Haltiner, M. M., Smale, S. T., and Tjian, R. (1986). Two distinct promoter elements in the human rRNA gene identified by linker scanning mutagenesis. Mol Cell Biol *6*, 227-235.

Hampel, K. J., Crosson, P., and Lee, J. S. (1991). Polyamines favor DNA triplex formation at neutral pH. Biochemistry *30*, 4455-4459.

Hamperl, S., Wittner, M., Babl, V., Perez-Fernandez, J., Tschochner, H., and Griesenbeck, J. (2013). Chromatin states at ribosomal DNA loci. Biochim Biophys Acta *1829*, 405-417.

Han, H., and Dervan, P. B. (1993). Sequence-specific recognition of double helical RNA and RNA.DNA by triple helix formation. Proc Natl Acad Sci U S A *90*, 3806-3810.

Hancock, R. (2000). A new look at the nuclear matrix. Chromosoma *109*, 219-225.

Hart, C. M., and Laemmli, U. K. (1998). Facilitation of chromatin dynamics by SARs. Current Opinion in Genetics & Development *8*, 519-525.

He, D. C., Nickerson, J. A., and Penman, S. (1990). Core filaments of the nuclear matrix. J Cell Biol *110*, 569-580.

Henderson, A. S., Warburton, D., and Atwood, K. C. (1972). Location of ribosomal DNA in the human chromosome complement. Proc Natl Acad Sci U S A *69*, 3394-3398.

Henderson, S., and Sollner-Webb, B. (1986). A transcriptional terminator is a novel element of the promoter of the mouse ribosomal RNA gene. Cell *47*, 891-900.

Heng, H. H., Goetze, S., Ye, C. J., Liu, G., Stevens, J. B., Bremer, S. W., Wykes, S. M., Bode, J., and Krawetz, S. A. (2004). Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. J Cell Sci *117*, 999-1008.

Holland, J. A., and Hoffman, D. W. (1996). Structural features and stability of an RNA triple helix in solution. Nucleic Acids Res *24*, 2841-2848.

Hoogsteen, K. (1959). The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. Acta Crystallographica *12*, 822-823.

Hoyne, Gacy, McMurray, and Maher (2000). Stabilities of intrastrand pyrimidine motif DNA and RNA triple helices. Nucleic Acids Res *28*, 770-775.

Hsu, L. C., Liu, S., Abedinpour, F., Beech, R. D., Lahti, J. M., Kidd, V. J., Greenspan, J. A., and Yeung, C. Y. (2003). The murine G+C-rich promoter binding protein mGPBP is required for promoter-specific transcription. Mol Cell Biol *23*, 8773-8785.

Huppert, J. L., and Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res *35*, 406-413.

Huth, J. R., Bewley, C. A., Nissen, M. S., Evans, J. N., Reeves, R., Gronenborn, A. M., and Clore, G. M. (1997). The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. Nat Struct Biol *4*, 657-665.

Jackson, D. A., Yuan, J., and Cook, P. R. (1988). A gentle method for preparing cyto- and nucleo-skeletons and associated chromatin. J Cell Sci *90 (Pt 3)*, 365-378.

Jain, A., Akanchha, S., and Rajeswari, M. R. (2005). Stabilization of purine motif DNA triplex by a tetrapeptide from the binding domain of HMGBI protein. Biochimie *87*, 781-790.

Jain, A., Wang, G., and Vasquez, K. M. (2008). DNA triple helices: biological consequences and therapeutic potential. Biochimie *90*, 1117-1130.

James, Brown, T., and Fox (2003). Thermodynamic and kinetic stability of intermolecular triple helices containing different proportions of C+*GC and T*AT triplets. Nucleic Acids Res *31*, 5598-5606.

Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P., and Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. Assay Drug Dev Technol *9*, 342-353.

Jimenez-Garcia, E., Vaquero, A., Espinas, M. L., Soliva, R., Orozco, M., Bernues, J., and Azorin, F. (1998). The GAGA factor of Drosophila binds triple-stranded DNA. J Biol Chem *273*, 24640-24648.

