

Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

Structural and biochemical characterization  
of ctTel1, an ATM kinase ortholog from  
*Chaetomium thermophilum*



Marijke Jansma

aus

Groningen, die Niederlande

2020



## **Erklärung**

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Karl-Peter Hopfner betreut.

## **Eidesstattliche Versicherung**

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 4. Februar 2020,

.....

Marijke Jansma

Dissertation eingereicht am: 6. Februar 2020

1. Gutachter: Prof. Dr. rer. nat. Karl-Peter Hopfner

2. Gutachter: Prof. Dr. rer. nat. Roland Beckmann

Mündliche Prüfung am: 22. April 2020



This thesis was prepared from October 2015 to February 2020 in the laboratory of Prof. Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-Universität München.

## **Publications**

During the work of this thesis, the following paper was published:

Jansma M, Linke-Winnebeck C, Eustermann S, Lammens K, Kostrewa D, Stakyte K, Litz C, Kessler B & Hopfner K-P (2020) Near-Complete Structure and Model of Tel1ATM from *Chaetomium thermophilum* Reveals a Robust Autoinhibited ATP State. *Structure* 28, 1-13.



# Contents

Summary . . . . .	VII
<b>1 Introduction</b>	<b>1</b>
1.1 DNA damage and repair . . . . .	1
1.1.1 Origins and forms of DNA damage . . . . .	1
1.1.2 Repair pathways . . . . .	2
1.1.3 DNA double strand breaks . . . . .	3
1.1.3.1 Physiological DSBs . . . . .	3
1.1.3.2 Endogenous DSBs . . . . .	4
1.1.3.3 Exogenous DSBs . . . . .	4
1.2 DSB repair pathways . . . . .	5
1.2.1 Homologous recombination . . . . .	6
1.2.2 Non-homologous end-joining . . . . .	7
1.2.3 Repair pathway choice . . . . .	8
1.3 DNA damage signalling in DSB repair: the MRN-ATM axis . . . . .	9
1.3.1 MRN complex: structure and functions . . . . .	9
1.3.1.1 Structure . . . . .	9
1.3.1.2 Functions in vitro . . . . .	11
1.3.1.3 Functions in vivo . . . . .	11

1.4	Signalling in DSB repair: ATM kinase . . . . .	12
1.4.1	Discovery of ATM and its role in DNA repair . . . . .	12
1.4.2	The PIKK family of signalling kinases . . . . .	13
1.4.3	PIKK family members . . . . .	15
1.4.4	Cellular function of ATM . . . . .	16
1.4.5	ATM activation . . . . .	19
1.4.6	Regulation of ATM activity . . . . .	20
1.4.7	Aim of the project . . . . .	22
<b>2</b>	<b>Materials and Methods</b>	<b>23</b>
2.1	Materials . . . . .	23
2.1.1	Oligonucleotides . . . . .	23
2.1.2	Plasmids . . . . .	24
2.1.3	Strains . . . . .	25
2.1.4	Antibodies . . . . .	25
2.1.5	Common buffers, media and antibiotics . . . . .	25
2.1.6	Software and algorithms . . . . .	26
2.2	Methods . . . . .	26
2.2.1	Molecular Biology Methods . . . . .	26
2.2.1.1	Molecular cloning . . . . .	26
2.2.1.2	Site-directed mutagenesis . . . . .	27
2.2.2	Microbiology methods: <i>E. coli</i> . . . . .	28
2.2.2.1	Generation of competent <i>E. coli</i> . . . . .	28
2.2.2.2	Transformation into <i>E. coli</i> . . . . .	28
2.2.2.3	Isolation of plasmid DNA . . . . .	29

2.2.2.4	Purification of MultiBacmid . . . . .	29
2.2.3	Microbiology methods: <i>C. thermophilum</i> . . . . .	29
2.2.3.1	Culturing <i>C. thermophilum</i> on agar plates . . . . .	29
2.2.3.2	Culturing <i>C. thermophilum</i> in liquid medium . . . . .	29
2.2.3.3	Harvesting and lysis of mycelium . . . . .	30
2.2.3.4	Generation of spores . . . . .	30
2.2.4	Insect cell methods . . . . .	30
2.2.4.1	Production of baculovirus in Sf21 cells . . . . .	30
2.2.5	Protein biochemistry methods . . . . .	31
2.2.5.1	Protein expression in <i>E. coli</i> . . . . .	31
2.2.5.2	Test expression ctTel1 in HighFive insect cells . . . . .	31
2.2.5.3	Purification of ctTel1(1-1864) . . . . .	31
2.2.5.4	Purification of nanobodies against ctTel1 . . . . .	32
2.2.5.5	Test expression of recombinant ctTel1 . . . . .	32
2.2.5.6	Purification of endogenous ctTel1 . . . . .	33
2.2.5.7	Denaturing polyacrylamide gel electrophoresis (SDS-PAGE) . . . . .	33
2.2.5.8	Western blot analysis . . . . .	34
2.2.6	Assays . . . . .	34
2.2.6.1	Annealing of DNA . . . . .	34
2.2.6.2	Affinity measurement by fluorescence anisotropy with DNA . . . . .	34
2.2.6.3	Affinity measurement by fluorescence anisotropy with FAM-Nbs1 peptide . . . . .	35
2.2.6.4	Pull-down assays . . . . .	35
2.2.6.5	Analytical gel filtration . . . . .	35
2.2.6.6	Kinase activity assays . . . . .	36

2.2.7	Cryo-EM . . . . .	36
2.2.7.1	Cryo-EM grid preparation . . . . .	36
2.2.7.2	Cryo-EM data acquisition . . . . .	36
2.2.8	Cryo-EM data analysis . . . . .	37
2.2.8.1	Data analysis . . . . .	37
2.2.8.2	Model building . . . . .	38
2.2.9	Bioinformatics and software . . . . .	39
<b>3</b>	<b>Results</b>	<b>40</b>
3.1	Protein biochemistry . . . . .	40
3.1.1	Purification ctTel1(1-1864) . . . . .	40
3.1.2	Purification of endogenous ctTel1 using nanobodies . . . . .	41
3.1.3	Expression trials and purification of recombinant ctTel1 . . . . .	43
3.2	Binding assays ctTel1(1-1864) . . . . .	45
3.2.1	DNA binding properties . . . . .	45
3.2.2	Nbs1 binding properties . . . . .	46
3.2.3	Pull-down assay with ctMRN( $\Delta$ CC) and ctMR( $\Delta$ CC) . . . . .	46
3.2.4	Analytical gel filtration . . . . .	47
3.3	Kinase activity of ctTel1 . . . . .	49
3.3.1	Initial characterisation kinase activity . . . . .	49
3.3.2	Kinase inhibitors . . . . .	50
3.4	Cryo-electron microscopy . . . . .	51
3.4.1	Structure determination of ctTel1 . . . . .	51
3.4.2	Overall architecture of ctTel1 . . . . .	55
3.4.3	Structure of the N-terminal $\alpha$ -solenoid . . . . .	57

3.4.4	FATKIN domain . . . . .	60
3.4.5	Dimer interface . . . . .	62
3.4.6	Kinase domain and active site . . . . .	64
3.4.7	Potential DNA binding sites . . . . .	67
3.4.8	Mapping of disease mutations . . . . .	68
<b>4</b>	<b>Discussion</b>	<b>71</b>
4.1	The recombinant expression of ctTel1 . . . . .	71
4.2	Binding activities of ctTel1: interaction with DNA and MRN . . . . .	73
4.2.1	Interaction with MRN . . . . .	73
4.2.2	Interaction with DNA . . . . .	76
4.3	Structural biology . . . . .	78
4.3.1	Comparison to other ATM kinase structures . . . . .	78
4.3.1.1	Comparison $\alpha$ -solenoid . . . . .	78
4.3.1.2	FATKIN domain . . . . .	79
4.3.1.3	Kinase domain . . . . .	80
4.3.2	Comparison to other PIKKs . . . . .	82
4.3.2.1	Comparison to ATR . . . . .	82
4.3.2.2	Comparison to DNA-PKcs . . . . .	85
4.3.2.3	Comparison to mTOR . . . . .	86
4.3.2.4	Comparison to SMG1 . . . . .	87
4.3.2.5	Comparison to Tra1 . . . . .	88
4.3.3	Summary . . . . .	89
4.4	A model for ATM activation . . . . .	89
4.5	Conclusion and outlook . . . . .	91

List of abbreviations . . . . .	XXI
Appendix: Multiple sequence alignment of ATM kinases . . . . .	XXV
Acknowledgements . . . . .	XXXI

## Summary

Ataxia-Telangiectasia Mutated (ATM) is an apical signalling kinase that responds to DNA double strand breaks (DSBs). DSBs are the most dangerous form of DNA damage. Therefore cells have evolved multiple pathways to repair such lesions and signal the damage to the cell to organise the DNA damage response (DDR). DSBs are sensed by the Mre11-Rad50-Nbs1 (MRN) complex, which initiates repair via homologous recombination (HR). At the beginning of this process, MRN activates the ATM kinase, which phosphorylates hundreds of substrates, including other DNA repair factors and kinases to orchestrate a cell-wide DNA damage response.

ATM belongs to the family of PI3-kinase like kinases (PIKKs). It has a very large HEAT-repeat domain, which plays a role in interacting with other proteins, and a highly conserved kinase domain, showing more homology to PI3-lipid kinases than protein kinases. ATM is normally present in the nucleus as an autoinhibited dimer, but rapidly becomes active upon DNA damage. As a part of its activation mechanism ATM is postulated to monomerise and become autophosphorylated.

In order to understand the molecular mechanism of ATM activation better, the aim of this work was to solve the structure of ctTel1, a fungal ATM ortholog from the mould *Chaetomium thermophilum*, using cryo-electron microscopy (cryo-EM). Due the large size of the protein (650 kDa as a dimer) recombinant expression was not possible. Therefore ctTel1 was purified from endogenous sources. The other aim was to characterise its interaction with MRN and DNA, for which a truncated ctTel1 construct was used which could be obtained in large quantities.

Using the truncated construct, the interactions between ATM and Nbs1, and the MRN complex, were characterised using pull-downs, analytical gel filtration and fluorescence anisotropy, showing that it forms a stable complex. Furthermore we found that truncated ctTel1 has DNA binding properties, preferring blunt-ended double stranded DNA and displaying a length-dependency.

Using cryo-EM we determined the first complete and the highest resolution structure of any ATM kinase to date, with the kinase domain at 2.8 Å and an overall structure including the HEAT repeat domain at 3.7 Å. This allowed for the building of a complete atomic model, except for a number of flexible loops. The structure reveals a hydrophobic dimer interface, suggesting monomerisation is unlikely to be an aspect of the activation mechanism.

In the active site of the kinase the loops required for catalysis are structured, coordinating ATP $\gamma$ S and a Mg<sup>2+</sup> ion. This suggests that ATM is in a catalytic proficient state. However, access to the active site is physically blocked by the PIKK regulatory domain (PRD) a regulatory helix. The PRD thus functions as a pseudosubstrate, suggesting that part of the activation mechanism entails its removal in response to allosteric binding of activators. Cancer mutations and ataxia-telangiectasia (A-T) mutations were mapped onto the structure, revealing the elements which are important for the activation. Taken together, the ctTel1 structure reveals the autoinhibitory circuitry and provides the first step towards understanding the activation mechanism of this important kinase.

# Chapter 1

## Introduction

### 1.1 DNA damage and repair

All information that is needed for the development and function of any living organism is encoded in deoxyribonucleic acid (DNA). Therefore it is of high importance that this genomic information is maintained. The DNA molecule is vulnerable to insults coming from both external sources and internal sources. Due to its function in storing information, errors in the DNA will last much longer than other biomolecules, which only have short life spans in the cell, and could even be transferred to daughter cells. Furthermore, DNA damage can interfere with transcription and replication. Unsurprisingly, DNA damage underlies many diseases, most famously cancer, and also ageing. To maintain genomic integrity, cells in all domains of life have developed mechanisms to repair DNA damage.<sup>1</sup>

#### 1.1.1 Origins and forms of DNA damage

As there are many different types of DNA damage, different mechanisms have evolved to resolve the different types of lesions that can occur. This DNA damage response (DDR) consists out of multiple highly coordinated steps, which start with sensing of the DNA damage, followed by damage signalling and ends with the coordinated repair of the lesion.<sup>2</sup>

Many factors, whether physical or chemical, can lead to chemical modifications and structural changes in the DNA. These can be extrinsic, such as radiation or toxins in fumes, or intrinsic, such as reactive oxygen species (ROS) which occur as by-products of cellular metabolism. Cells can also introduce DNA damage in a programmed manner in order to increase genetic diversity, for example in meiosis.

DNA damage can occur spontaneously within the chemical environment of the cell. ROS that are generated during cellular respiration can lead to multiple forms of DNA damage, such as base modifications, single strand breaks, double strand breaks and protein-DNA linkages.<sup>3,4</sup> Abasic sites in the DNA can occur due to spontaneous hydrolysis of the bases.<sup>5</sup> Another common endogenous source of DNA damage is replication, in which DNA polymerases misincorporate nucleotides, or incorporate chemically altered precursors.<sup>2,6</sup>

Exogenous sources of DNA damage, such as physical stress, ionizing radiation (IR), ultra-violet (UV) light and toxic chemicals, for example found in smoke, can also lead to a wide variety of lesions.<sup>7</sup> DNA damaging chemicals can lead to single strand breaks, double strand breaks and even lead to covalent linkages of protein to the DNA. UV light often induces pyrimidine dimers, in which two bases are covalently linked. IR can induce damage in a multitude of direct and indirect ways, either by generating ROS or by directly causing breaks in the DNA.<sup>8</sup>

### 1.1.2 Repair pathways

To repair this wide variety of different lesions that arise, cells evolve a multitude of different DNA repair pathways.

Minor modifications in the DNA that do not distort the structure of the DNA or base pairing, for example alkylations and certain photoproducts induced by UV-light, can be repaired by Direct DNA-lesion reversal. This is typically mediated by only one enzyme, directly removing the modification without further alterations/cutting in the DNA.<sup>8</sup>

During replication, erroneous base pairing mismatches can occur. Although mismatches do not affect any processes occurring on the DNA, they can lead to permanent mutations as the sequence is altered. Such lesions, and also other insertions and deletions occurring during replication are resolved via mismatch repair (MMR).<sup>8,9</sup>

The insertion of an abnormal DNA base or a simple base adduct is typically resolved by Base Excision Repair (BER) and single-strand break repair (SSBR). The damaged base is removed in a coordinated process of different enzymes that recognize the damaged base, mediate its excision, fill the gap and ligate the break.<sup>8,10</sup>

More complex forms of DNA damage, such as inter- and intra-strand crosslinks which induces distortions in the DNA helix are repaired by Nucleotide Excision Repair (NER). Once a distortion in the DNA is sensed, duplex melting is induced and up to 30 bases are excised around the lesion. The resulting gap is filled in by polymerase I and then ligated. NER is divided in two sub-pathways depending on the lesion recognition: Transcription coupled repair (TRC) if occurring cotranslationally, where the DNA damage is sensed by RNA Polymerase II, and global-genome NER.<sup>8,11</sup> Furthermore, inter-strand DNA crosslinks can also be repaired via the Fanconi Anaemia (FANC) pathway.<sup>12</sup>

Translesion DNA synthesis (TLS) is rather a damage-tolerance mechanism than a direct DNA repair mechanism. This process takes place if certain lesions persist and block genome duplication. To avoid stalled replication forks, which are highly detrimental to a cell, specialised polymerases can skip over the damage to give the cell more time to repair the lesion.<sup>13</sup>

For double strand break (DSB) repair, multiple pathways are available depending on the cell cycle: homologous recombination (HR) which depends on the presence of a sister chromatid as a template, and non-homologous end joining (NHEJ) which involves the direct ligation of DNA ends. These pathways will be discussed in more detail below, as the proteins that signal these forms of damage are the focus of this thesis.<sup>8</sup>

### 1.1.3 DNA double strand breaks

DSBs involve breaks closely located to each other in the sugar-phosphate backbone of both strands of the DNA chain. DSBs are considered to be the most dangerous form of DNA damage. If left unrepaired, DSBs can lead to chromosomal rearrangements, aneuploidy and cell death. One DSB is sufficient to induce apoptosis. Due to their deleterious nature, DSBs are often at the root of the development of cancer and other diseases.<sup>1</sup>

DSBs can arise spontaneously both due to various extra- or intracellular factors. Additionally, numerous physiological processes involve DSB formation in a tightly programmed manner, as will be discussed in more detail below.

#### 1.1.3.1 Physiological DSBs

In a physiological context, cells generate DSBs as intermediates in the recombination and rearrangement of genetic material within or between chromosomes.

In sexually reproducing organisms, DSBs are generated at specific hotspots by SPOrulation 11 (Spo11), a topoisomerase II like protein, during meiotic prophase I. The break is repaired through homologous recombination. Through the formation of cross-overs genetic material is exchanged between the homologous chromosomes, leading to genetic diversification.<sup>14</sup>

Programmed DSBs and genetic diversification are also essential in V(D)J and class switch recombination to generate immunoglobulins (Igs) and T-cell receptors (TCRs), which are expressed in B- and T-cells. For the adaptive immune system, it is essential to recognise a wide range of antigens, which requires high genetic diversity. This diversity is obtained by reshuffling of genetic coding segments Variable (V), Diversity (D) and Joining (J) and a signal segment in developing lymphocytes.<sup>15</sup> Recombination is induced by the RAG recombinase, which creates programmed DSBs. This leads to rearrangement of the coding segments, which are ligated via non-homologous end-joining.<sup>16-18</sup>

DSBs are also introduced during mating-type switching in the yeast *Saccharomyces cerevisiae*. This yeast can exist either in a diploid or a haploid form. In its haploid form, yeast have a primitive form of sexual differentiation which is based on the allele present at the MAT locus. Haploid yeast cells can only mate with a cell of an opposite cell type to produce a diploid cell. During mating-type switching, a site-specific DSB is introduced at the MAT locus by the HO-endonuclease, which is resected and repaired via homologous recombination. This way, the mating gene type is exchanged for a silenced copy of the other MAT allele.<sup>19</sup>

### 1.1.3.2 Endogenous DSBs

While programmed DSBs are essential for meiosis and adaptive immunity, DSBs occur more frequently by accident during other cellular processes, especially during DNA replication.<sup>20</sup> The progression of replication can be hindered by many different obstacles in and on the DNA, such as lesions, secondary structure formation in the DNA and other DNA-bound proteins. The replicative stress can lead to replication fork stalling and collapse. This in turn may lead to incomplete replication and exposed DNA ends, triggering a DSB repair response, typically via HR, as there is a sister chromatid available as a template.<sup>21–24</sup>

Cellular processes such as respiration can lead to the formation of ROS which can directly induce many different types of DNA damage, including DSBs. Additionally, ROS-induced single-strand breaks (SSBs) and crosslinks can also lead to replication fork collapse as described above, hereby indirectly inducing DSBs as well.<sup>25</sup>

### 1.1.3.3 Exogenous DSBs

DSBs can be caused by extracellular sources, such as IR, UV light and numerous genotoxic agents.

IR includes high energy radiation such as cosmic radiation, gamma rays and x-rays. IR can damage the DNA directly or indirectly through the formation of ROS. In both cases, SSBs are created which can lead to DSBs either through replication fork collapse as described above, or when two SSBs are formed in close proximity.<sup>26,27</sup>

UV light can induce covalent dimerization between neighbouring pyrimidines, which are normally repaired through NER. If this pathway fails, it can lead to replication stress and DSB formation as described before.<sup>8,28–30</sup>

A vast array of DNA damaging chemicals, or clastogens, have been reported, which induce different forms of DNA damage. Some can induce breaks in the DNA directly, whereas other chemicals induce damage that leads to replication fork collapse. Topoisomerase inhibitors, such as camptothecin, block the ligation step performed by topoisomerase after it cleaves a DNA strand. This way topoisomerase is covalently trapped on the DNA, which can lead to SSBs and DSBs.<sup>31</sup>

Nonribosomal peptides such as bleomycin directly induce breaks in the DNA. The exact mechanism has not been elucidated, but appears to be a process which depends on oxygen and metal ions for catalysis.<sup>32,33</sup>

Many other chemicals induce other types of DNA damage which can lead to DSB generation through replication fork collapse. DNA alkylating agents, such as methyl methanesulfonate (MMS), and crosslinking agents like cisplatin directly modify the DNA and cause replication stress.<sup>34</sup> Chemicals including, N-nitrosamines, polycyclic aromatic hydrocarbons, and aldehydes which are ubiquitous in our environment and are for example found in processed food and smoke, also can lead to the formation of bulky adducts on the DNA and inter-strand crosslinks, causing replication stress and DNA damage as described above.<sup>2,12</sup>

While exposure to these physical and chemical sources is harmful for the cell, ionizing radiation and clastogens can be used in a positive manner in the form of radiation therapy and chemotherapeutics in cancer therapy to kill fast growing cells by damaging their genome.<sup>33,35</sup>

## 1.2 DSB repair pathways

As DSBs are highly detrimental, cells in all kingdoms of life have evolved dedicated repair pathways, which consists out of proteins which sense the DSB, repair the lesion and signal the damage to the cell.

DSBs are repaired via two pathways, namely canonical non-homologous end joining (c-NHEJ), in which the DNA ends are directly ligated, and HR, which depends on resection and a sister chromatid as a template for repair. A third pathway, alternative non-homologous end joining (alt-NHEJ) has characteristics of the two other pathways, involving microhomology with a template close to the site of the lesion, and a limited amount of resection, prior to ligation of the ends (Figure 1.1).<sup>36,37</sup>

Reflecting the physiological functions of DSBs and the danger occurring when such lesions are not repaired, mutations in any component of these repair pathways are typically associated with diseases leading to sterility, immune deficiency, neurodegeneration, congenital anomalies and other developmental disorders and cancer. Some repair proteins are essential, leading to embryonic lethality in the case of mutations.<sup>38</sup>

The double-strand break repair mechanisms described here are specific for eukaryotes.

### 1.2.1 Homologous recombination

HR consists out of a related set of mechanisms to repair the different types of DSBs. Which pathway is chosen depends on the cell type, the phase of cell cycle, and the organism, although the main players in the process highly conserved throughout evolution.<sup>39-41</sup>

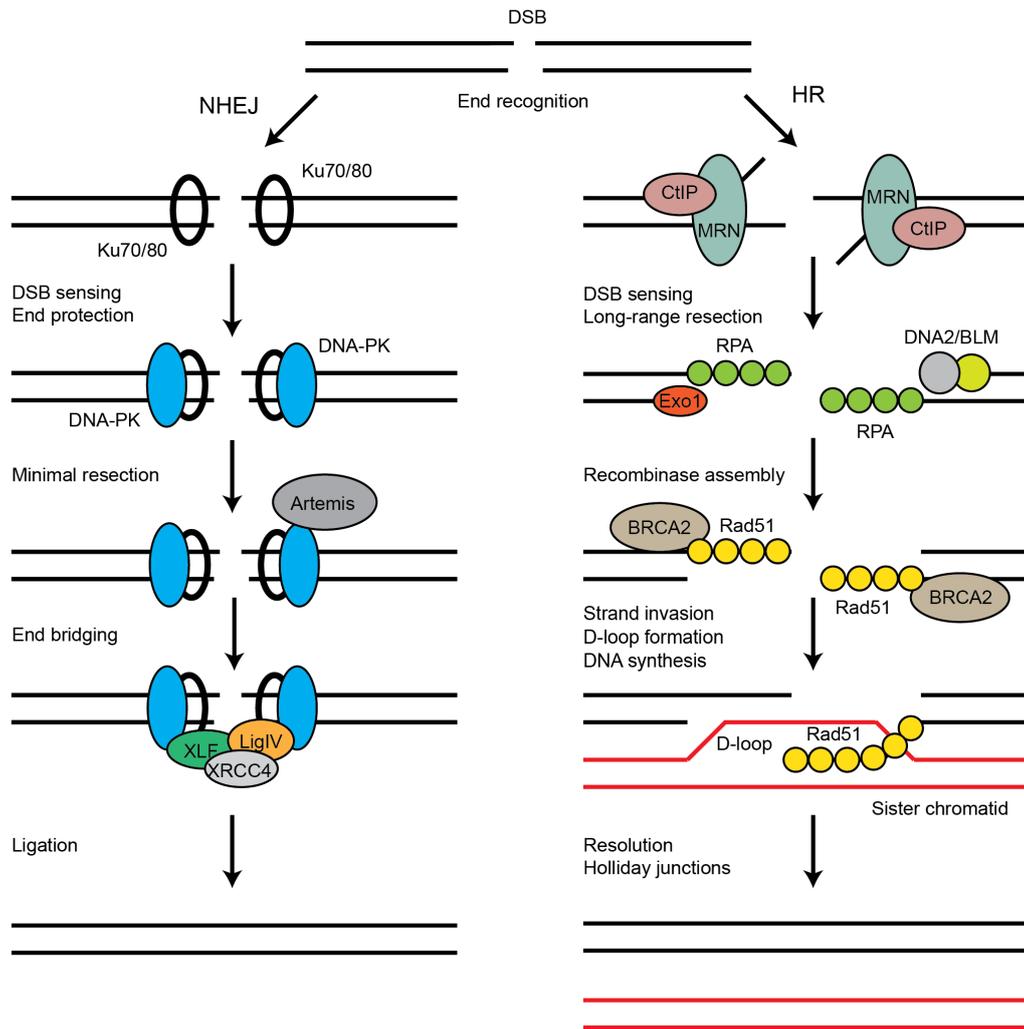
Typically, HR depends on the presence of a non-damaged homologous DNA fragments and therefore occurs primarily in Synthesis phase (S phase), when a sister chromatid is available as a template. Due to the use of a template, HR is suggested to be less mutagenic than NHEJ, in which the ends are directly ligated.<sup>42</sup>

When a DSB occurs, it is sensed by the meiotic recombination 11 - radiation sensitive 50 - Nijmegen breakage syndrome 1 (Mre11-Rad50-Nbs1 or MRN) complex.<sup>43</sup> MRN subsequently recruits Ataxia-Telangiectasia mutated (ATM), a signalling kinase to the lesion, where it becomes activated and phosphorylates many substrates to orchestrate the DNA damage response and halt the cell cycle.<sup>44-46</sup> MRN also initiates resection at the site of the lesion and catalysis together with CtBP-interacting protein (CtIP, Sae2 in yeast) the removal of an oligonucleotide from each 5' end to create a clean template for further resection.<sup>47</sup> Another function of CtIP is the recruitment of breast cancer 1(BRCA1), which stimulates HR further.<sup>48</sup>

The initial short processing is followed by extensive resection in 5'-3' direction. This resection is either performed by the exonuclease exonuclease 1 (Exo1), or by the DNA replication ATP-dependent helicase/nuclease 2 (Dna2) together with the helicases Bloom syndrom protein (BLM) or Werner syndrome ATP-dependent helicase (WRN). This generates a long 3'-overhang which is immediately coated by replication protein A (RPA).<sup>41,49</sup> The coating by RPA leads to removal from secondary structure from the DNA. Furthermore RPA-ssDNA activates the kinase Ataxia telangiectasia and Rad3 related (ATR) which similarly to ATM regulates the DNA damage response. RPA is replaced by Rad51, a process which is mediated by multiple proteins, including BRCA2, which function it is to stabilise the Rad51-ssDNA nucleofilament.<sup>50,51</sup>

The Rad51 nucleofilament then searches for a homologous DNA sequence and subsequently invades the homologous DNA segment, displacing the complementary DNA and forming a displacement loop (D-loop).<sup>52</sup> Rad51 subsequently dissociates from the DNA, exposing a single stranded 3'-prime end for DNA synthesis mediated by polymerase  $\eta$ . This process is followed by ligation by DNA ligase  $\iota$ , yielding a Holliday junction or, if the D-loop undergoes a second-end invasion, a double Holliday junction.<sup>39,53,54</sup>

Such Holliday junctions can be resolved in different ways, resulting either in cross-over or non-cross-over products.<sup>55</sup> Most D-loops are disrupted and then repaired by synthesis-dependent strand annealing (SDSA). If the D-loop undergoes second-end capture, the resulting double Holliday junction is resolved by the action of topoisomerases, which can lead to crossover or non-crossover events.<sup>54</sup>



**Figure 1.1: DSB repair by non-homologous end joining and by homologous recombination.** Depending on the stage of the cell cycle and the nature of the DSB, broken ends are either sensed by Ku or by the MRN complex. In NHEJ, Ku recruits the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) to form DNA-PK. This is followed by minimal processing by Artemis and other factors. The DSBs are ligated by X-ray repair cross-complementing protein 4 (XRCC4), Ligase IV and XRCC4-like factor (XLF). HR consists out of more steps, starting with long-range resection stimulated by MRN and CtIP. The resulting ssDNA is coated by RPA, which is replaced by Rad51 in later steps. Rad51 mediates homology search in a sister chromatid. This way the sister chromatid functions as a template for DNA synthesis to repair the lesion. The resulting Holliday junctions is resolved through multiple mechanism.

### 1.2.2 Non-homologous end-joining

NHEJ is mainly favoured in the Gap 1 phase ( $G_1$  phase) and early S phase when no sister chromatid available. Essentially, in NHEJ the DNA ends are directly ligated, which is considered a more error prone process than HR, due to the absence of a template, and because frequently a few base pairs are deleted in the process. This may lead to the loss of genetic information, although other sources claim that NHEJ is in fact a reliable DNA repair mechanism.<sup>42,56</sup>

In the first step, the DNA break is recognised by Ku70/80, a dimeric ring-shaped complex which has a very high affinity for DNA ends.<sup>57</sup> Binding of Ku protects the ends from premature processing by other enzymes, and mediates the recruitment other NHEJ factors, most importantly DNA-PKcs (catalytic subunit), a large kinase. DNA-PKcs is recruited to the break via the C-terminus of Ku80.<sup>58–60</sup> Together with the Ku-dimer it forms DNA-PK (DNA dependent protein kinase). Once bound to the DNA ends, DNA-PK autophosphorylates and undergoes conformational changes, which leads to the exposure of the DNA ends.<sup>42,61–64</sup>

Due to the great variety of ends that can occur in a DSB, additional steps are required to generate blunt ended DNA required for ligation. This can be done by minimal resection and gap filling. Resection is performed by the nuclease Artemis, which is recruited by DNA-PK and leaves either a blunt end or a short 3' overhang.<sup>65,66</sup> Blunt overhangs can also be obtained by filling in the gap to limit sequence loss, which is catalysed by polymerase  $\mu$  and  $\lambda$  or by terminal deoxynucleotidyl transferase (Tdt). An additional step in end-processing is the addition of a 5' phosphate group by polynucleotide kinase (PNK) to enable ligation.<sup>59,67–70</sup>

The gap between the broken ends is bridged by XLF (XRCC4-like factor), which links DNA ligase IV-XRCC4 complex to DNA-PK and stimulates its activity, which leads to the final ligation of the broken ends.<sup>71–73</sup>

An alternative pathway in NHEJ is alt-NHEJ or microhomology mediated end joining (MMEJ), which is used when there are short stretches of homologous DNA available close to the lesion. This pathway is considered to be highly error-prone and seems to be used mainly during Ig heavy chain class switching in B cells.<sup>74–76</sup>

### 1.2.3 Repair pathway choice

Throughout the years many aspects the mechanisms of double-strand break repair have become elucidated, but it is still not fully understood why cells choose one pathway over the other. Several factors that are known to play a role in the decision are the nature and the origin of the double-strand break, the cell type and the stage of the cell cycle. HR is for example only possible in the S phase, when there is a sister chromatid available, whereas NHEJ is favoured in the ( $G_1$  phase) phase.<sup>39</sup> The type of end found in a DSB seems to be indeed a major determinant, with "dirty" ends, for example ends with protein blocks, often required resection, whereas clean ends can be ligated more easily via NHEJ.<sup>77</sup> Furthermore, the choice of pathway can also be determined by relative abundance of repair factors, posttranslational modifications and the accessibility of the chromatin.<sup>78</sup> Furthermore, there seem to be differences between organisms as well. NHEJ appears to be the preferred DSB repair pathway in mammals, whereas in yeast HR is more common. Why this difference exists is not understood, but may be related to the complexity of the genomes.<sup>79</sup>

The choice between the pathways furthermore appears to involve an intricate interplay between 53BP1-Rif1 and BRCA1-CtIP. NHEJ is stimulated by 53BP1-Rif1, which associates with a double strand break and promotes activity by the NHEJ machinery during (G<sub>1</sub> phase) phase. Furthermore p53-binding protein 1 and Replication Timing Regulatory Factor 1 (53BP1-Rif1) inhibits resection.<sup>80–82</sup> During the S-phase and G2 phase, this complex is inhibited by BRCA1-CtIP, which thereby stimulates HR and enables long range resection.<sup>42,48</sup>

Long range resection by Exo1 and Dna2 requires remodelling of the chromatin around the lesion to enable access to the DNA. This reorganisation, catalysed by chromatin remodellers such as INO80 requiring 80 (Ino80), is possible due to inhibition of 53BP1 by BRCA1.<sup>48,83,84</sup>

The preference for a pathway also depends on the phase of the cell cycle. Cyclin dependent kinases (CDKs) have been found to directly modify proteins involved in either HR or NHEJ, thereby tuning their activities and enabling recruitment of other repair factors.<sup>85–87</sup> CDK dependent phosphorylation of Nbs1 plays a role in colocalization of BRCA1 with MRN, thereby stimulating HR as described above. CDK also phosphorylates Exo1, directly stimulation its nuclease activity.<sup>87</sup>

The cell cycle is furthermore also directly affected by the recruitment and activation of ATM by MRN. ATM phosphorylates many substrates, such as histone H2AX, 53BP1 and effector kinases such as CHK1, thereby directly influencing both the phase of the cell cycle and also the repair pathway choice.<sup>35,44</sup>

Although how cells choose the repair pathway is not entirely clear, it is evident that this involves a high degree of crosstalk between different kinases and other proteins that not only directly affect the cell cycle, but are also affected themselves by the phase of the cell cycle.

### **1.3 DNA damage signalling in DSB repair: the MRN-ATM axis**

The repair processes described above are events that require to be regulated carefully. When a lesion is sensed by the MRN complex, it recruits the ATM kinase to the damage site, where it is activated and orchestrates the DNA damage response. In this section more details about the MRN complex and the ATM kinase are discussed.