Kanak, M., Alseiari, M., Balasubramanian, P., Addanki, K., Aggarwal, M., Noorali, S., Kalsum, A., Mahalingam, K., Pace, G., Panasik, N., and Bagasra, O. (2010). Triplex-forming MicroRNAs form stable complexes with HIV-1 provirus and inhibit its replication. Appl Immunohistochem Mol Morphol *18*, 532-545.

Keppel, F. (1986). Transcribed human ribosomal RNA genes are attached to the nuclear matrix. J Mol Biol *187*, 15-21.

Kiessling, L. L., Griffin, L. C., and Dervan, P. B. (1992). Flanking sequence effects within the pyrimidine triple-helix motif characterized by affinity cleaving. Biochemistry *31*, 2829-2834.

Kimura, H., Tao, Y., Roeder, R. G., and Cook, P. R. (1999). Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble

holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Mol Cell Biol *19*, 5383-5392.

Koziol, M. J., and Rinn, J. L. (2010). RNA traffic control of chromatin complexes. Curr Opin Genet Dev *20*, 142-148.

Kramer, J. A., Singh, G. B., and Krawetz, S. A. (1996). Computer-assisted search for sites of nuclear matrix attachment. Genomics *33*, 305-308.

la Cour, T., Kiemer, L., Molgaard, A., Gupta, R., Skriver, K., and Brunak, S. (2004). Analysis and prediction of leucine-rich nuclear export signals. Protein Eng Des Sel *17*, 527-536.

Laemmli, U. K., Kas, E., Poljak, L., and Adachi, Y. (1992). Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. Curr Opin Genet Dev *2*, 275-285.

Lang, S. (2008) RNA vermittelte Mechanismen in der sequenzspezifischen Erkennung von Chromatin, University of Regensburg, Regensburg.

Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J. L., Thuong, N. T., Lhomme, J., and Helene, C. (1987). Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-[alpha]-thymidylate covalently linked to an azidoproflavine derivative. Nucleic Acids Res *15*, 7749-7760.

Learned, R. M., Learned, T. K., Haltiner, M. M., and Tjian, R. T. (1986). Human rRNA transcription is modulated by the coordinate binding of two factors to an upstream control element. Cell *45*, 847-857.

Lee, J. S., Burkholder, G. D., Latimer, L. J., Haug, B. L., and Braun, R. P. (1987). A monoclonal antibody to triplex DNA binds to eucaryotic chromosomes. Nucleic Acids Res *15*, 1047-1061.

Lee, J. S., Johnson, D. A., and Morgan, A. R. (1979). Complexes formed by (pyrimidine)n . (purine)n DNAs on lowering the pH are three-stranded. Nucleic Acids Res *6*, 3073-3091.

Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V., and Fresco, J. R. (1988). Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns. Biochemistry *27*, 9108-9112.

Li, G., Tolstonog, G. V., and Traub, P. (2002). Interaction in vitro of type III intermediate filament proteins with triplex DNA. DNA Cell Biol *21*, 163-188.

Li, J., Langst, G., and Grummt, I. (2006). NoRC-dependent nucleosome positioning silences rRNA genes. EMBO J *25*, 5735-5741.

Li, J., Santoro, R., Koberna, K., and Grummt, I. (2005). The chromatin remodeling complex NoRC controls replication timing of rRNA genes. EMBO J *24*, 120-127.

Linnemann, A. K., Platts, A. E., Doggett, N., Gluch, A., Bode, J., and Krawetz, S. A. (2007). Genomewide identification of nuclear matrix attachment regions: an analysis of methods. Biochem Soc Trans *35*, 612-617.

Little, R. D., Platt, T. H., and Schildkraut, C. L. (1993). Initiation and termination of DNA replication in human rRNA genes. Mol Cell Biol *13*, 6600-6613.