#### **1.3.1 MRN complex: structure and functions**

##### **1.3.1.1 Structure**

The MRN complex is one of the central complexes in DNA double strand break repair, in particular in HR, which it initiates upon sensing DNA ends, although MRN also appears to play roles in NHEJ.<sup>43</sup> MRN also performs the first steps towards repair of the lesion and is believed to

tether the DNA ends. Furthermore, MRN signals DNA damage by recruiting the ATM kinase to the lesion through interactions with the very C-terminus of Nbs1 (X-ray sensitive 2 (Xrs2) in *S. cerevisiae*).<sup>60,88</sup> In this course of this process, ATM becomes active and signals damage to the cell to orchestrate the DNA damage response and halt the cell cycle.<sup>89</sup>

Mre11 and Rad50 are conserved in all kingdoms of life, whereas Nbs1 is only found in eukaryotes. MR forms the core of the complex, consisting of a heterotetramer made up out of two Mre11 and two Rad50 molecules. Either one or two copies of Nbs1 are associated with MR.<sup>90,91</sup>

Mre11 interacts with Rad50 and Nbs1, forming the core of the MRN complex. Mre11 is a phosphoesterase which has both endo- and exonuclease activity. The activity requires the presence of two  $Mn^{2+}$  ions.<sup>92,93</sup> Mre11 interacts with the Rad50 coiled-coils via a helix-loop-helix motif at its C-terminus.<sup>94,95</sup> In crystal structures of Mre11 orthologs from eukaryotes, further stabilisation is obtained via a latching loop. This loop becomes ordered upon Nbs1 binding. Nbs1 interacts with Mre11 via its flexible C-terminal tail, binding across the Mre11 dimer and breaking its symmetry.<sup>90,96,97</sup>

Rad50 belongs to the Structural Maintenance of Chromosome (SMC) family of ATP Binding Cassette (ABC) ATPases superfamily, a family which also includes the SMC proteins found in cohesin and condensin. As is characteristic for SMC proteins, Rad50 contains long, rod-like coiled-coils, which connect the N-terminal half of its nucleotide binding domain (NBD) to its C-terminal half. Binding of ATP leads to tight complex of the NBD domains.<sup>92,93,98</sup>

Depending on the domain of life, the length of the coiled-coil varies from 15-60 nm and may be correlated with the complexity of the organisation of the genome.<sup>91</sup> At the apical end of the coiled-coil is a  $Zn^{2+}$ -hook which is coordinated by a conserved CxxC motif. It is believed to play a role in the intermolecular interactions with other MRN complexes.<sup>99-101</sup> Mutations in this  $Zn^{2+}$  hook, as well as deletions of the  $Zn^{2+}$  hook impair Rad50 function, and abolish repair and signalling functions associated with the globular domain, despite the relatively large distance. Deletion of the coiled-coils has the same effect, suggesting that the MRN complex is controlled allosterically.<sup>101,102</sup>

Nbs1 only exists in eukaryotes and has no enzymatic activities, but is important for recruitment of other DNA repair factors, functioning primarily as a binding scaffold. Nbs1 has a structured N-terminus, and a long unstructured C-terminus. The N-terminus contains a Fork-head associated (FHA) domain and two BRCA1 carboxy-terminal (BRCT) in tandem, which play a role in the recruitment of nuclease CtIP, helicase BLM, BRCA1, and checkpoint protein Mediator Of DNA Damage Checkpoint 1 (MDC1), which are all proteins that stimulate HR.<sup>83,103-106</sup> The unstructured C-terminal domain contains a conserved Mre11 binding motif.<sup>90,107</sup> The very C-terminus contains a conserved motif that mediates binding to ATM.<sup>60,88</sup>

### 1.3.1.2 Functions in vitro

Although the MRN complex has been studied for many years, many open questions remain concerning its biochemical activities, which seem to differ between bacteria, archaea and eukaryotes.

Mre11 displays both exonuclease and endonuclease activity. The requirement for ATP in these processes is unclear, as it appears to be required for exonuclease activity in bacteria and archaea, but not in eukaryotes. The ATP hydrolysis rate of Rad50 is rather slow, turning over 0.1 ATP molecule per minute, meaning that Rad50 maintains a relatively long ATP-bound state, although this hydrolysis rate can be stimulated by double stranded DNA. The ATP bound state appears to be important to maintain an DNA end-tethered state, and is required for damage signalling through ATM.<sup>108–112</sup>

ATP binding was found in some crystal structures to induce a conformational change forming a DNA binding groove between the coiled-coils of Rad50. Curiously, this would block the nuclease site of Mre11. The drawback of these studies was that the coiled-coils were removed to enable crystallisation.<sup>94,113–116</sup> A more recent structure of intact *E. coli* MR (SbcCD) shows that Mre11 undergoes a massive conformational change upon DNA binding and ATP hydrolysis, giving the DNA access to the nuclease while the Rad50 coiled-coil forms a rigid rod.<sup>117</sup> Combined with mutational studies of the Zn<sup>2+</sup> hook and the coiled-coils, this suggests a high degree of allosteric regulation within the MRN complex.

Nbs1, which is only present in eukaryotes, regulates the activities of the MR complex, by stimulating DNA binding and processing. For DNA damage signalling it recruits ATM with a conserved C-terminal domain. Interestingly, some studies have shown that ATM can also directly bind to MR in the absence of Nbs1, with ATM activation depending on Rad50.<sup>111,112</sup>

Taken together, the MRN complex is very flexible and is subject to a high degree of allosteric control, which allows the complex to respond adequately to many different types of DSBs. The ATP-bound state of Rad50 and its coiled-coils are important for functions in the globular domain of MRN, including the activation of the ATM kinase.

### 1.3.1.3 Functions in vivo

As described above, the main role of MRN is in double strand break repair, mainly through initiating and mediating HR. MRN is also playing a role in NHEJ, although the interplay with the Ku-dimer, which binds to the end first as part of the first step of NHEJ, remains to be elucidated.<sup>118,119</sup> Indeed, MRN has been shown to be able to cleave Ku-bound ends, possibly due to its endonuclease activity, suggesting there are more factors, dependent on the nature of the break and the cell cycle, that regulate the pathway choice.<sup>120,121</sup>

Resection, which is characteristic for HR, depends largely on CtIP, which is recruited by Nbs1. Depending on the organism, CtIP/Sae2 was shown to directly stimulate the endonuclease activities of Mre11, whereas in higher eukaryotes BRCA1 recruitment via CtIP seems to be more important to promote long range resection.<sup>48,122–124</sup>

Apart from its role in DNA repair, MRN has been found to play a role in the maintenance of telomeres, the capping structures at the ends of eukaryotic chromosomes. They consist out of a repetitive sequence and a single stranded tail, which forms a T-loop. This structure is stabilised by shelterin, which protects the ends from degradation by protein repair machinery. In yeast, MRN associates with the telomeres and recruits the telomerase Cdc13-Stn1-Ten1 (CST) complex. Whether MRN actually processes telomeric DNA, or whether it works only as a scaffold for telomerase is not well understood.<sup>125</sup> In higher eukaryotes, MRN associates with shelterin and plays a role in generating the single stranded DNA that is part of the telomere cap. MRN also activates ATM to orchestrate a DNA damage response to broken or otherwise dysfunctional telomeres.<sup>42,89,126,127</sup>

MRN also plays a role in the repair of programmed DSBs that are induced in meiosis by Spo11. DSBs generated this way have topoisomerases trapped covalently on the DNA. These fusions are removed by MRN and another set of factors which are specific for meiosis.<sup>14</sup>

During replication topoisomerases can also get covalently attached to the DNA, which can lead to replication fork stalling. Several HR factors stabilise these forks to avoid their degradation, which would lead to further genome instability. MRN induced signalling via ATR-ATRIP leads to phosphorylation events that in turn lead to restart of the replication fork.<sup>128–130</sup>

Apart from its direct role in DNA repair and the initiation of resection, MRN furthermore induces the orchestration of a cell-wide DNA damage response through its activation of the ATM kinase, which plays an apical role in the phosphorylation cascade that organises the response. As the ATM kinase is the main topic of this thesis, this will be discussed in detail below.

## 1.4 Signalling in DSB repair: ATM kinase

### 1.4.1 Discovery of ATM and its role in DNA repair

The ATM kinase was originally identified by genetic mutations in the rare autosomal recessive disorder ataxia-telangiectasia (A-T) which is characterised by neurodegeneration leading to ataxia, dilation of the blood vessel (telangiectasia) and genomic instability.<sup>131</sup> Their genomic instability leads to immunodeficiency, radiosensitivity and a high susceptibility to malignancies.<sup>89,132</sup>

The mutations occurred in a very large gene, encoding a gigantic 350 kDa protein (700 kDa as a dimer), which was identified as a kinase that displays homology to the phosphatidylinositol-3 (PI3)

family of lipid kinases, which phosphorylate phosphatidylinositol to generate phospholipid second messengers. However, ATM does not phosphorylate lipids, but targets proteins.<sup>131</sup>

The functional relationship between MRN and ATM was discovered when patients with a syndrome similar to A-T, called A-T Like Disorder (ATLD), were identified to have mutations in Mre11. Cells with this mutated Mre11 showed indeed much lower ATM activation, demonstrating that the MRN complex is needed for ATM activation.<sup>133–135</sup>

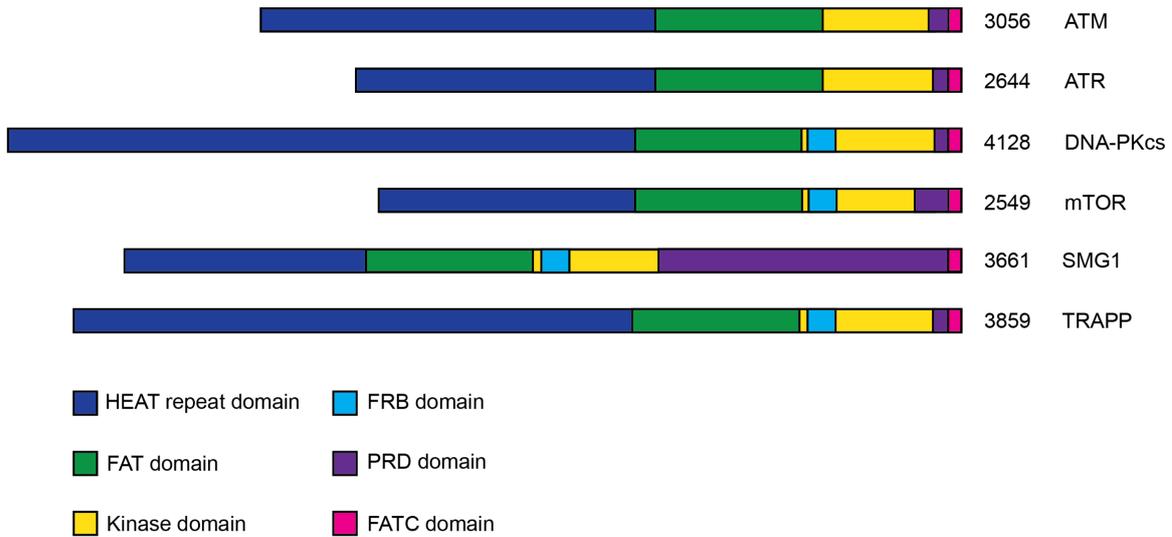
This led to the establishment of ATM as an apical DNA damage signalling kinase, mainly phosphorylating proteins that play a role in genome maintenance and the regulating of the cell cycle. In the light of these findings is not surprising that ATM is frequently mutated in many forms of cancer.<sup>136</sup> Phosphorylations are among the most common post-translational modifications (PTM) in signalling and can have a variety of effects on the target protein. Typically, phosphorylation leads to modulation of substrate activity, protein stability, interactions with other proteins and localisation. A cascade of phosphorylation events therefore can have a wide range of effects on the cell.<sup>45</sup> ATM is a prime example, as (auto)phosphorylation leads to activation, recruitment of other proteins to the lesion and stabilises interactions between DNA repair factors and modulators. This way, ATM orchestrates the DNA damage response.

ATM was found to have an ortholog in lower eukaryotes such as *S. cerevisiae* as well, where the protein was named telomere length regulation protein 1 (Tel1), reflecting its role in telomere maintenance. As in higher eukaryotes, Tel1 is activated by MRX, the fungal homolog of MRN, in response to DSBs, showing that this mechanism is conserved amongst eukaryotes.<sup>137,138</sup>

#### 1.4.2 The PIKK family of signalling kinases

ATM belongs to the family of PI3 like Protein Kinases (PIKKs) together with ATR (denoted Mec1 in *S. cerevisiae*), mammalian target of rapamycin (mTOR, TOR in yeast), DNA-PK (not found in lower eukaryotes), Transformation/transcription domain-associated protein (TRRAP, Tra1 in yeast) and Suppressor with morphogenetic effect on genitalia 1 (SMG1), which only exists in plants and higher eukaryotes. All family members are very large kinases (300 to 500 kDa for monomers) involved in homeostasis surveillance and stress signalling.<sup>89,139,140</sup>

ATR/Mec1 and DNA-PK are also involved in DNA damage signalling, which like ATM and MRN associate with the lesions via DNA binding cofactors. All three are serine/threonine kinases, with a preference for an S/TQ motif.<sup>42,89,141,142</sup> The other PIKKs are involved in other pathways, such as nutrient sensing and the regulation of metabolism. Regulation of metabolism and nutrient sensing is orchestrated by mTOR. SMG1 controls nonsense-mediated RNA decay (NMD). TRAPP/Tra1 is the only pseudokinase amongst PIKKs and is part of multiple protein complexes playing a role in the regulation of transcription.<sup>140</sup>



**Figure 1.2: The domain organisation of PIKKs.** All PIKKs are enormous enzymes. The numbers refer to the length of the amino acid chain in human PIKKs. The HEAT and FAT domains, which form the largest part of the polypeptide chain length, are entirely  $\alpha$ -helical. The kinase domain in some PIKKs contains an FRB domain, which is a regulatory four-helical bundle. All PIKKs have a regulatory PRD and FATC domain at their C-terminus.

All PIKKs share a similar domain architecture, consisting out of a number of conserved elements (Figure 1.2). Apart from their kinase domain, which shows homology to PI3-kinases, they are entirely  $\alpha$ -helical, forming repeat elements. The N-terminal  $\alpha$ -solenoid domain consisting out of a 1000-3000 residues and consists mainly out of Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR1 (HEAT) repeats.<sup>143</sup> In terms of sequence, the  $\alpha$ -solenoid is not highly conserved. However, HEAT repeats are often implied to play a role in interactions with other proteins and DNA, suggesting that the  $\alpha$ -solenoid domain functions as an interaction hub or a scaffold for other proteins.<sup>144</sup> Such a function would reflect the apical roles the kinases play in their pathways they regulate.<sup>140</sup>

The N-terminal  $\alpha$ -solenoid is followed by a sequence-wise more conserved FRAP-ATM-TRRAP (FAT) domain, which plays a role in regulating kinase activity. The FAT domain is followed by a highly conserved kinase domain, which is only 5- 10 % of the overall peptide chain length. The kinase domain contains numerous elements that are specific for PIKKs. At the C-terminus, there is a PIKK Regulatory Domain (PRD) followed by a FATC (C-terminus of FAT) domain. Some PIKKs, but not ATM and ATR, contain an FKBP12-Rapamycin binding (FRB) domain at the beginning of the kinase domain.<sup>89,145,146</sup>

The PRD was discovered to be important for the regulation of kinase activity in ATR.<sup>147</sup> Later it was found to be present in all PIKKs, where it likely has a similar role. The FRB, which

is missing in ATM and ATR, was originally identified in mTOR and is important for binding of rapamycin.<sup>148</sup> Together with the FAT domain, the kinase and its insertions form an integral domain often denoted as FATKIN.<sup>140,146</sup> The crystal structure of the mTOR FATKIN domain shows how the FAT domain indeed wraps around the kinase, showing how the FAT domain could directly regulate the kinase.<sup>149</sup> In addition, many elements in the active site are highly conserved, such as the catalytic loop, the activation loop and the P-loop which shields the active site. Furthermore, the kinase domain contains an Lethal with Sec Thirteen (LST8) binding element (LBE), a helical insert in the C-terminal domain, which binds to LST8 in mTOR, but may also have regulatory roles in other PIKKs.<sup>146,149</sup>

### 1.4.3 PIKK family members

Of the PIKK family, ATR and DNA-PKcs are together with ATM involved in genome maintenance.

ATR (Mec1 in yeast) is in terms of sequence, and as later structural work showed, also in architecture most similar to ATM, as it also forms a dimer. Unless DNA-PKcs and ATM, ATR is essential in proliferating cells, where it is the apical kinase in response to DNA replication stress. Comparable to ATM it also phosphorylates hundreds of substrates.<sup>44</sup> ATR is activated in response to more forms of DNA damage than than DSBs, which mainly activate DNA-PK and ATM. ATR forms a complex with ATRIP, which is recruited to RPA-ssDNA.<sup>150</sup> Furthermore ATR is activated by DNA topoisomerase 2-binding protein 1 (TopBP1).<sup>151</sup> Another activator is ETAA1, which also contains an ATR-activating domain like TopBP1. Both activators might be relevant for different signalling pathways.<sup>42</sup>

DNA-PKcs is an enormous 460 kDa protein that also becomes activated in response to DSBs.<sup>152</sup> Together with Ku70/Ku80, which recruits the kinase to the DSBs, it forms DNA-PK and mediates and promotes the first steps in NHEJ, as has been described above. DNA-PK and NHEJ seem to be of high importance especially during lymphocyte maturation and V(D)J and class-switch maturation. Furthermore, mutations in DNA-PKcs are also associated with neurological defects, suggesting that NHEJ is important for DSB repair in neuronal progenitor cells. DNA-PK was the first PIKK of which a structure was solved, revealing that the HEAT repeat domain forms an enormous  $\alpha$ -solenoid cradle, on top of which the kinase domain is located.<sup>42,153</sup>

While the other PIKKs are not involved in DNA damage signalling, they play important roles in other aspects of the maintenance of cellular homeostasis.

TOR (mTOR in mammals) controls cell growth by promoting anabolic processes and suppressing catabolic processes. TOR was identified to be the protein to which the immunosuppressant rapamycin is binding. Yeast has two variants of the TOR protein, whereas higher eukaryotes have one mTOR protein, which is part of two different complexes, namely mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2), which regulate different cellular processes. Furthermore, many

other proteins associate with these complexes to modulate their activity, such as the GTPase Rheb, which activates mTORC1.<sup>148</sup> The preferred phosphorylation motif is a hydrophobic or aromatic residue followed by a serine or threonine.<sup>154</sup> A crystal structure of the human mTOR FATKIN domain provided the first high resolution insight into this class of kinases. It was revealed that the regulatory FAT domain cradles the kinase domain, revealing a way of allosteric control of the kinase. The active site itself was shown to be highly structured, reflecting an active state. The kinase is rendered inactive nevertheless by elements such as the PRD, and the proximity of the FRB, which block access to the active site.<sup>149</sup>

SMG1 is involved in NMD, a process in which transcripts with premature termination codons are degraded. It is only found in higher eukaryotes and some plant species, but not in fungi. In contrast to some other PIKKs, SMG1 phosphorylates only one substrate, UPF1, a conserved effector protein which activates the NMD pathway. SMG1 forms a complex together with SMG8/9 which inhibit its kinase activity.<sup>155</sup>

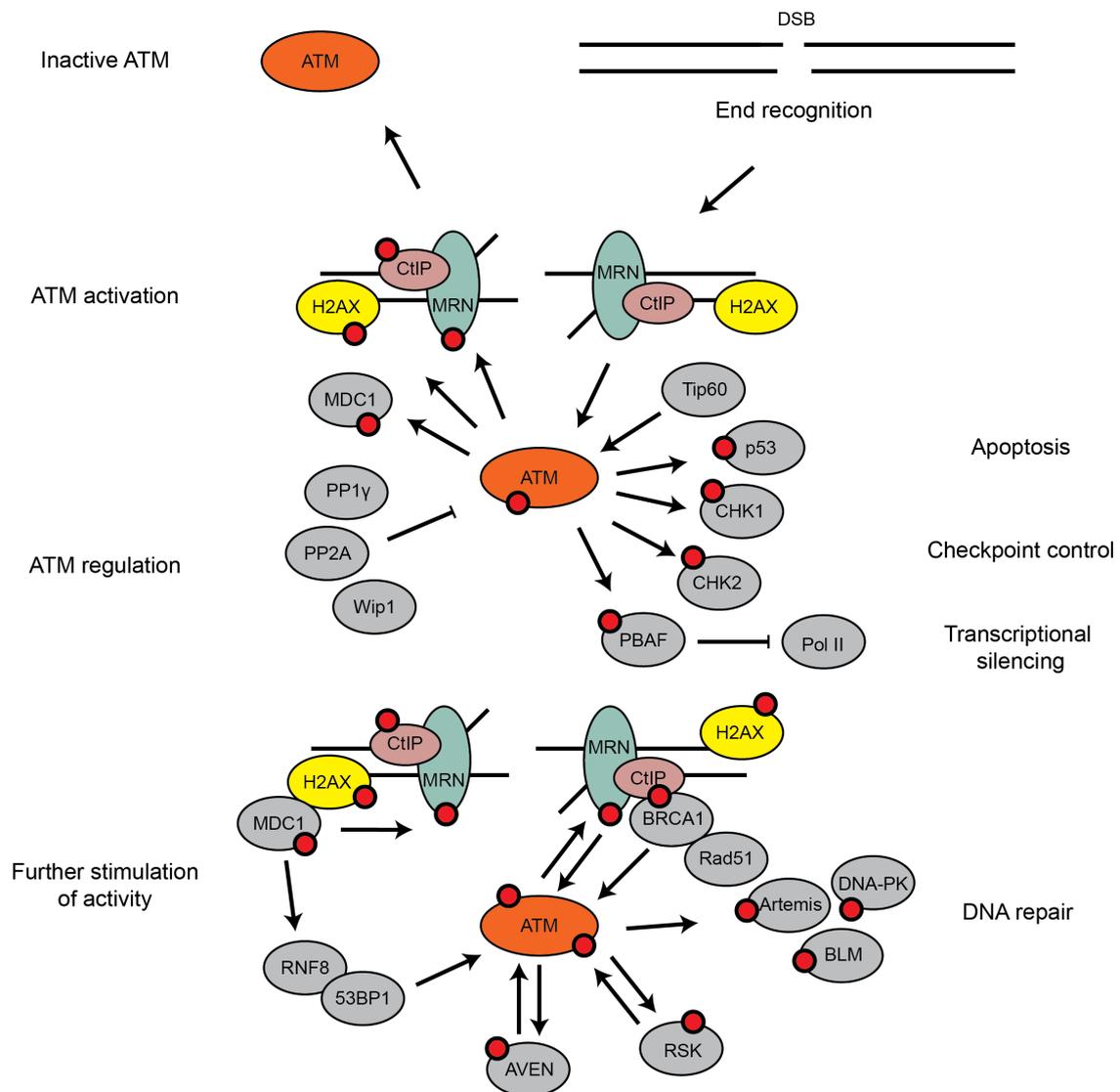
Tra1 (in yeast, TRAPP in higher eukaryotes) is the only pseudokinase of the PIKK family, as it lacks the conserved catalytic DFG motif in its active site. Tra1 forms a scaffolding subunit in two large complexes involved in the regulation of transcription. In yeast Tra1 is found in SAGA (Spt-Ada-Gcn5-Acetyltransferase) and NuA4 (nucleosome-Acetyltransferase-of-histone-H4). Furthermore Tra1 is also found in other chromatin-related complexes, such as the ASTRA complex in yeast and its human ortholog TRAPP in complexes such as Tip60. Both Tra1 and TRAPP are essential, reflecting the importance of the many complexes they are part of.<sup>140,156,157</sup>

#### 1.4.4 Cellular function of ATM

ATM is recruited to the site of the DSB by the MRN complex.<sup>45,135</sup> During this process, ATM becomes activated and phosphorylates hundreds of proteins with an S/TQ motif, including MRN itself and many other DNA repair factors.<sup>44,46,141,142</sup> ATM furthermore phosphorylates other kinases, such as CHK1 and 2. This way ATM orchestrates the DDR, which has a wide range of effects on the cell, such as cell cycle arrest, cell death, regulation of gene expression and protein degradation, as well as more localised effects such as modulation of the chromatin, mainly through phosphorylation of histone H2AX (which is denoted upon phosphorylation as  $\gamma$ H2AX) flanking the DSB to enable subsequent DNA repair (Figure 1.3).<sup>158,159</sup>

Despite its apical role in the damage response against DSBs, ATM is not essential, as cells in which is knock-out are still viable. As there is a high degree of crosstalk between the DNA damage signaling kinases, it is likely that ATR and DNA-PK can take over some of the signalling function of ATM. Mutations in ATM, especially truncations, lead to more severe phenotypes than knockouts.<sup>42,89</sup>

Although ATM is mainly recruited and activated by MRN in the first steps of HR, ATM has been found to be activated in NHEJ. In general, there is a significant degree of crosstalk between



**Figure 1.3: ATM activation and signalling.** In its resting state, ATM is present in the nucleus as an autoinhibited dimer. It is activated by the MRN complex upon the occurrence of a DSB. ATM is autophosphorylated in this process and reported to monomerise. Activated ATM phosphorylates hundreds of substrates, of which only a few are depicted here. Most importantly, the formation of  $\gamma$ H2AX leads to reorganisation of the chromatin and recruitment of more repair factors. ATM can also phosphorylate p53, leading to apoptosis, CHK1 and CHK2, which control cell cycle checkpoints. ATM-dependent phosphorylation of PBAF leads to transcriptional silencing by inhibiting Pol II. Furthermore, many factors, including ATM targets, regulate ATM activity, either through stimulating ATM more to amplify the signal, or by tempering its activity, e.g. through phosphatase activity. These concerted activities lead to a robust DNA damage response.

the different pathways and the signalling kinases. About 80 % of lesions repaired by NHEJ are independent of ATM, but a subset of lesions, mainly located in heterochromatin and with blocked ends, depends on ATM, ATR and DNA-PK. The mechanism may be that ATM, activated by

MRN, activates DNA-PK in these situations.<sup>160–162</sup> Another role of ATM in NHEJ could be DSB-end bridging.<sup>42,163</sup> ATM indeed phosphorylates many proteins that are involved in NHEJ, such as Artemis, PKN and DNA-PKcs itself.<sup>45</sup>

The main role of ATM, however, is in stimulating HR upon recruitment by MRN. An important step in this process is ATM-dependent phosphorylation of Nbs1, which recruits CtIP, which then initiates long-range resection. ATM-mediated phosphorylation of CtIP is also implied to play a role in removing Ku from single stranded DSBs to allow HR in the case of replication fork collapse.<sup>42,164,165</sup> Despite its stimulatory role in HR, ATM is not essential for HR progression. ATR may take over part of the DNA damage signalling during resection when the single stranded DNA is coated by RPA.<sup>42</sup>

One of the key targets of ATM phosphorylation is histone H2AX. The resulting  $\gamma$ H2AX is the foundation for a further chromatin-based signalling cascade. This mark is primarily read by MDC1, which is further stabilized on the chromatin by ATM. Further phosphorylation of MDC1 leads to further retention of MRN on the chromatin via Nbs1. This in turn could lead to more ATM recruitment, further  $\gamma$ H2AX formation, and subsequently more recruitment of MDC1, initiating an amplification loop to robustly signal DSBs.<sup>42,158,159</sup>

Another effect of ATM-dependent phosphorylation of MDC1 is the recruitment of the Ring Finger Protein 8 (RNF8) ubiquitin ligase, which leads to recruitment of 53BP1. As described above, this protein stimulates NHEJ.<sup>78</sup> Curiously, ATM also phosphorylates BRCA1 in response to DSBs, which antagonizes 53BP1, leading to HR. The choice for HR or NHEJ seems to depend on many factors, including the stage of the cell cycle, but the role of ATM herein is poorly understood.<sup>42</sup>

ATM-dependent phosphorylation of H2AX also leads to transcriptional silencing, ensuring that the transcription machinery does not collide with the DNA repair machinery. ATR and DNA-PKcs can also phosphorylate H2AX, but the exact effect of this depends on the cellular context. ATM appears to be mainly responsible for nucleolar DNA silencing.<sup>166</sup> Another way in which ATM induces transcriptional silencing is by phosphorylating the Polybromo-associated BAF complex (PBAF) SWItch/Sucrose Non-Fermentable (SWI/SNF) chromatin remodeler, which leads to silencing of RNA polymerase II.<sup>42,167</sup> Furthermore,  $\gamma$ H2AX has also been suggested to promote DNA repair by concentrating DNA repairs factor via phase separation.<sup>168,169</sup>

In order to ensure that the cell does not divide as long as damage is present, CHK1 and CHK2 are also phosphorylated by ATM, leading to cell cycle arrest until the lesion is repaired. If the damage is too severe for repair, ATM-dependent phosphorylation of p53 can lead to apoptosis.<sup>45</sup>

In addition to its role in DNA repair, ATM plays a role in the maintenance of telomeres, which also depends on activation by MRN, as has been described in previous sections. ATM and ATR both become activated when telomeres become too short, which leads to cell-cycle arrest and senescence.<sup>42</sup>

ATM also has been found to have roles in responding to transcriptional stress and regulating RNA metabolism. For example, the spliceosome has been found to be both a target and an effector for ATM in response to the formation of R-loops, although many details remain elusive.<sup>42,170</sup>

Some studies suggest that ATM can also be directly activated through oxidation by ROS, a mechanism which would not require MRN nor DNA.<sup>89,171</sup> ATM can indeed be stimulated *in vivo* by the addition of hydrogen peroxide. It is suggested that this ROS pathway is mainly relevant in neurons, in which ATM is more frequently located in the cytoplasm than in other cell types. This phenomenon might provide an explanation for the neurodegeneration observed in A-T.<sup>89</sup> Many questions remain to be answered, however, as ROS can also induce DSBs, which in turn are known to activate ATM.

ATM located in the cytoplasm has furthermore been reported to play a role in innate immunity in the detection of viral DNAs via MRN.<sup>172</sup> The presence of viral genomes have indeed been found to trigger a DDR too. As a self-defence mechanism, certain viruses have been described to encode proteins that inhibit ATM, enabling them to multiply in the their host cell.<sup>173–175</sup>

ATM thus appears to be involved in many aspects of cell physiology, but its main role is in the response to DNA damage. There is a significant degree of crosstalk between the different DSB repair pathways and the other kinases, ATR and DNA-PKcs, that mediate these. In addition, ATM may play a role in completely different pathways too, such as innate immunity and the response to ROS, although many aspects of the role of ATM there remain to be investigated.

#### 1.4.5 ATM activation

As ATM activation has large consequences for the cell, ATM is normally residing in the nucleus in an autoinhibited dimeric state. Autoinhibition is also a hallmark of the related PI3-kinases, which are typically bound to regulatory subunits. These interact with the kinase domain and active site via multiple interfaces, thereby physically blocking access to the active sites. This way, these lipid kinases remain inactive, until the blockage is lifted by an activating factor.<sup>89</sup> Earlier structures of PIKKs reveal similar mechanisms for autoinhibition, suggesting that ATM is probably also redundantly autoinhibited, requiring several strong cue to induce an active conformation.<sup>89,149,176</sup>

MRN activates ATM by recruiting the kinase through the C-terminus of Nbs1.<sup>60,88</sup> Mre11 and Rad50 have also been shown to interact directly with ATM and can activate ATM independently of Nbs1. Several studies have shown that MR in the ATP bound state can activate ATM, not through ATP-hydrolysis. As Rad50 hydrolyses ATP slowly, it has a longer ATP-bound state which could activate ATM.<sup>111,112,129,177,178</sup> Some studies suggested that ATM can also be activated in the absence of MRN and Ku.<sup>179</sup> This would suggest that ATM may also be activated by other factors. A plausible candidate would be DNA. Pull-down experiments with *Xenopus laevis* cell extracts have shown that ATM in the absence of MRN can be activated by an excess of DNA ends.<sup>180,181</sup>

ATM has been demonstrated to have DNA binding properties. It has a preference for long DNA strands (1 kb or larger) with blunt end or short overhangs.<sup>111,112,182,183</sup> Recent work on the activation of scTel1 showed that double stranded DNA and MRN synergistically activate ATM. ATM has a preference for long nucleosome free DNA, which appears to be in agreement with earlier studies, but there is no requirement for double stranded ends, suggesting that scTel1 binds to internal parts of the DNA.<sup>111</sup>

In addition to its affinity for DNA, the chromatin context itself is also important for ATM activation. Certain chromatin modifications indirectly stimulate ATM activation, leading to recruitment of factors that modify ATM. One of these factors, Tip60, acetylates ATM at the Lys3016 in the C-terminus, a modification which stimulates kinase activity. Furthermore, certain histone methylations affect the level and duration of ATM activation, and can additionally also lead to the recruitment of remodellers, which modify the chromatin in such a way that ATM can be fully activated. Binding of MDC1 to  $\gamma$ H2AX, a process which stimulates HR, indeed promotes persisting ATM activity as DSB sites, as described above. This activation may be distinct from the initial activation of ATM at the DSBs by MRN.<sup>89,184,185</sup>

Although more and more factors are found to activate ATM, it still poorly understood what changes take place in ATM to turn it into an active conformation. Early on it was reported that autophosphorylation of Ser1981, which is located in the FAT domain, is essential for ATM activation. This modification occurs in *trans* and leads to monomerisation.<sup>46</sup> However, this mechanism is controversial, as modification of the equivalent serine in mouse is not associated with activation.<sup>186,187</sup> Also, the monomerization is unlikely in light with recent structural work, including the structure described in this thesis, as the dimer interface is largely hydrophobic.<sup>188</sup> Although an intermediate resolution structure of an ATM monomer has been reported, it is not clear whether this represents an activated state.<sup>189</sup>

#### 1.4.6 Regulation of ATM activity

Although MRN activates ATM in the ATP-bound state, nuclease activity appears to require ATP hydrolysis. This raises the question what DNA substrate is presented to ATM *in vivo*, because the double-stranded DNA is resected. There seems to be a significant degree of cross-talk between ATM, ATR and DNA-PK. Upon resection, the ssDNA is coated by RPA, which stimulates ATR but not ATM. This may lead to changes in the signalling cascade, as ATR could take over from ATM. DNA-PK has been reported to phosphorylate ATM, which leads to the downregulation of its activity.<sup>42,190</sup>

MRN also becomes phosphorylated by ATM and other PIKKs, which has effects both on DNA repair and ATM signaling. Phosphorylation of Mre11 has been associated with a reduction of its affinity for DNA, also reducing further ATM recruitment to the DNA.<sup>191</sup> Phosphorylation of Rad50 and Nbs1 has been associated with checkpoint control through Smc1 and other factors.