Ma, H., Siegel, A. J., and Berezney, R. (1999). Association of chromosome territories with the nuclear matrix. Disruption of human chromosome territories correlates with the release of a subset of nuclear matrix proteins. J Cell Biol *146*, 531-542.

Maher, L. J., 3rd, Dervan, P. B., and Wold, B. J. (1990). Kinetic analysis of oligodeoxyribonucleotide-directed triple-helix formation on DNA. Biochemistry *29*, 8820-8826.

Manning, G. S. (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. Quarterly Reviews of Biophysics *11*, 179-246.

Marqués, N., Sesé, M., Cánovas, V., Valente, F., Bermudo, R., de Torres, I., Fernández, Y., Abasolo, I., Fernández, P. L., Contreras, H.*, et al.* (2013). Regulation of protein translation and c-Jun expression by prostate tumor overexpressed 1. Oncogene.

Martianov, I., Ramadass, A., Serra Barros, A., Chow, N., and Akoulitchev, A. (2007). Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature *445*, 666-670.

Mayer, C., Neubert, M., and Grummt, I. (2008). The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. EMBO Rep *9*, 774-780.

Mayer, C., Schmitz, K. M., Li, J., Grummt, I., and Santoro, R. (2006). Intergenic transcripts regulate the epigenetic state of rRNA genes. Mol Cell *22*, 351-361.

McDonald, C. D., and Maher, L. J. (1995). Recognition of duplex DNA by RNA polynucleotides. Nucleic Acids Res *23*, 500-506.

McStay, B., and Grummt, I. (2008). The epigenetics of rRNA genes: from molecular to chromosome biology. Annu Rev Cell Dev Biol *24*, 131-157.

McStay, B., and Reeder, R. H. (1986). A termination site for Xenopus RNA polymerase I also acts as an element of an adjacent promoter. Cell *47*, 913-920.

Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. Nat Rev Genet *10*, 155-159.

Mercer, T. R., and Mattick, J. S. (2013). Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol *20*, 300-307.

Metcalf, C. E., and Wassarman, D. A. (2006). DNA binding properties of TAF1 isoforms with two AT-hooks. J Biol Chem *281*, 30015-30023.

Milbradt, A. G., Kulkarni, M., Yi, T., Takeuchi, K., Sun, Z. Y., Luna, R. E., Selenko, P., Naar, A. M., and Wagner, G. (2011). Structure of the VP16 transactivator target in the Mediator. Nat Struct Mol Biol *18*, 410-415.

Mirkovitch, J., Mirault, M. E., and Laemmli, U. K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell *39*, 223-232.

Mitton-Fry, R. M., DeGregorio, S. J., Wang, J., Steitz, T. A., and Steitz, J. A. (2010). Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. Science *330*, 1244-1247.

Morgan, A. R., and Wells, R. D. (1968). Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. J Mol Biol *37*, 63-80.

Moser, H. E., and Dervan, P. B. (1987). Sequence-specific cleavage of double helical DNA by triple helix formation. Science *238*, 645-650.

Moss, T., and Stefanovsky, V. Y. (1995). Promotion and regulation of ribosomal transcription in eukaryotes by RNA polymerase I. Prog Nucleic Acid Res Mol Biol *50*, 25-66.

Nakanishi, M., Weber, K. T., and Guntaka, R. V. (1998). Triple helix formation with the promoter of human alpha1(I) procollagen gene by an antiparallel triplex-forming oligodeoxyribonucleotide. Nucleic Acids Res *26*, 5218-5222.

Nelson, H. C., Finch, J. T., Luisi, B. F., and Klug, A. (1987). The structure of an oligo(dA).oligo(dT) tract and its biological implications. Nature *330*, 221-226.

Nelson, L., Bender, C., Mannsperger, H., Buergy, D., Kambakamba, P., Mudduluru, G., Korf, U., Hughes, D., Van Dyke, M., and Allgayer, H. (2012). Triplex DNA-binding proteins are associated with clinical outcomes revealed by proteomic measurements in patients with colorectal cancer. Molecular Cancer *11*, 38.