Mechanistically the effects of phosphorylation of MRN are poorly understood and require further investigation.<sup>89</sup>

As (auto)phosphorylation is important for the regulation ATM activity, several phosphatases have been described to modulate ATM activity, although full details remain unclear. The protein phosphatase PP2A constitutively associates with ATM, but the association is lost upon DNA damage.<sup>192</sup> Similarly, PP1 $\gamma$ , which associates with ATM through Repo-Man, which is also lost upon DNA damage, leading to an increase of autophosphorylation.<sup>193</sup> Inhibiting both phosphatases indeed appears to lead to a near-complete DNA damage response as ATM becomes hyperphosphorylated. Another phosphatase, Wild-Type P53-Induced Phosphatase 1 (Wip1), regulates ATM activity after is induced via p53-dependent expression occurring upon DNA damage, suggesting a way of reducing the signalling cascade.<sup>89,194</sup>

Many other types of factors that modulate ATM activity have been described, and only a few will be described here as examples. The protein ATM interacting factor (ATMIN) binds to ATM with similar C-terminal motif as Nbs1, suggesting it might compete with Nbs1. ATMIN does not appear to play a role in the DDR, however, but may be of importance in other pathways, such as the cellular response against osmotic shock.<sup>89,195</sup> In situations in which MRN levels in the cell are low, other factors can compensate for the lack of MRN to activate ATM. ATM has indeed been reported to be directly stimulated by 53BP1 and BRCA1 *in vitro*, although the stimulatory effect appears to correlate with direct binding between MRN, ATM and both of these factors.<sup>196</sup> In mammalian cells, proteins of the homeodomain family of transcriptional regulators, NKX3.1 and homeobox B9, which are both found to mutated in certain forms of cancer, can bind directly to ATM, stimulating autophosphorylation and further phosphorylation of targets upon DNA damage.<sup>89,197,198</sup>

Some targets of ATM also directly stimulate ATM activity. For example, the apoptotic inhibitor Aven, which plays a role in controlling Gap 2 (G<sub>2</sub> phase)/Mitotic (M) checkpoint, is phosphorylated by ATM, but also binds to ATM upon DNA damage and stimulates its activity.<sup>199</sup> Ribosomal S6 kinase, which plays a role in controlling the same checkpoint, is also an ATM target and activates ATM further upon DNA damage. This provides a further way in which ATM is involved in regulating the cell cycle.<sup>89,200</sup>

It is currently unclear how such a large number of proteins can bind to and modulate ATM, as there is no common binding motif. The answer probably lies in the large size of the HEAT repeat domain of ATM, which could accommodate many modulators. This is indeed the case for e.g. Tra1, which functions as a hub within large complexes. As ATM activation has major consequences for the cell, it is not surprising that many modulators regulate ATM activity to fine-tune the signalling cascade.<sup>89</sup>

### 1.4.7 Aim of the project

ATM is an apical signalling kinase in the response to DSBs. Its importance in genome maintenance is reflected in the fact that it is frequently mutated in cancer, and also in the variety of severe symptoms that characterise A-T. When ATM is activated by MRN, it phosphorylates hundreds of substrates, including other kinases, leading to the recruitment of many DNA repair factors, cell cycle arrest and even cell death, if the damage is beyond repair. Furthermore, emerging roles of ATM in completely different areas have been identified, such as innate immunity and redox signalling.

However, many key questions remain. Due to the large size of ATM, it is not possible to express it recombinantly in quantities required for biochemistry and crystallography. Therefore, structural data and insights in its *in vitro* biochemistry have remained elusive. It is for example not understood how MRN activates ATM, and especially not what structural rearrangements need to take place within ATM in order to turn it into an active conformation. Furthermore, the binding to MRN and any potential DNA binding properties were not characterized in high detail.

The aims of this thesis, therefore, were to determine a structure of ctTel1, an ATM ortholog from the thermophilic fungus *Chaetomium thermophilum* and to characterise its binding properties to DNA and MRN, in order to gain more insight in its activation mechanism. *C. thermophilum* has become popular amongst structural biologists in recent years as its proteins are more stable, making them well suited for structural studies.

Although it was not possible to express full-length ctTel1 recombinantly, a truncated construct, consisting of the HEAT-repeat domain, was identified which could be expressed in large quantities in *Escherichia coli*. This construct was then used to probe the binding properties towards MRN and DNA. Furthermore, nanobodies were raised against this protein, enabling the purification of the full-length endogenous ctTel1 from *C. thermophilum*. Although the amounts were small, they proved to be sufficient for high-resolution structure determination using cryo-EM. These two approaches combined give more insight in the activation mechanism of this highly important kinase.

# Chapter 2

## Materials and Methods

### 2.1 Materials

All chemicals used in this study were obtained from Carl Roth (Karlsruhe), Merck (Darmstadt) or Sigma Aldrich (Deisenhofen) unless stated otherwise. Enzymes for molecular biology were purchased from Fermentas (St. Leonrot) or New England Biolabs (Frankfurt). Desalted oligonucleotides for molecular cloning and HPLC purified oligonucleotides for biochemical assays were obtained from Metabion (Planegg). Media for insect cell culture was obtained from GIBCO/Invitrogen. Chromatography columns and materials were obtained from GE Healthcare (Penzberg) unless indicated otherwise.

Sterilised glassware and laboratory materials were used for all experiments.

#### 2.1.1 Oligonucleotides

**Table 2.1: DNA oligonucleotides used DNA binding assays.** Bp: base pair, Fw: forward, rv: reverse, 6-FAM: 6-carboxyfluorescein.

Oligo	Sequence (5'-3')
60 bp fw	6-FAM-AAGGGAACATCGGGTAGTTGAACAGCATAGGCCCGGTCTTGCCTTAACCTCTAATTCTCG
60 bp rv	CGAGAATTAGAGGTTAAGGCAAGACCGGGCCTATGCTGTTCAACTACCCGATGTTCCCTT
30 bp fw	6-FAM-ACCGTAGGTTTACTGTACGATATCAGTGC
30 bp rv	GCACTGATATCGTGACAGTAAACCTACGGT
25 bp fw	6-FAM-GAGATCCGTATGCCTACTGACATTG
25 bp rv	CAATGTCAGTAGGCATACGGATCTC
20 bp fw	6-FAM-CCTATATGAAGCTGAGTGCC
20 bp rv	GGCACTCAGCTTCATATAGG

**Table 2.2: DNA oligonucleotides used for cloning.** Fw: forward, rv: reverse, his: His<sub>6</sub>-tag, KD: kinase dead, FLAG: triple FLAG-tag, ctTel1-#: number of the DNA fraction for Gibson assembly, G<sub>4</sub>S: protein linker consisting out of four glycines and one serine.

Amplified gene	Sequence (5'-3' direction)	Cloning method
ctTel1(1-1864)-his fw	GCGGCCGCACTCGAGCACCAC	Blunt end ligation
ctTel1(1-1864)-his rv	GGCTCGGTCCGGCAATCGAGGATTC	Blunt end ligation
MBP-ctTel1(1-1864) fw	AAGGAGATATACATAATGAAACATCACCATCACCATCACC	InFusion
MBP-ctTel1(1-1864) rv	TTGACCCGCGGCATAGCGCCCTGAAAATAAAGATTCTCG	InFusion
FLAG-ctTel1(KD) for	TTGACCCGCGGCATAGCGCCCTGAAAATAAAGATTCTCG	QuikChange
FLAG-ctTel1(KD) rev	CCAGCAGGATTTTGTGGCCGTGACGCGCACCCAGTCCCA	QuikChange
FLAG-ctTel1-1 fw	GCTGGGCGGCATACGCATAAAGAGCCTGCAGTCTCGACA	Gibson Assembly
FLAG-ctTel1-1 rv	ACGTCCAGCCAGTGTGCTCTGCTTCCAGATTTATCGTCCG	Gibson Assembly
FLAG-ctTel1-2 fw	ACGACGATAAATCTGGAAGCAGAGCACACTGGCTGGAC	Gibson Assembly
FLAG-ctTel1-2 rv	TTGTGCGAGACTGCAGGCTCTTTATGCGTATGCCGCCAG	Gibson Assembly
FLAG-ctTel1-3 fw	GAGAATCCTCGATTGCGGACAGAGCACACTGGCTGGAC	Gibson Assembly
FLAG-ctTel1-3 rv	CCTGAGAGACCCGCTAGTTGGCTTCCAGATTTATCGTCGTCGT	Gibson Assembly
FLAG-ctTel1-4 fw	ACGACGATAAATCTGGAAGCCAACCTAGCGGGTCTCTCAGGC	Gibson Assembly
FLAG-ctTel1-4 rv	ACGTCCAGCCAGTGTGCTCTGTCCGCAATCGAGGATTCTCC	Gibson Assembly
FLAG-ctTel1-5 fw	TCCCAGAGCCCATCATCAACCACTAGCGGGTCTCTCAG	Gibson Assembly
FLAG-ctTel1-5 rv	TAACTCTTGACCCGCGGCATGCTTCCAGATTTATCGTCGTCGT	Gibson Assembly
FLAG-ctTel1-6 fw	ACGACGATAAATCTGGAAGCATGCCGCGGGTCAAGAG	Gibson Assembly
FLAG-ctTel1-6 rv	CCTGAGAGACCCGCTAGTTGGTTGATGATGGGCTCTGGGAT	Gibson Assembly
Nanobody-G <sub>4</sub> S-FLAG fw	ACAAGGATGACGATGACAAGGCGGCCGCACTCG	Gibson Assembly
Nanobody-G <sub>4</sub> S-FLAG rv	TGGTGTCTCGAGTGCGGCCGCTTGTTCATCGTC	Gibson Assembly

### 2.1.2 Plasmids

**Table 2.3: Plasmids for recombinant protein expression in *E. coli* or in insect cells.** Labels and abbreviations as in table 2.1.

Construct	Encoded Sequence	Vector
ctTel1(1-1864)-his	ctTel1 aa 1-1864	pET28
MBP-ctTel1(1-1864)-his	ctTel1 aa 1-1864	pET28
FLAG-ctTel1	ctTel1	pACEbac
FLAG-ctTel1(KD)	ctTel1 D2760A and N2765K	pACEbac
D4-FLAG-his	anti-ctTel1 nanobody D4	pMJS187
F7-FLAG-his	anti-ctTel1 nanobody F7	pMJS187
D4-G <sub>4</sub> S-PreScission-G <sub>4</sub> S-FLAG-his	anti-ctTel1 nanobody D4	pMJS187
F7-G <sub>4</sub> S-PreScission-G <sub>4</sub> S-FLAG-his	anti-ctTel1 nanobody F7	pMJS187

### 2.1.3 Strains

**Table 2.4: Strains used for molecular cloning, protein expression and pull-downs.** *E. coli* XL1Blue was used for molecular cloning, *E. coli* DH10 for the generation of bacmids for insect cell expression and *E. coli* Rosetta for recombinant protein expression. *C. thermophilum* was used for endogenous ctTel1 pull-downs. Sf21 and High five insect cells were used for virus production and recombinant protein expression, respectively.

Organism	Strain	Source
<i>C. thermophilum</i>	var. DSM 1495	DSMZ and Prof. Dr. Hurt, University of Heidelberg
<i>E. coli</i>	XL1Blue	Stratagene, Heidelberg
<i>E. coli</i>	Rosetta(DE3)	Novagen, Madison, USA
<i>E. coli</i>	DH10 MultiBac	Redbiotech, Schlieren, Switzerland
<i>S. frugiperda</i>	Sf21	Invitrogen, Karlsruhe
<i>T. ni</i>	HighFive	Invitrogen, Karlsruhe

### 2.1.4 Antibodies

**Table 2.5: Antibodies for Western blots.** Labels and abbreviations as in table 2.1.

Antibody	Primary or Secondary	Origin	Dilution	Source
anti-ctTel1	Primary	Mouse	1:500	Prof. Kremmer Lab, LMU Munich
anti-FLAG	Primary	Rabbit	1:1000	Sigma Aldrich, Munich
anti-mouse IgG2a-HRP	Secondary	Sheep	1:1000	GE Healthcare, Penzberg
anti-rabbit-IgG2	Secondary	Goat	1:1000	Biorad, Munich

### 2.1.5 Common buffers, media and antibiotics

*Escherichia coli* cultures were grown in liquid Lysogeny Broth (LB) medium (1% (w/v) bacto-tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) or on LB agar plates (LB + 1.5% (w/v) agar). Antibiotics were added to the following final concentrations: 100 µg/mL ampicillin, 50 µg/mL kanamycin, 34 µg/mL chloramphenicol.

*E. coli* DH10 MultiBac cultures were grown in liquid LB medium or on LB plates containing 0.1 mM IPTG, 50 µg/mL kanamycin, 10 µg/mL tetracycline, 10 µg/mL X-Gal and 7 µg/mL gentamycin.

Sf21 cells were cultured in Sf-900 III serum-free medium (GIBCO/Invitrogen, Karlsruhe) with 10 µg/mL gentamycin. HighFive cells were grown in ExpressFive serum-free medium (prepared according to instruction from GIBCO/Invitrogen, Karlsruhe) with 10 µg/mL gentamycin.

*C. thermophilum* was cultured in liquid Chaetomium Culture Medium (CCM), composed of 3 g sucrose, 0.5 g NaCl, 0.65 g K<sub>2</sub>HPO<sub>4</sub> · 3 H<sub>2</sub>O, 0.5 g MgSO<sub>4</sub> · 7 H<sub>2</sub>O, 0.01 g Fe(III)sulfate-hydrate, 5 g tryptone, 1 g peptone, 1 g yeast extract and 15 g potato dextrin per L, pH 7.0.

**Table 2.6: Common buffers.**

Buffer	Composition
4x Laemmli buffer	0.11 M TRIS pH 6.8, 16% (v/v) ethanol, 4% (w/v) SDS 5% (v/v) $\beta$ -mercaptoethanol, 0.05% bromophenol blue
Coomassie stain	50% (v/v) ethanol, 7% (v/v) acetic acid, 0.2% (w/v) Coomassie Brilliant Blue R250
2x TG buffer	50 mM TRIS, 384 mM glycine, 20% (v/v) ethanol, 0.05% SDS
10x TBS	200 mM TRIS-HCl pH 7.4, 1.37 M NaCl, 26 mM KCl
1x TBST	1x TBS, 0.1% TWEEN-20
20x ECL	2 M TRIS pH 8.5
1x ECL (10 mL)	10 mL 20x ECL solution, 3 $\mu$ L H <sub>2</sub> O <sub>2</sub> , 25 $\mu$ L cumaric acid (90 mM), 50 $\mu$ L luminol (250 mM)

## 2.1.6 Software and algorithms

**Table 2.7: Software.**

Software	Source
COOT	Emsley et al., 2010
CTFFind 4.1	Rohou and Grigorieff, 2015
EPU	FEI
JaView	Waterhouse et al., 2009
MAFFT	Katoh et al., 2005
MotionCor2	Zheng et al., 2017
PHENIX	Adams et al., 2010
Prism	GraphPad
PyMOL	Schrödinger LLC
Relion 3.0	Zivanov et al., 2018
Sequence Manipulation Suite	Stothard, 2000
SnapGene	GSL Biotech LLC
SWISS-MODEL server	Waterhouse et al., 2018
UCSF Chimera	Pettersen et al., 2004
UCSF ChimeraX	Goddard et al., 2017

## 2.2 Methods

### 2.2.1 Molecular Biology Methods

#### 2.2.1.1 Molecular cloning

Genes were amplified using the polymerase chain reaction technique (PCR) using a self-made Phusion High-Fidelity PCR Master Mix (ThermoFisher Scientific). For a typical reaction 10-50 ng

template DNA was used and 0.5  $\mu$ M (final concentration) oligonucleotide primer, to a total volume of 20  $\mu$ L. The standard PCR protocol is shown in table 2.8.

Bands were separated using agarose gel electrophoresis and were visualised using UV light. Bands of the correct size were cut out and purified using the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren). PCR products and plasmid DNA were assembled using Gibson Assembly (New England Biolabs) or via blunt end ligation according to the manufacturers instructions. For blunt end ligation, the primers were phosphorylated using PNK prior to their use for PCR. The reactions were transformed into chemically competent *E. coli*.

The TEV cleavage site from the original ctTel1(1-1864)-TEV-his construct was removed to clone ctTel1(1-1864)-his using blunt end ligation. Nanobodies with a G<sub>4</sub>S-PreScission-G<sub>4</sub>S-FLAG-his tag were cloned into a pMJS187 vector using Gibson assembly. The codon-optimised kinase domain and the catalytically-dead mutant (GenScript, Hong Kong), with N-terminal FLAG-tags, and fragments of ctTel1(1-1864)-his were cloned sequentially into a pACEbac vector using Gibson assembly in order to obtain full-length FLAG-ctTel1.

The integrity of all constructs was confirmed by DNA sequencing (GATC/Eurofins, Ebersberg).

**Table 2.8: PCR program used for molecular cloning.** The annealing temperatures X depended on the temperatures calculated in SnapGene. The extension time was adjusted according to the length of the gene or vector.

Cycle step	Temperature	Time	Cycles
Initial denaturation	98 °C	2 min	1x
Denaturation	98 °C	30 sec	30x
Annealing	X °C	30 sec	
Extension	72 °C	30 sec/kb	
Final extension	72 °C	10 min	1x
Cooling	10 °C	$\infty$	

### 2.2.1.2 Site-directed mutagenesis

To generate kinase-dead full-length ctTel1, site-directed mutagenesis using QuikChange (Agilent, Oberhaching) was performed using the Phusion kit. Primers for site-directed mutagenesis were designed using SnapGene. A typical Quikchange reaction contained 100-200 ng of template and 25 nM (final concentration) of each DNA oligonucleotide. The standard QuikChange PCR program is shown in table 2.9.

After DpnI digestion, 5  $\mu$ L of the reaction was transformed into 50  $\mu$ L chemically competent XL1Blue *E. coli* cells. The presence of the desired point mutation and the integrity of the gene was confirmed using DNA sequencing (GATC/Eurofins, Ebersberg).

**Table 2.9: PCR program used for site-directed mutagenesis.** The annealing temperatures were adjusted depending on the reaction. The extension time was adjusted according to the length of the gene.

Cycle step	Temperature	Time	Cycles
Initial denaturation	98 °C	30 sec	1x
Denaturation	98 °C	10 sec	30x
Annealing	65 - 72 °C	30 sec	
Extension	72 °C	30 sec/kb	
Final extension	72 °C	10 min	1x
Cooling	10 °C	∞	

## 2.2.2 Microbiology methods: *E. coli*

### 2.2.2.1 Generation of competent *E. coli*

To prepare chemically competent cells, an overnight culture of *E. coli* was prepared and used to inoculate 200 mL of LB medium with the appropriate antibiotics and grown until an OD<sub>600</sub> of 0.3-0.5. The cells were centrifuged at 3000g for 5 min at 4 °C and resuspended in chilled TSS buffer, aliquoted in 100 µL, flash-frozen in liquid nitrogen and stored at -80 °C.