Nemeth, A., Guibert, S., Tiwari, V. K., Ohlsson, R., and Langst, G. (2008). Epigenetic regulation of TTF-I-mediated promoter-terminator interactions of rRNA genes. EMBO J *27*, 1255-1265.

Nemeth, A., and Langst, G. (2011). Genome organization in and around the nucleolus. Trends Genet *27*, 149-156.

O'Mahony, D. J., Xie, W. Q., Smith, S. D., Singer, H. A., and Rothblum, L. I. (1992). Differential phosphorylation and localization of the transcription factor UBF in vivo in response to serum deprivation. In vitro dephosphorylation of UBF reduces its transactivation properties. J Biol Chem *267*, 35-38.

Ohno, M., Fukagawa, T., Lee, J. S., and Ikemura, T. (2002). Triplex-forming DNAs in the human interphase nucleus visualized in situ by polypurine/polypyrimidine DNA probes and antitriplex antibodies. Chromosoma *111*, 201-213.

Olivas, W. M., and Maher, L. J., 3rd (1995a). Competitive triplex/quadruplex equilibria involving guanine-rich oligonucleotides. Biochemistry *34*, 278-284.

Olivas, W. M., and Maher, L. J., 3rd (1995b). Overcoming potassium-mediated triplex inhibition. Nucleic Acids Res *23*, 1936-1941.

Ottaviani, D., Lever, E., Takousis, P., and Sheer, D. (2008). Anchoring the genome. Genome Biol *9*, 201.

Paes, H. M., and Fox, K. R. (1997). Kinetic studies on the formation of intermolecular triple helices. Nucleic Acids Res *25*, 3269-3274.

Panayotatos, N., and Wells, R. D. (1981). Cruciform structures in supercoiled DNA. Nature *289*, 466-470.

Pardoll, D. M., and Vogelstein, B. (1980). Sequence analysis of nuclear matrix associated DNA from rat liver. Exp Cell Res *128*, 466-470.

Pederson, T. (2000). Half a century of "the nuclear matrix". Mol Biol Cell *11*, 799-805.

Piergentili, R., and Mencarelli, C. (2008). Drosophila melanogaster kl-3 and kl-5 Y-loops harbor triple-stranded nucleic acids. J Cell Sci *121*, 1605-1612.

Pilch, Brousseau, R., and Shafer, R. H. (1990). Thermodynamics of triple helix formation: spectrophotometric studies on the d(A)10.2d(T)10 and d(C+3T4C+3).d(G3A4G3).d(C3T4C3) triple helices. Nucleic Acids Res *18*, 5743-5750.

Postel, E. H., Flint, S. J., Kessler, D. J., and Hogan, M. E. (1991). Evidence that a triplex-forming oligodeoxyribonucleotide binds to the c-myc promoter in HeLa cells, thereby reducing c-myc mRNA levels. Proc Natl Acad Sci U S A *88*, 8227-8231.

Potaman, V. N., and Sinden, R. R. (1998). Stabilization of intramolecular triple/single-strand structure by cationic peptides. Biochemistry *37*, 12952-12961.

Qiao, F., and Cech, T. R. (2008). Triple-helix structure in telomerase RNA contributes to catalysis. Nat Struct Mol Biol *15*, 634-640.

Radichev, I., Parashkevova, A., and Anachkova, B. (2005). Initiation of DNA replication at a nuclear matrix-attached chromatin fraction. J Cell Physiol *203*, 71-77.

Reeves, R. (1984). Transcriptionally active chromatin. Biochim Biophys Acta *782*, 343-393.

Reyes, J. C., Muchardt, C., and Yaniv, M. (1997). Components of the human SWI/SNF complex are enriched in active chromatin and are associated with the nuclear matrix. J Cell Biol *137*, 263-274.

Rinn, J. L., and Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem *81*, 145-166.

Roberts, and Crothers (1991). Specificity and stringency in DNA triplex formation. Proc Natl Acad Sci USA *88*, 9397-9401.

Roberts, R. W., and Crothers, D. M. (1992). Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. Science *258*, 1463-1466.

Roche (2011). LabFAQs: Roche).

Rougée, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangeli, C., Garestier, and Hélène, C. (1992). Kinetics and thermodynamics of triple-helix formation: effects of ionic strength and mismatches. Biochemistry *31*, 9269-9278.

Sakamoto, N., Chastain, P. D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J. D., and Wells, R. D. (1999). Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia. Mol Cell *3*, 465-475.

Sambrook, J. J., and Russell, D. D. W. (2001). Molecular cloning: a laboratory manual. Vol. 2: Cold Spring Harbor laboratory Press).

Sandor, Z., and Bredberg, A. (1995). Triple helix directed psoralen adducts induce a low frequency of recombination in an SV40 shuttle vector. Biochim Biophys Acta *1263*, 235-240.

Santamaría, A., Castellanos, E., Gómez, V., Benedit, P., Renau-Piqueras, J., Morote, J., Reventós, J., Thomson, T. M., and Paciucci, R. (2005). PTOV1 enables the nuclear translocation and mitogenic activity of flotillin-1, a major protein of lipid rafts. Mol Cell Biol *25*, 1900-1911.

Santamaría, A., Fernández, P. L., Farré, X., Benedit, P., Reventós, J., Morote, J., Paciucci, R., and Thomson, T. M. (2003). PTOV-1, a novel protein overexpressed in prostate cancer, shuttles between the cytoplasm and the nucleus and promotes entry into the S phase of the cell division cycle. Am J Pathol *162*, 897-905.

Santoro, R., and Grummt, I. (2001). Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. Mol Cell *8*, 719-725.

Santoro, R., and Grummt, I. (2005). Epigenetic mechanism of rRNA gene silencing: temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation. Mol Cell Biol *25*, 2539-2546.

Santoro, R., Li, J., and Grummt, I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. Nat Genet *32*, 393-396.

Santoro, R., Schmitz, K. M., Sandoval, J., and Grummt, I. (2010). Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. EMBO Rep *11*, 52-58.

Schmitz, K. M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes Dev *24*, 2264-2269.

Schroth, G. P., and Ho, P. S. (1995). Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA. Nucleic Acids Res *23*, 1977-1983.

Schubert, T., Pusch, M. C., Diermeier, S., Benes, V., Kremmer, E., Imhof, A., and Langst, G. (2012). Df31 protein and snoRNAs maintain accessible higher-order structures of chromatin. Mol Cell *48*, 434-444.

Schwarzacher, H. G., and Mosgoeller, W. (2000). Ribosome biogenesis in man: current views on nucleolar structures and function. Cytogenet Cell Genet *91*, 243-252.

Segal, E., and Widom, J. (2009). Poly(dA:dT) tracts: major determinants of nucleosome organization. Curr Opin Struct Biol *19*, 65-71.

Seidel, S. A., Dijkman, P. M., Lea, W. A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J. S., Srinivasan, P., Baaske, P., Simeonov, A., *et al.* (2013). Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. Methods *59*, 301-315.

Seither, P., Zatsepina, O., Hoffmann, M., and Grummt, I. (1997). Constitutive and strong association of PAF53 with RNA polymerase I. Chromosoma *106*, 216-225.

Sen, D., and Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature *334*, 364-366.

Shindo, H., Kamiya, M., Torigoe, H., and Sarai, A. (1993). Mechanism of DNA triplex formation and its specificity as studied by filter binding assay. Nucleic Acids Symp Ser, 17-18.

Shiue, C. N., Berkson, R. G., and Wright, A. P. (2009). c-Myc induces changes in higher order rDNA structure on stimulation of quiescent cells. Oncogene *28*, 1833-1842.

Simon, D. N., and Wilson, K. L. (2011). The nucleoskeleton as a genome-associated dynamic 'network of networks'. Nat Rev Mol Cell Biol *12*, 695-708.