### 2.2.2.2 Transformation into *E. coli*

For transformation, 10-100 ng of plasmid DNA was added to 100 µL of competent *E. coli* and incubated on ice for 30 min. The cells were heat shocked for 45 sec at 42 °C and incubated on ice for 5 min. Then 900 µL of LB was added and the cells were incubated at 37 °C for 30-60 min. Afterwards, the cells were centrifuged at 3000g for 5 min and 800 µL of the supernatant was removed. The cells were then resuspended before plating out on LB agar plates supplemented with the correct antibiotics.

*E. coli* DH10Multibac cells were thawed on ice and 1-2 µL of a few 100 ng of pACEbac vector were added and incubated on ice for 15-30 min and heat shocked for 45 sec at 42 °C, and rested on ice for 2 min. 900 µL LB was added and the cells were incubated at 37 °C overnight. The culture was centrifuged at 3000 g for 5 min and plated onto LB agar bacmid plates supplemented with the appropriate antibiotics, X-gal and IPTG. The plates were incubated for minimally 24 h until blue and white colonies could be distinguished.

### 2.2.2.3 Isolation of plasmid DNA

Single colonies were picked and used to inoculate 5 mL LB medium with the corresponding antibiotic. Cells were shaken overnight at 37 °C 200 rpm and harvested by centrifugation 3000 g for 5 min.

For cloning purposes and for expression in *E. coli*, plasmid DNA was isolated from a 5 mL overnight *E. coli* XL1Blue culture using Nucleospin Plasmid EasyPure kit (Macherey-Nagel, Düren) following the instructions of the manufacturer. Nucleic acid concentrations were determined using a NanoDrop (NanoDrop™ 1000 Spectrophotometer, ThermoScientific).

### 2.2.2.4 Purification of MultiBacmid

Several of the white colonies were restreaked to confirm they were the correct colonies. Then these were used to seed LB cultures for a midiprep. For the isolation of bacmid DNA for virus generation and protein expression in insect cells, the NucleoBond Xtra Midi prep (Macherey-Nagel, Düren) was used according to the instructions.

## 2.2.3 Microbiology methods: *C. thermophilum*

Culturing of *C. thermophilum* was done according to instructions from the Hurt laboratory (University of Heidelberg).<sup>201</sup>

### 2.2.3.1 Culturing *C. thermophilum* on agar plates

*C. thermophilum* was either grown from freeze-dried mycelium, from dried spores and from fresh mycelium. To start a new culture on plates, freeze-dried mycelium was cut into pieces and put onto the agar plate and then hydrated with sterile CCM medium. Alternatively, dried spores were plated out the same way. To plate out frozen spores, they were thawed and directly pipetted onto the agar plate. To propagate plates with mycelia, pieces were cut out and used to inoculate fresh CCM agar plates. In all cases, the plate was incubated at 55 °C in a sealed plastic bag supplemented with a wet tissue to avoid dehydration. The plates were incubated until the plate was fully covered with mycelium (48-72 h). Plates with mycelia were stored at room temperature until further use.

### 2.2.3.2 Culturing *C. thermophilum* in liquid medium

Two CCM agar plates with mycelium were cut into pieces and used to inoculate a preculture of 500 mL of liquid CCM medium in a 1 L baffled Erlenmeyer flask supplied with ampicillin and

incubated in a rotary shaker at 90 rpm for 24 h at 55 °C. The preculture was transferred into a kitchen blender and shredded until a blended mycelium broth was obtained. The broth was used to inoculate 6 x 2 L CCM media in 5 L baffled Erlenmeyer flasks, which were grown overnight at 55 °C at 90 rpm. After harvesting the mycelium (see below) the flasks were returned to the incubator to obtain a second batch of mycelium.

### **2.2.3.3 Harvesting and lysis of mycelium**

The cultures were strained through a metal sieve with a very fine pore size. The mycelium was transferred to vacuum filter to remove residual medium. The mycelium was divided into smaller pieces and flash frozen in liquid nitrogen. The frozen mycelia were stored at -80 °C until lysis.

The frozen mycelia were lysed by means of cryo-milling (SPEX 6970EFM). The milling cycle consisted of 5 min of precooling, 3 min grinding at a rate of 9 and 3 min of cooling down. The lysed mycelia were stored at -80 °C.

### **2.2.3.4 Generation of spores**

*C. thermophilum* spores were generated by culturing mycelium on rice agar. 75 g rice was cooked for 2 h in 1 L water. The obtained rice broth was filtered through a sieve to remove the rice seeds and filled up to 1 L with water. 15 g of agar was added prior to autoclaving. 50 mL rice agar was filled in sterile glass beakers. 2 mL of liquid mycelium culture (cultured as described above) was added to the agar and incubated at 37 °C for minimally one week until black spores appeared on the surface of the agar. The spores were harvested in 1 M sterile sorbitol solution by scraping of the spores with a sterile inoculating loop. The spore slurry was filtered through a sterile gauze to remove remaining agar or mycelium. The filtered spore solution was aliquoted and stored at -20 °C.

## **2.2.4 Insect cell methods**

### **2.2.4.1 Production of baculovirus in Sf21 cells**

For transfection, 200 µL medium was mixed with 3 µL FuGene transfection reagent (Promega, Wallheim) and 1-2 µg bacmid and then incubated at room temperature for 45 min. Then 2 mL of  $0.4 \times 10^6$ /mL Sf21 cells was incubated in a 6-well culture plate at 27.5 °C for 30 min. The transfection mixture was added to the cells after these had attached to the bottom of the plate and incubated for 72-84 h at 27.5 °C. The resulting supernatant contained the P0 virus.

A suspension of 10 mL  $1 \times 10^6$ /mL Sf21 cells was prepared in a 100 mL Erlenmeyer. To this, 1 mL of the supernatant containing the P0 baculovirus was added. Cells were incubated for 3-4 days at

27.5 °C and 95 rpm, until signs of transfection, such as changes in viability and morphology, could be observed.

In case of successful transfection, cells were pelleted (10 min 3000 g) and 0.5 mL of the supernatant containing P1 virus was added to 50 mL Sf21 cells  $0.4 \times 10^6$ /mL in a 500 mL Erlenmeyer. This was incubated for 72 h at 27.5 °C while shaking at 95 rpm. To obtain the P2 virus, cells were centrifuged at 3000 g for 10 min and the virus-containing supernatant was filtered through a 0.22 µm filter and stored at 4 °C.

## **2.2.5 Protein biochemistry methods**

### **2.2.5.1 Protein expression in *E. coli***

For large scale protein expression, 100 mL overnight culture of LB medium with the right selection antibiotics was prepared by streaking colonies of a plate with *E. coli* Rosetta (DE3). The culture was grown overnight at 37 °C. A 3 L LB medium culture with the appropriate antibiotics was inoculated 1:100 with the overnight culture. Cells were grown until an  $OD_{600} = 0.6-0.8$  and expression was induced with 0.5 mM IPTG. The temperature was lowered to 18 °C and cells were grown overnight. Cells were harvested the next day by centrifugation at 3000 g for 10 min and either frozen in liquid nitrogen and stored until further use, or used immediately for purification.

### **2.2.5.2 Test expression ctTel1 in HighFive insect cells**

For test expressions, 50 mL media with  $1 \times 10^6$ /mL in 250 mL flasks were used and baculovirus P2 titrated in different ratios. Cells were incubated at 27.5 °C shaking 95 rpm for 1-4 days and then harvested. Every 24 h the cell count was monitored and a sample was harvested for analysis of the expression levels of ctTel1.

### **2.2.5.3 Purification of ctTel1(1-1864)**

Cells were resuspended in Lysis buffer (50 mM TRIS-HCl pH 8.0, 300 mM NaCl, 5 % (v/v) glycerol, 30 mM imidazole, 1 mM β-mercaptoethanol) supplemented with protease inhibitor tablets. Cells were lysed by sonication on ice (Branson Sonifier 250). Lysate was cleared by centrifugation at 30.600 g 4 °C for 45 min (rotor SS34) and loaded onto an immobilised metal affinity chromatography column (HisTrap), which was preequilibrated in Lysis buffer. After loading the column was washed with lysis buffer. Protein was eluted using Elution buffer (50 mM TRIS-HCl pH 8.0, 50 mM NaCl, 5 % (v/v) glycerol, 500 mM imidazole, 1 mM β-mercaptoethanol). The purity of the fractions was analysed using SDS-PAGE. The best samples were pooled and loaded on a HiTrap Q anion exchange column which was preequilibrated in QA buffer (25 mM TRIS-HCl pH 8.0, 100 mM NaCl, 5 % (v/v)

glycerol, 2 mM imidazole, 1 mM  $\beta$ -mercaptoethanol). The column was washed with three column volumes of QA buffer prior to elution using a gradient over 5 column volumes with buffer QB (25 mM TRIS-HCl pH 8.0, 1 M NaCl, 5 % (v/v) glycerol, 2 mM imidazole, 1 mM  $\beta$ -mercaptoethanol). The eluted fractions were analysed using SDS-PAGE gel. The purest fractions were pooled and spin-concentrated. The protein was further purified using size exclusion chromatography with a HiLoad Superdex 200 16/60 column, which was preequilibrated in SEC buffer (25 mM HEPES NaOH pH 8.0, 300 mM NaCl, 2 mM imidazole, 2 mM TCEP). The purest fractions were spin-concentrated and either using directly or flash-frozen in liquid nitrogen and stored at -80 °C.

#### **2.2.5.4 Purification of nanobodies against ctTel1**

Nanobodies were generated against ctTel1(1-1864) by C. Linke-Winnebeck, in collaboration with ChromoTek GmbH, Martinsried.

Nanobodies were expressed in *E. coli* Rosetta from a pMJS187 vector, which is optimised for expression of disulfide-containing proteins.<sup>202-204</sup> The cell pellet was resuspended in Buffer A (50 mM TRIS-HCl pH 8.0, 300 mM NaCl, 20 mM imidazole) supplemented with protease inhibitor tablets. Cells were lysed by sonication 15 min at 30 % output. The lysate was cleared by centrifugation at 30.600 g 4 °C for 45 min. The supernatant was incubated for 30 min with Ni<sup>2+</sup>-NTA agarose (IBA Life Sciences), preequilibrated in Buffer A. After incubation, the beads were transferred to a gravity chromatography column (Biorad) and were washed with Buffer A. Protein was eluted using Buffer B (50 mM TRIS-HCl pH 8.0, 300 mM NaCl, 500 mM imidazole). The lysate, supernatant, wash fraction and the eluted fractions were analysed on an SDS-PAGE gel. The purest fractions were pooled and dialysed overnight into Buffer C (25 mM TRIS-HCl pH 8.0, 150 mM NaCl). The protein was then concentrated using an Amicon spinconcentrator with a 5 or 13 kDa cut-off, depending on the exact construct. The concentrated protein was purified further using size exclusion chromatography using a HiLoad 75 16/600 column. The eluted fractions were analysed using SDS-PAGE and the purest fractions were pooled, concentrated, flash frozen in liquid nitrogen and stored at -80 °C until further use.

#### **2.2.5.5 Test expression of recombinant ctTel1**

Codon-optimised FLAG-ctTel1 and FLAG-ctTel1(KD) were expressed in HighFive cells infected with different baculovirus ratios. Cells were harvested after three days by centrifuging at 3000 g for 10 min. The cells were flash frozen in liquid nitrogen and then thawed on ice. Cells were resuspended in Lysis buffer (50 mM HEPES NaOH pH 8.0, 250 mM NaCl, 10 % (v/v) glycerol, 0.1 % (w/v) CHAPS) supplemented with protease inhibitor tablets. The cells were lysed on ice using a glass homogenizer. Additionally, the cells were sonicated for 30 sec at 20 % output. The lysate was cleared by centrifugation at 30.600 g 45 min at 4 °C. The supernatant was filtered with a 0.45  $\mu$ m syringe filter. The supernatant was incubated with 150  $\mu$ L FLAG-beads (preequilibrated in Lysis

buffer) for 60 min in the cold room under rotation. The beads were washed with Lysis buffer and boiled. The fractions were analysed using SDS-PAGE.

#### **2.2.5.6 Purification of endogenous ctTell**

Endogenous ctTell was purified from *C. thermophilum* var DSM1495 using a nanobody-mediated pull-down. 100 g of *C. thermophilum* powder was resuspended in Resuspension Buffer (25 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 5 % (v/v) glycerol, 0,1 % (w/v) CHAPS) supplemented with protease inhibitor tablets. The lysate was cleared by centrifugation at 30.600 g for 45 min at 4 °C.

M2 FLAG affinity resin (150 µL, Sigma Aldrich) was preequilibrated in Resuspension Buffer and incubated with a nanobody for 1 h at 10 °C on a rotating wheel. Excess nanobody was washed off 4 x 1 mL lysis buffer. Nanobody coated beads were incubated with the cleared lysate overnight at 10 °C under rotation in an insect cell harvesting vessel.

The beads were harvested by gentle centrifugation at 4 °C 5 min and washed using 4 times 15 mL washing buffer (25 mM HEPES-NaOH pH 7.5, 150 mM NaCl). The nanobody-ctTell complex was eluted by proteolytic cleavage using PreScission protease. ATP $\gamma$ S and MgCl<sub>2</sub> were added to the ctTell eluate to a final concentration of 0.1 mM and 1 mM, respectively.

Concentrations of endogenous ctTell were estimated on SDS-PAGE gel, by comparing the size of the ctTell band to bands from a dilution series of a similarly-sized protein, of which the concentration can be accurately determined.

#### **2.2.5.7 Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)**

The purity of protein samples was assessed by means of denaturing polyacrylamide gel electrophoresis (sodium dodecyl sulfate polyacrylamide gel electrophoresis, SDS-PAGE) using 4-20 % precast Bis-Tris Gels and RunBlue running buffer (Expedeon, Heidelberg).

Protein samples were mixed with 4x Laemmli buffer and denatured for 5 min at 95 °C. Cell lysate samples were diluted further in water before adding Laemmli buffer. Proteins were separated by running the gels at 120 V in 1x RunBlue SDS buffer. The gel was stained either by incubated the gel for 30 min with Coomassie Brilliant Blue staining solution followed by destaining in hot water, or by staining with InstantBlue and destaining with water. Alternatively, the gel was used for Western blot.

### **2.2.5.8 Western blot analysis**

The sample of interest was separated by size using an SDS-PAGE gel run at 100 V and transferred to a PVDF membrane using wet transfer with a Mini-Trans Blot electrophoretic transfer cell (Bio-Rad). The PVDF membrane was active by soaking it for 5 min in 100 % (v/v) methanol. All other components were pre-soaked in cold 2x TG buffer (50 mM TRIS, 384 mM glycine, 20 % (v/v) methanol, 0,05 % (w/v) SDS) for 30 min. The gel was sandwiched with the PVDF membrane between a foam pad and two Whatman papers on both sides and clamped tightly prior to transfer to the blotting chamber. The transfer was run for 2 h at 250 mA in the cold room/write actual temperature.

After the transfer, the membrane was blocked by incubation in 50 mL TBST + 5 % (w/v) milk powder for 1 h at room temperature. The membrane was then incubated with the primary antibody mAB anti-ctTel1 8D9 (Kremmer lab, work by C. Linke-Winnebeck) diluted 1:500 in TBST + 1 % (w/v) milk powder overnight in the cold room on a rocker. The membrane was then washed using 4x 50 mL TBST for 15 min at room temperature. The membrane was then incubated with the secondary antibody mouse anti rat Ig2a-HRP horse radish peroxidase conjugated diluted 1:1000 in 5 mL TBST + 1 % (w/v) milk powder for 1 h at room temperature. The membrane was then washed 4x for 15 min in TBST. To detect ctTel1, the membrane was incubated in 10 mL ECL solution for 1 minute. The chemiluminescence was directly measured using an Amersham™ Imager 600 (GE Healthcare, Penzberg). Alternatively, an antibody against the FLAG-tag was used (see table 2.5).

## **2.2.6 Assays**

### **2.2.6.1 Annealing of DNA**

Oligonucleotides of various lengths were ordered commercially (20, 25, 30 and 60 bp). Labelled oligonucleotides were mixed with a 1.1-fold molar excess of the unlabelled oligonucleotide in annealing buffer (25 mM TRIS pH 7.5, 50 mM NaCl and 10 mM MgCl<sub>2</sub>). The mixture was incubated at 95 °C for 5 min in a thermocycler, and cooled down gradually to 20 °C at a cooling rate of 0.1 °C/sec.

### **2.2.6.2 Affinity measurement by fluorescence anisotropy with DNA**

CtTel1(1-1864) was diluted in Assay Buffer (25 mM HEPES-NaOH pH 8.0, 100 mM NaCl and 5 mM β-mercaptoethanol) and mixed with an equal volume of 6-FAM 5'-labelled oligonucleotide (5 nM final concentration). After incubation for 30 min at 25 °C, the fluorescence anisotropy was measured at an excitation wavelength of 470 nm and an emission wavelength of 520 nm, using a plate reader (Tecan). The results were analysed using Prism (GraphPad) and fit using an equation assuming one specific binding site. Experiments were performed in triplicates.

### 2.2.6.3 Affinity measurement by fluorescence anisotropy with FAM-Nbs1 peptide

FAM-Nbs1 peptide (2643.9 Da, aa 939-954, Genscript, Hong Kong) was dissolved in Assay Buffer (25 mM HEPES-NaOH pH 8.0, 100 mM NaCl and 5 mM  $\beta$ -mercaptoethanol) to a concentration of 1 mM. For the experiment, the peptide was diluted to a final concentration of 5 nM in Assay Buffer. CtTel1(1-1864) was diluted stepwise from 5  $\mu$ M to 0.08  $\mu$ M and mixed 1:1 with the prediluted peptide and incubated for 30 min at room temperature. The fluorescence anisotropy was measured as described before. The experiment was performed in triplicates and the results were analysed using GraphPad Prism and fit according an equation assuming one specific binding site.

### 2.2.6.4 Pull-down assays

Anti-MBP-nanobody beads (ChromoTek) were used for pull-down experiments, as ctMRN( $\Delta$ CC) (work by Aaron Alt and Brigitte Kessler) has a tendency to stick to most other resins. Therefore, a construct of ctTel1(1-1864) containing an MBP-tag was cloned and purified according to the protocol described above.

The anti-MBP slurry was preequilibrated in Binding Buffer (100 mM KCl, 25 mM HEPES pH 7.5, 5 % (v/v) glycerol and 0.5 mM DTT). 10  $\mu$ g of MBP-ctTel1(1-1864) was incubated with the beads for 1 h under rotation in the cold room. The supernatant was collected for analysis on SDS-PAGE gel and the beads were washed three times with Binding buffer. The beads were then incubated with ctMRN( $\Delta$ CC) for 1 h and the same procedure was repeated. The final wash fraction and the beads were analysed on SDS-PAGE gel. The experiment was also performed with ctMR( $\Delta$ CC). As controls, MBP-ctTel1(1-1864) and ctMR(N)( $\Delta$ CC) were subjected to the same procedure separately.

### 2.2.6.5 Analytical gel filtration

To assess the stability of the interaction between ctTel1(1-1864) and ctMRN( $\Delta$ CC), analytical gel filtration assays were performed using a Superose 6 Increase 5/150 column coupled to an ÄKTA Micro system (GE Healthcare, Penzberg), with a flow of 0.2 mL/min.

ctMRN( $\Delta$ CC) and ctTel1(1-1864) were prediluted in Gel Filtration Buffer (25 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 1 mM TCEP) to a final concentration typically between 5  $\mu$ M to 10  $\mu$ M. The reactions were incubated on ice for 30 min prior to gel filtration analysis. The collected fractions were analysed on SDS-PAGE gel.

To test the effect of ATP, a gel filtration assay was performed with 200  $\mu$ M ATP added to the gel filtration running buffer and 1 mM ATP $\gamma$ S added to the Gel Filtration Buffer.

### 2.2.6.6 Kinase activity assays

To monitor the kinase activity of endogenous ctTel1 and its response to potential stimulating factors, kinase assays were performed. Endogenous ctTel1 were used as purified (25 nM final concentration). As a test substrate the histone tail of ctH2AX, *in vivo* one of the main targets of ATM, coupled to GST (GST-ctH2AX(120-133)) was used to a final concentration of 1  $\mu$ M. To determine the preferred metal ion for ATM activity, MgCl<sub>2</sub> or MnCl<sub>2</sub> were added to a final concentration of 250  $\mu$ M. The reactions were started by the addition of ATP, typically in a final concentration of 0.5 to 1 mM. Reactions were incubated at 42 °C for 30 min to 60 min. Reactions were stopped by the addition of Laemmli buffer and denaturing of the proteins at 95 °C for 5 min. Reactions were run on an SDS-PAGE gel at 100 V for 60 min for a better separation of the bands.

Phosphorylated proteins were visualised using either of two approaches. The gel was either stained using a phospho-stain (ProQ Diamond, ThermoScientific), which consists of a fluorescent dye coupled to a phosphate binding moiety, according to the instructions from the supplier. The other approach was to use  $\gamma$ -<sup>32</sup>P-ATP, diluted 1:10 in the ATP solution, and visualisation of the incorporated radiolabelled phosphate using autoradiography on a phosphoscreen. For both approaches bands were visualised using a Typhoon imaging system (GE Healthcare, Penzberg).

To test whether observed kinase activity is specific for ctTel1, different PIKK and ATM-specific inhibitors were tested. Prior to the addition of ATP to start the reaction, ctTel1 was incubated for 30 min with either the general PIKK inhibitor wortmannin (0.5  $\mu$ M), KU55933 (10  $\mu$ M), KU60019 (10  $\mu$ M), caffeine (2 mM) or Torin-2 (1 mM). The final concentrations are in brackets. All inhibitors were diluted in reaction buffer.

## 2.2.7 Cryo-EM

### 2.2.7.1 Cryo-EM grid preparation

Holey R2/1 copper grids with 2 nm continuous amorphous carbon (Quantifoil, Grosslöbichau) were glow-discharged (EMS GloQube). Directly before applying the full-length ctTel1 sample to the grid, TWEEN-20 was added to a final concentration of 0.05 %. Grids were prepared on a Leica EM GP cryo-plunger by applying 4.5  $\mu$ L of sample to the grid. The sample was then preincubated for 45 sec on the grid at 15 °C and 95 % humidity and blotted for 2 sec before vitrification in liquid ethane. The grids were stored in liquid nitrogen until further use.

### 2.2.7.2 Cryo-EM data acquisition

Three datasets were collected on an FEI Titan Krios transmission electron microscope operated at 300 kV with a GIF quantum energy filter (slit width 40 eV) and a Gatan K2 Summit direct electron

detector operated in counting mode. Automated data collection was done using EPU (FEI) with a defocus ranging from  $-1.2\ \mu\text{m}$  to  $-3.5\ \mu\text{m}$  at a calibrated pixel size of  $1.059\ \text{\AA}/\text{pixel}$  and a dose rate of  $1.34\ \text{e}/\text{\AA}^2$  frame, 8 sec exposure over 40 frames.

## 2.2.8 Cryo-EM data analysis

### 2.2.8.1 Data analysis

All data processing was performed using Relion 3.0 and resolutions reported are based on the gold standard FSC 0.143 correlation. Motioncor (as implemented within Relion 3) was used for beam induced motion correction of the movie frames. CTFFind 4.1 was used to determine the CTF parameters of the motion corrected micrographs. Micrographs that showed abnormal Fourier patterns, such as ice rings due to contamination, and did not show any information beyond  $4.5\ \text{\AA}$  were discarded.<sup>205–207</sup>

Particles were picked within Relion 3 using a 3D reference which was obtained from an earlier, lower resolution dataset and was filtered to  $20\ \text{\AA}$  to avoid model bias. After particle extraction particles were sorted according to their figure of merit. Particles that were obvious contaminants, such as ethane and ice, were removed. 3D classifications were performed using a reference filtered to  $60\ \text{\AA}$  to clean the dataset. 286,284 particles from dataset 1, 269,295 particles from dataset 2 and 308,368 particles from dataset 3 were selected for initial 3D refinement. As a reference a 3D reference from an earlier test-dataset that was filtered to  $40\ \text{\AA}$  was used, resulting for each dataset in a  $3.9\ \text{\AA}$  map, which showed significant flexibility in the N-terminal region. The particles of each dataset were then subjected to Bayesian polishing, CTF refinement and subsequently another round of 3D refinement, resulting in three maps at  $3.5\ \text{\AA}$ , with still significant flexibility in the N-terminal region. The datasets were then normalised and merged. The resulting 863,937 particles were used for another round of 3D refinement, which resulted in a  $3.2\ \text{\AA}$  map.

Further 3D classifications with a mask encompassing the kinase domain and fixed angles yielded a class of 613,262 particles that led to a  $3.0\ \text{\AA}$  map of the kinase domain after 3D refinement and a  $2.8\ \text{\AA}$  map with C2 symmetry applied, which allowed for *de novo* modelling.

In order to sort out the conformational heterogeneity of the N-terminal domain, another 3D classification using fixed angles derived from a previous refinement and a mask for one of the N-terminal domains was performed. This yielded a class of 111,910 particles of the  $\alpha$ -solenoid in an open conformation, and a class of 132,127 particles in a closed conformation. Subsequent 3D refinement yielded a map of  $3.6\ \text{\AA}$  for the open conformation and  $3.4\ \text{\AA}$  for the closed conformation.

Using a mask encompassing the  $\alpha$ -solenoid of the other protomer, the 132,127 particles for the closed conformation were used for another round of masked 3D classification with fixed angles. This way a class was obtained, containing 51,116 particles, with both protomers in a closed conformation.

Autorefinement of these particles without applying symmetry led to a 4.0 Å map and with applying C2 symmetry to 3.8 Å.

The map of the two protomers in a closed conformation still displayed some degree of heterogeneity in the N-terminal region. In order to improve the density further, a mask was created enclosing the N-terminal regions of both protomers and another round of masked 3D classification with fixed angles was performed. This yielded a class consisting of 32,764 particles in which both  $\alpha$ -solenoids were in a better-defined closed conformation. Autorefinement with C2 symmetry applied led to a more uniform density for the entire particle at an overall resolution of 3.7 Å.

### 2.2.8.2 Model building

A homology model of the ctTel1 kinase domain was generated using SwissServer based on the kinase domain of hsATM (5NP0) and was docked into the density of the 2.8 Å ctTel1 FATKIN domain using USCF Chimera.<sup>208–210</sup> Rigid body fitting was performed in PHENIX.<sup>211</sup> All residues beyond the kinase domain itself were built manually in Coot.<sup>212</sup> As ctTel1 is predominantly  $\alpha$ -helical, sequence based secondary structure predictions could be used to guide model building. Side chains were visible throughout the density it was possible to find the sequence register for most of. The protein and could build an atomic model for 2652 residues out of 2944 residues. For some loops connecting helices, and for some highly flexible parts of the  $\alpha$ -solenoid no density was observed, preventing modelling for these regions. For the very N-terminus of ctTel1 density for  $\alpha$ -helices was visible, but no side chain information. Based on secondary structure predictions (PsiPred)  $\alpha$ -helices with tentative registers were places in this region.<sup>213</sup>

The open conformation of ctTel1 was modelled using a rigid body fit in PHENIX of the Spiral of the  $\alpha$ -solenoid as had been built in the closed conformation, as no significant conformational changes within the  $\alpha$ -solenoid were observed at this resolution.

After initial building and fitting of all residues for which density was observed, further rounds of refinement and model building were performed in PHENIX. The model of the FATKIN domain and of the closed conformation were refined against their corresponding maps with real space refinement in PHENIX using Ramachandran and  $C_{\beta}$ -restraints and further restraints for secondary structure. For the modelling of the closed and open conformation of the  $\alpha$ -solenoid, self-restraints on the initial coordinates for the Pincer and the FATKIN were included.

To generate a model of the full dimer in the closed conformation, the other protomer was generated by applying C2 symmetry and then performed a rigid body fit of both protomers in the corresponding map encompassing the entire ctTel1 particle at 3.8 Å in PHENIX.

### 2.2.9 Bioinformatics and software

All figures containing models were prepared using PyMOL (Schrödinger LLC, 2015) and EM maps were visualised using Chimera and ChimeraX.<sup>210,214</sup> Multiple sequence alignments were performed using MAFFTS.<sup>215</sup> and visualised in JalView.<sup>216</sup> Percentages of identity and similarity were calculated using the Sequence Manipulation Suite.<sup>217</sup> Primers were designed using SnapGene (GSL Biotech LLC).

# Chapter 3

## Results

### 3.1 Protein biochemistry

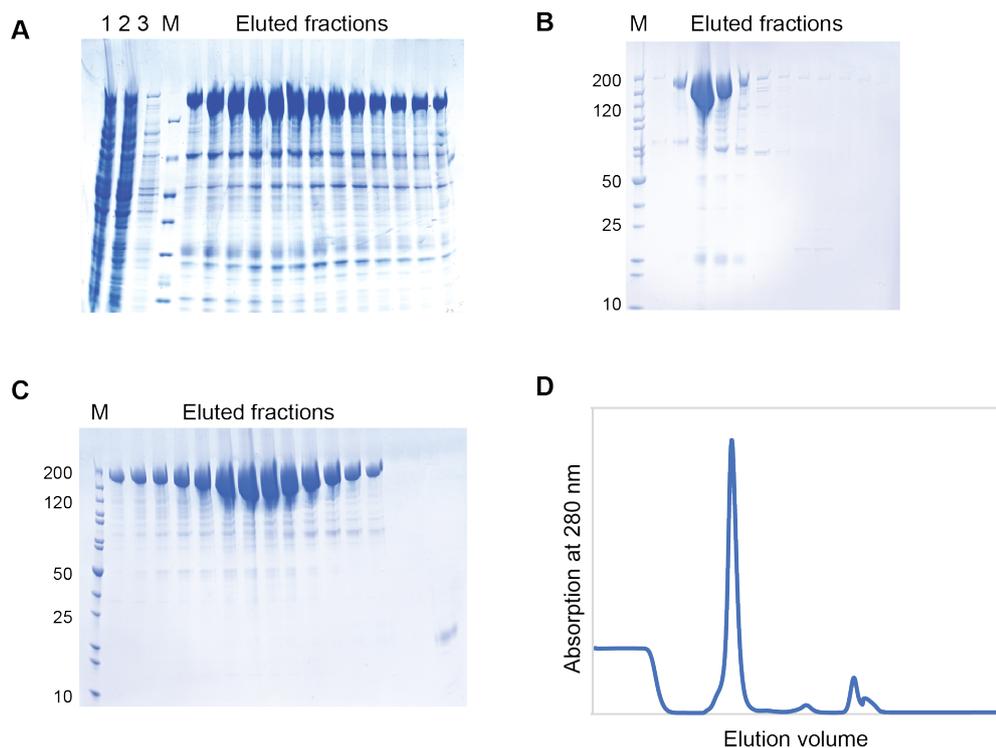
#### 3.1.1 Purification ctTel1(1-1864)

Previous attempts to express full-length ctTel1 in any recombinant system tested were unsuccessful (work by C. Linke-Winnebeck). The expression tests of different constructs of full-length ctTel1 led to the identification of a truncated construct comprising the first 35 predicted HEAT repeats, consisting of the first 1864 residues, which could be expressed in *E. coli* in large quantities. This construct, ctTel1(1-1864) enabled biophysical assays and crystallography trials.

The original ctTel1(1-1864) contained a C-terminal TEV-cleavage site before the His<sub>6</sub>-tag, which was removed by blunt-end ligation to make the construct shorter, as TEV-cleavage was not used because it would lead to much lower yields due to inefficient cleavage.

CtTel1(1-1864) was expressed in *E. coli* Rosetta, yielding soluble overexpressed protein, as is shown by the band at 210 kDa in the lysate and the supernatant (Figure 3.1). Purification using Ni<sup>2+</sup>-NTA and the His<sub>6</sub>-tag fused to the C-terminus of ctTel1(1-1864) removed many impurities, although strong contamination bands were still visible in the eluted fractions. Most of these impurities were removed by anion exchange chromatography using a HiTrap Q column. The purest fractions were then applied on a HiLoad Superdex 200 16/60 size exclusion chromatography column, yielding a sharp peak for CtTel1(1-1864). Minor impurities were still observed after this final purification step. However, when comparing the intensity of the bands on SDS-PAGE gel, the CtTel1(1-1864) sample is >95% pure.

The yield for 1 L *E. coli* culture was 5 mg of pure protein. Due to these high yields, CtTel1(1-1864) was suitable for further biochemical and biophysical assays.



**Figure 3.1: Purification of ctTel1(1-1864) from *E. coli* Rosetta.** **A.** Coomassie-stained SDS-PAGE gel of the His-tag purification step. The upper band corresponds to ctTel1(1-1864). 1: lysate, 2: supernatant, 3: wash fraction, M: marker. **B.** Coomassie-stained SDS-PAGE gel of the anion exchange purification step. **C.** Coomassie-stained SDS-PAGE gel of the fractions obtained after gel filtration. **D.** Gel filtration profile. The first peak corresponds to ctTel1(1-1864).

### 3.1.2 Purification of endogenous ctTel1 using nanobodies

Endogenous ctTel1 was purified using nanobodies, as we did not succeed in endogenously tagging ctTel1 or recombinantly expressing full-length ctTel1.

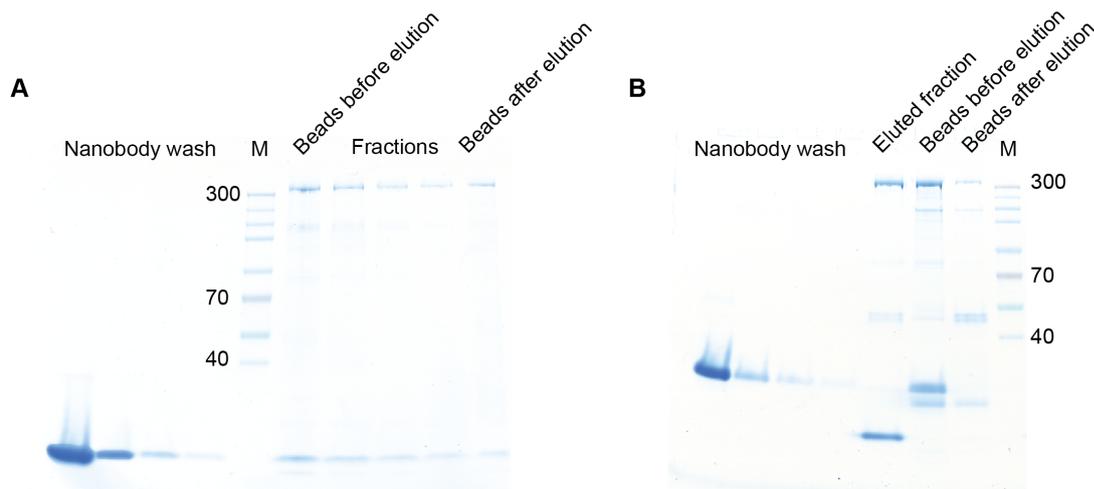
Nanobodies were raised against ctTel1(1-1864) in order to stabilise the protein further for assays and crystallography (work by C. Linke-Winnebeck in collaboration with ChromoTek GmbH, Martinsried). Nanobodies were expressed in *E. coli* Rosetta from a pMJS187 vector, which was engineered to enhance the yield of disulphide bond containing proteins, giving high overexpression rates.<sup>203,204</sup> Purification using the His<sub>6</sub>-tag yielded a very pure sample after one step. The subsequent gel filtration using a HiLoad 75 16/600 column resulted in pure protein without any major contaminants. The yield was 3 to 5 mg/L culture, depending on the different nanobody constructs.