Singleton, S. F., and Dervan, P. B. (1993). Equilibrium association constants for oligonucleotide-directed triple helix formation at single DNA sites: linkage to cation valence and concentration. Biochemistry *32*, 13171-13179.

Smith, H. C., and Rothblum, L. I. (1987). Ribosomal DNA sequences attached to the nuclear matrix. Biochem Genet *25*, 863-879.

Soragni, E., Herman, D., Dent, S. Y., Gottesfeld, J. M., Wells, R. D., and Napierala, M. (2008). Long intronic GAA*TTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedreich ataxia. Nucleic Acids Res *36*, 6056-6065.

Stephanova, E., Stancheva, R., and Avramova, Z. (1993). Binding of sequences from the 5'- and 3'-nontranscribed spacers of the rat rDNA locus to the nucleolar matrix. Chromosoma *102*, 287-295.

Stollar, B. D., and Raso, V. (1974). Antibodies recognise specific structures of triple-helical polynucleotides built on poly(A) or poly(dA). Nature *250*, 231-234.

Strohner, R., Nemeth, A., Jansa, P., Hofmann-Rohrer, U., Santoro, R., Langst, G., and Grummt, I. (2001). NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. EMBO J *20*, 4892-4900.

Strohner, R., Nemeth, A., Nightingale, K. P., Grummt, I., Becker, P. B., and Langst, G. (2004). Recruitment of the nucleolar remodeling complex NoRC establishes ribosomal DNA silencing in chromatin. Mol Cell Biol *24*, 1791-1798.

Suda, T., Mishima, Y., Takayanagi, K., Asakura, H., Odani, S., and Kominami, R. (1996). A novel activity of HMG domains: promotion of the triple-stranded complex formation between DNA containing (GGA/TCC)11 and d(GGA)11 oligonucleotides. Nucleic Acids Res *24*, 4733-4740.

Sugimoto, N., Wu, P., Hara, H., and Kawamoto, Y. (2001). pH and cation effects on the properties of parallel pyrimidine motif DNA triplexes. Biochemistry *40*, 9396-9405.

Suter, B., Schnappauf, G., and Thoma, F. (2000). Poly(dA.dT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters in vivo. Nucleic Acids Res *28*, 4083-4089.

Todd, A. K., Johnston, M., and Neidle, S. (2005). Highly prevalent putative quadruplex sequence motifs in human DNA. Nucleic Acids Res *33*, 2901-2907.

Vasquez, K. M., Marburger, K., Intody, Z., and Wilson, J. H. (2001). Manipulating the mammalian genome by homologous recombination. Proc Natl Acad Sci U S A *98*, 8403-8410.

Vasquez, K. M., Narayanan, L., and Glazer, P. M. (2000). Specific mutations induced by triplex-forming oligonucleotides in mice. Science *290*, 530-533.

Vasquez, K. M., Wensel, T. G., Hogan, M. E., and Wilson, J. H. (1995). High-affinity triple helix formation by synthetic oligonucleotides at a site within a selectable mammalian gene. Biochemistry *34*, 7243-7251.

Vecerova, J., Koberna, K., Malinsky, J., Soutoglou, E., Sullivan, T., Stewart, C. L., Raska, I., and Misteli, T. (2004). Formation of nuclear splicing factor compartments is independent of lamins A/C. Mol Biol Cell *15*, 4904-4910.

Vojnic, E., Mourao, A., Seizl, M., Simon, B., Wenzeck, L., Lariviere, L., Baumli, S., Baumgart, K., Meisterernst, M., Sattler, M., and Cramer, P. (2011). Structure and VP16 binding of the Mediator Med25 activator interaction domain. Nat Struct Mol Biol *18*, 404-409.

Volker, J., Botes, D. P., Lindsey, G. G., and Klump, H. H. (1993). Energetics of a stable intramolecular DNA triple helix formation. J Mol Biol *230*, 1278-1290.

Völker, J., and Klump, H. H. (1994). Electrostatic effects in DNA triple helices. Biochemistry *33*, 13502-13508.

Wang, A. H., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G., and Rich, A. (1979). Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature *282*, 680-686.

Wang, G., and Vasquez, K. M. (2006). Non-B DNA structure-induced genetic instability. Mutat Res *598*, 103-119.

Watson, J. D., and Crick, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-738.

Wells, R. D. (2008). DNA triplexes and Friedreich ataxia. FASEB J *22*, 1625-1634.

Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., and Wohlrab, F. (1988). The chemistry and biology of unusual DNA structures adopted by oligopurine.oligopyrimidine sequences. FASEB J *2*, 2939-2949.

Westin, L., Blomquist, P., Milligan, J. F., and Wrange, O. (1995). Triple helix DNA alters nucleosomal histone-DNA interactions and acts as a nucleosome barrier. Nucleic Acids Res *23*, 2184-2191.

Wilusz, J. E., Jnbaptiste, C. K., Lu, L. Y., Kuhn, C. D., Joshua-Tor, L., and Sharp, P. A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. Genes Dev.

Wu, P., Kawamoto, Y., Hara, H., and Sugimoto, N. (2002). Effect of divalent cations and cytosine protonation on thermodynamic properties of intermolecular DNA double and triple helices. J Inorg Biochem *91*, 277-285.

Wu, Q., Gaddis, S. S., MacLeod, M. C., Walborg, E. F., Thames, H. D., DiGiovanni, J., and Vasquez, K. M. (2007). High-affinity triplex-forming oligonucleotide target sequences in mammalian genomes. Mol Carcinog *46*, 15-23.

Yamamoto, K., Yamamoto, M., Hanada, K., Nogi, Y., Matsuyama, T., and Muramatsu, M. (2004). Multiple protein-protein interactions by RNA polymerase I-associated factor PAF49 and role of PAF49 in rRNA transcription. Mol Cell Biol *24*, 6338-6349.

Yao, H., Brick, K., Evrard, Y., Xiao, T., Camerini-Otero, R. D., and Felsenfeld, G. (2010). Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. Genes Dev *24*, 2543-2555.

Zhao, J., Bacolla, A., Wang, G., and Vasquez, K. M. (2010). Non-B DNA structure-induced genetic instability and evolution. Cell Mol Life Sci *67*, 43-62.

Zheng, R., Shen, Z., Tripathi, V., Xuan, Z., Freier, S. M., Bennett, C. F., Prasanth, S. G., and Prasanth, K. V. (2010). Polypurine-repeat-containing RNAs: a novel class of long non-coding RNA in mammalian cells. J Cell Sci *123*, 3734-3744.

Zhou, Y., Santoro, R., and Grummt, I. (2002). The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. EMBO J *21*, 4632-4640.

Zhou, Y., Schmitz, K. M., Mayer, C., Yuan, X., Akhtar, A., and Grummt, I. (2009). Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. Nat Cell Biol *11*, 1010-1016.

Zillner, K. (2013) Nucleolar Epigenomics by Dynamic Molecular Combing and Nucleolar Matrix Analysis, PhD, Regensburg, Regensburg.

Zillner, K., Filarsky, M., Rachow, K., Weinberger, M., Langst, G., and Nemeth, A. (2013). Large-scale organization of ribosomal DNA chromatin is regulated by Tip5. Nucleic Acids Res.

Zillner, K., Jerabek-Willemsen, M., Duhr, S., Braun, D., Langst, G., and Baaske, P. (2012). Microscale thermophoresis as a sensitive method to quantify protein: nucleic acid interactions in solution. Methods Mol Biol *815*, 241-252.

Zolotukhin, A. S., Uranishi, H., Lindtner, S., Bear, J., Pavlakis, G. N., and Felber, B. K. (2009). Nuclear export factor RBM15 facilitates the access of DBP5 to mRNA. Nucleic Acids Res *37*, 7151-7162.