The nanobodies were used to purify ctTel1 directly from its native source, *C. thermophilum*. For one purification, 100g of cryo-milled *C. thermophilum* was used, which corresponds to 12 L of liquid culture. CtTel1-nanobody complex was initially eluted using a FLAG-tag. Analysis on SDS-PAGE gel clearly shows a band that runs higher than 300 kDa, and a band corresponding to the nanobody at around 20 kDa (Figure 3.2). It was confirmed by liquid chromatography-tandem mass

spectrometry (LC-MS/MS) finger printing at the Protein Analysis Unit (ZfP, LMU Munich) that the 320 kDa band corresponds to ctTel1.

As the amounts of ctTel1 were too small for experiments to optimise the buffer, such as Thermo-Fluor, or further purification steps such as gel filtration, the protein was used as it was eluted from the anti-FLAG-beads for further kinase assays and cryo-EM analysis. As there was concern that the negatively charged FLAG-peptide might interfere with biochemical assays and cryo-EM grid preparation, the nanobodies were recloned to add a PreScission protease cleavage site in between two G<sub>4</sub>S linkers to enable efficient proteolytic cleavage.

Using PreScission protease, it was possible to cleave the nanobody-ctTel1 complex of the anti-FLAG beads. As this form of elution could be performed in one step, the sample volume could be reduced, leading to a more concentrated ctTel1. It was still not possible to determine the concentration of ctTel1 accurately with for example using aNanodrop, due to the small amounts and the presence of the PreScission protease and the nanobody. The concentration of ctTel1 was estimated by comparing the size of the band to the size of a band of a similarly large protein, of which the concentration and the dilution factor was known (data not shown). This way, it was estimated that the ctTel1 concentration after proteolytic cleavage is around 100 nM.



**Figure 3.2: Purification of endogenous ctTel1 from *C. thermophilum* with a nanobody.** **A.** Coomassie-stained SDS-PAGE gel of a nanobody-based ctTel1 purification. The ctTel1-nanobody complex was eluted using FLAG-peptide. The upper band corresponds to ctTel1, the lowest band to the nanobody. **B.** Coomassie-stained SDS-PAGE gel of a nanobody-based ctTel1 purification. The ctTel1-nanobody complex was eluted using PreScission protease. The upper band corresponds to ctTel1, the lowest band to the nanobody. Note the size difference between uncleaved nanobody in the wash fraction and cleaved nanobody in the eluted fraction.

### 3.1.3 Expression trials and purification of recombinant ctTel1

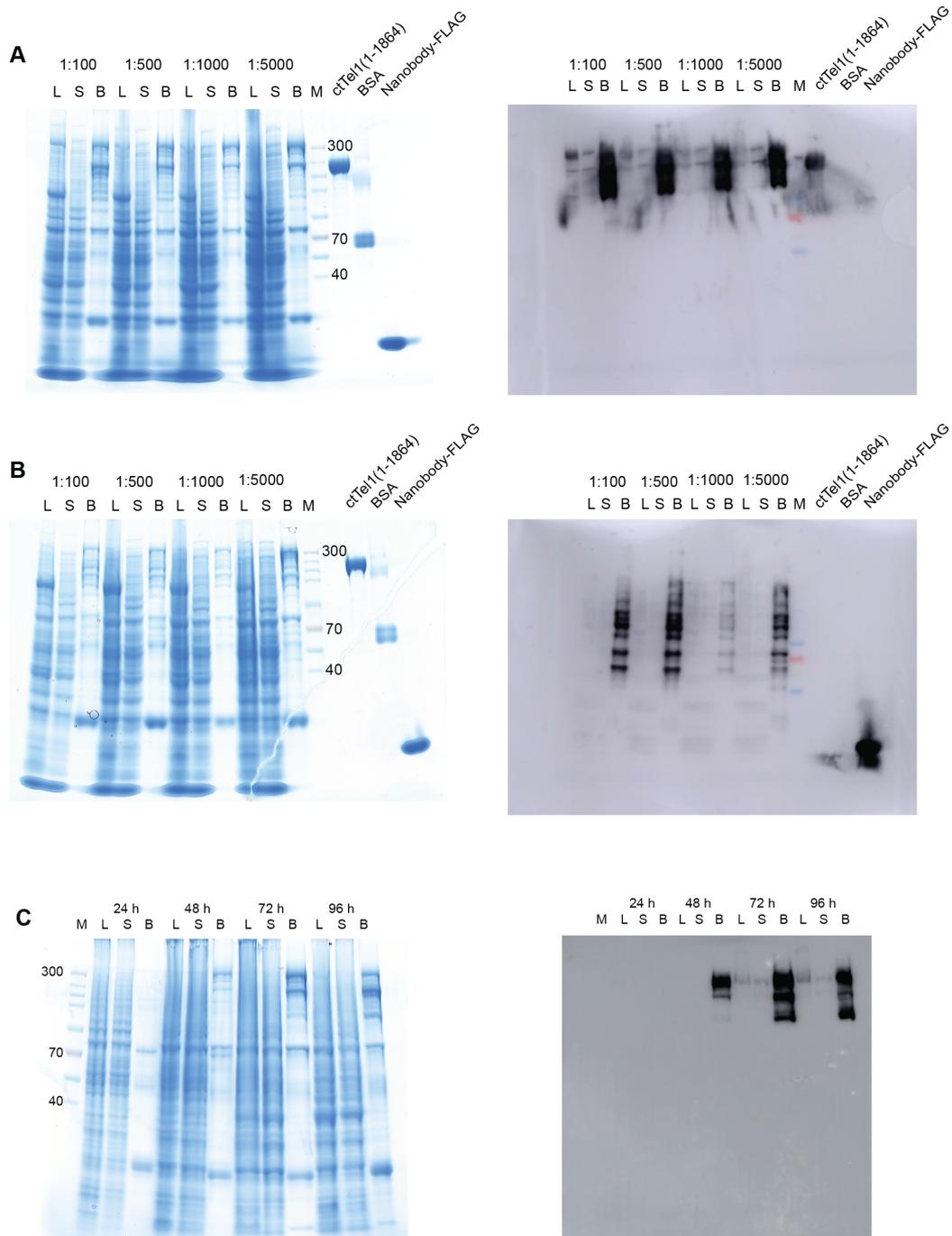
Previous attempts to express full-length ctTel1 were unsuccessful. The small quantities of endogenous ctTel1 were limiting for many experiments, mutants could not be generated and a nanobody was always bound to the protein, potentially influencing activities or interfering with complex formation. Therefore, it was attempted to express recombinant ctTel1 with an N-terminal FLAG-tag in insect cells using a codon-optimised gene for the kinase domain. Previous trials (work by C. Linke-Winnebeck) in insect cells had yielded some protein, albeit insoluble, and therefore insect cells seemed to be the most promising system to optimise.

Codon-optimisation of genes may improve expression levels and yield more stable protein. As CtTel1(1-1864) could be expressed in *E. coli* without problems, it was concluded that the kinase domain was the problematic region for expression. Therefore a codon-optimised gene for the ctTel1 kinase domains was ordered and fused to the truncated ctTel1(1-1864) construct using Gibson assembly. A kinase dead (KD) mutant (D2760A and N2765K) was generated using site-directed mutagenesis. The constructs were then cloned into a pACEbac vector for expression in HighFive insect cells.

Different virus titres were tested for both the wild-type and the KD mutant to find the optimal expression level. Expression was analysed using SDS-PAGE gel analysis and Western blot (Figure 3.3). For all titres tested, expression of soluble wild-type ctTel1 was observed. However, on the SDS-PAGE gel it is already evident there are many impurities in the sample. The Western blot indeed shows significant degradation of ctTel1. The expression levels for the KD mutant were lower, and furthermore the mutant suffered from more severe degradation than the wild-type. Therefore, work on the KD mutant was abandoned.

In order to assess whether the degradation is due to too long expression times, a time course expression test was performed, and a sample was taken every 24h and analysed as before. After 24 h, no expression is visible yet. The first bands appear after 48 h, but there was initial degradation is visible, although less severe than observed before. After 72 and 96 h more severe degradation was observed.

This indicated that the codon optimisation of the kinase domain was insufficient to enable recombinant expression in insect cells. As degradation started early on in the expression, no further attempts at recombinant expression of full-length ctTel1 were undertaken.



**Figure 3.3: Expression tests of FLAG-ctTel1 and FLAG-ctTel1(KD) in HighFive insect cells** **A.** Coomassie-stained SDS-PAGE gel of the expression levels of FLAG-ctTel1 with different virus titres. The upper band at 300 kDa corresponds to FLAG-ctTel1. L: lysate, S: supernatant, B: FLAG-beads, M: marker. CtTel1(1-1864), BSA and nanobody-FLAG were loaded as controls for the Western blot. **B.** Western blot the the same gel. An antibody against ctTel1 was used to identify FLAG-ctTel1. **C.** Coomassie-stained SDS-PAGE gel of the expression levels of FLAG-ctTel1(KD) with different virus titres. **D.** Western blot of the same gel. FLAG-ctTel1(KD) was visualised using an antibody against the FLAG-tag. **E.** Coomassie-stained SDS-PAGE gel of the expression levels of FLAG-ctTel1 over 96h with a virus titre of 1:1000. **F.** Western blot of the same gel. FLAG-ctTel1 was visualised with an antibody against ctTel1.

## 3.2 Binding assays ctTel1(1-1864)

### 3.2.1 DNA binding properties

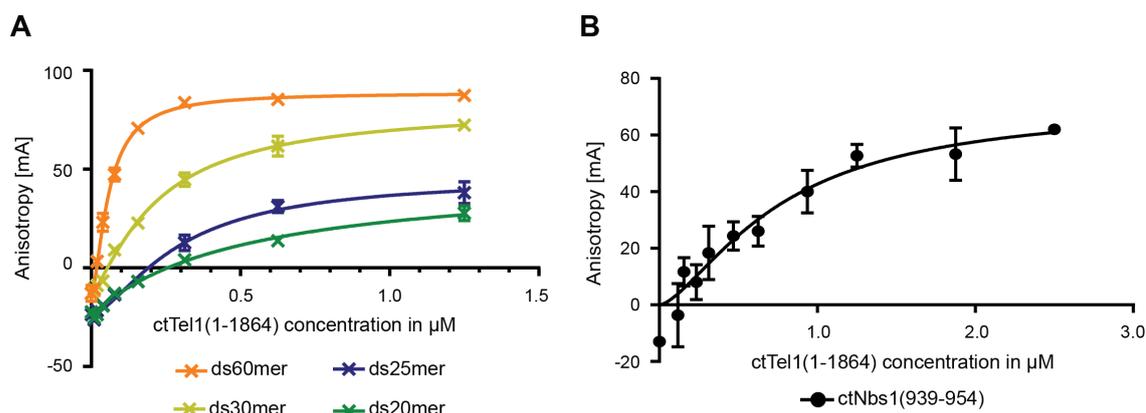
Several studies have demonstrated that ATM has DNA binding properties, and that its kinase activity is stimulated by DNA. Early work using atomic force microscopy (AFM) showed that hsATM can bind either to DNA ends or to interior parts of long DNA, and that ATM prefers long DNA. Kinase assays with hsATM and scTel1 also demonstrated that ATM is stimulated by longer DNA.<sup>111,182</sup>

As HEAT repeats often play a role in mediating protein-protein or protein-DNA interactions, ctTel1(1-1864), which only consists out of HEAT repeats, was considered a suitable construct to assay the DNA binding properties. Furthermore, assays to establish binding properties are not possible with the endogenous ctTel1, as they require large amounts of protein.

Previous work by C. Linke-Winnebeck showed that ctTel1(1-1864) also has DNA binding properties, and displays a preference for blunt ended, double stranded DNA, and has a preference for longer DNA (data not shown). The fluorescence anisotropy experiments with blunt-ended, double-stranded DNA of various lengths were repeated with optimised concentrations of protein and DNA oligonucleotides to determine the  $K_D$  more accurately. Different lengths of double-stranded blunt ended were tested (ds20mer, ds25mer, ds30mer and ds60mer) to see whether there is length dependency. All experiments were performed in triplicates.

CtTel1(1-1864) binds DNA and has a preference for longer DNA. While the ds60mer is bound with a  $K_D$  of  $58 \pm 4$  nM, the affinity for the ds30mer, ds25mer and ds20mer is  $K_D = 220 \pm 27$  nM,  $K_D = 300 \pm 49$  nM and  $K_D = 500 \pm 230$  nM, respectively (Figure 3.4A).

To see if the affinity for longer DNA is higher, it was attempted to perform fluorescence anisotropy with longer DNA. However, when the fluorophore is too far removed from the binding site and still can "tumble" independently, it is not possible to measure affinities reliably anymore. Therefore electrophoretic mobility shift assays (EMSAs), either using agarose gels or native-PAGE gels, were performed to assay alternative DNA structures and long DNA fragments unsuitable for fluorescence anisotropy. Furthermore, it would in principle also be possible to observe whether multiple ctTel1(1-1864) bind to one long oligonucleotide. However, the resolution of the agarose gels was too poor to see any differences between oligonucleotides of various lengths. While native-PAGE gels typically allow for a better resolution, ctTel1(1-1864) would not run into native-PAGE gels (data not shown) and therefore this method also did not allow for assaying DNA binding properties.



**Figure 3.4: Binding properties of ctTel1(1-1864) measured with fluorescence anisotropy.** **A.** DNA binding properties of ctTel1(1-1864) were assayed using different lengths of 6-FAM, 5'-labelled double stranded DNA oligonucleotides, ranging from a ds20mer to a ds60mer. The DNA concentration was kept constant at a final concentration of 5 nM, while the concentration of protein ranged from 0 to 1.25  $\mu\text{M}$ . **B.** Binding properties of ctTel1(1-1864) towards the C-terminal tail of ctNbs1 was assayed using a FAM-labelled peptide of ctNbs1(939-954). The peptide concentration was kept constant at a final concentration of 5 nM. The concentration of protein ranged from 0 to 2.5  $\mu\text{M}$ .

### 3.2.2 Nbs1 binding properties

Soon after the discovery of the ATM kinase, it was established it interacts with the MRN complex. MRN recruits ATM to the site of a DSB mainly via the C-terminus of Nbs1.<sup>60,88,135</sup> In this process ATM is activated. Previous work by C. Linke-Winnebeck already established that the C-terminus of ctNbs1 can be used to pull down ctTel1(1-1864), but it was not known how strong this interaction was.

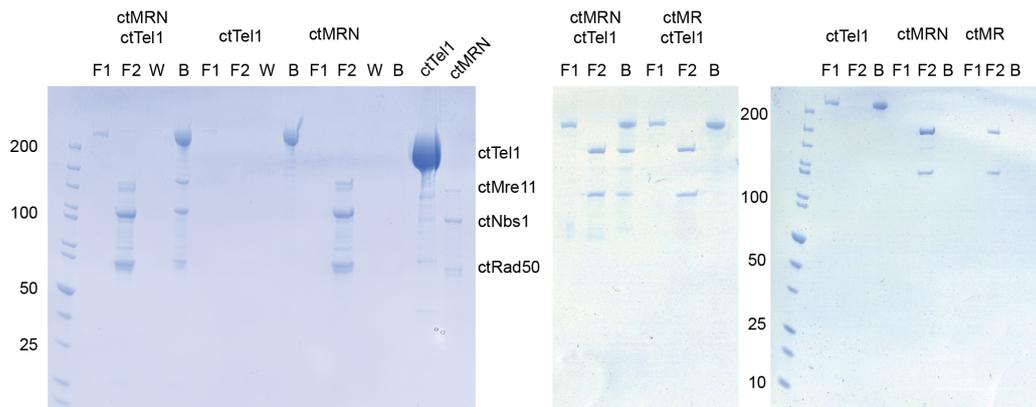
To determine the affinity of the C-terminal fragment of ctNbs1 for ctTel1(1-1864), fluorescence anisotropy was performed using a FAM-labelled Nbs1 peptide (939-954). The experiment was performed essentially the same way as the DNA-binding assays.

Under the conditions tested, the ctNbs1 peptide binds to ctTel1 with a  $K_D$  of  $764 \text{ nM} \pm 219 \text{ nM}$ . This indicates that ctTel1 binds to the ctNbs1 peptide with medium affinity (Figure 3.4B).

In the cell, Nbs1 is part of the MRN complex. It might be the case that the presence of MR and DNA further enhance the interaction with ATM *in vivo*.

### 3.2.3 Pull-down assay with ctMRN( $\Delta\text{CC}$ ) and ctMR( $\Delta\text{CC}$ )

In the cell, the C-terminus of Nbs1 recruits ATM to the DSB. However, in order to activate ATM, the complete MRN complex is needed. It has been reported that the ATP-bound state of Rad50 is required for ATM activation. Furthermore, several studies indicated that ATM does not only



**Figure 3.5: Pull-down assay of MBP-ctTel1(1-1864) and ctMR(N)( $\Delta$ CC).** F1: flow 1, F2: flow 2, W: wash, B: beads. MBP-ctTel1(1-1864) was immobilised on the beads and incubated with ctMRN( $\Delta$ CC) or ctMR( $\Delta$ CC). As controls, MBP-ctTel1(1-1864) and ctMR(N)( $\Delta$ CC) were loaded separately.

directly interact with Nbs1, but also with Mre11 and Rad50.<sup>111,112</sup> Therefore, the interactions of ctTel1(1-1864) with the ctMRN complex were investigated.

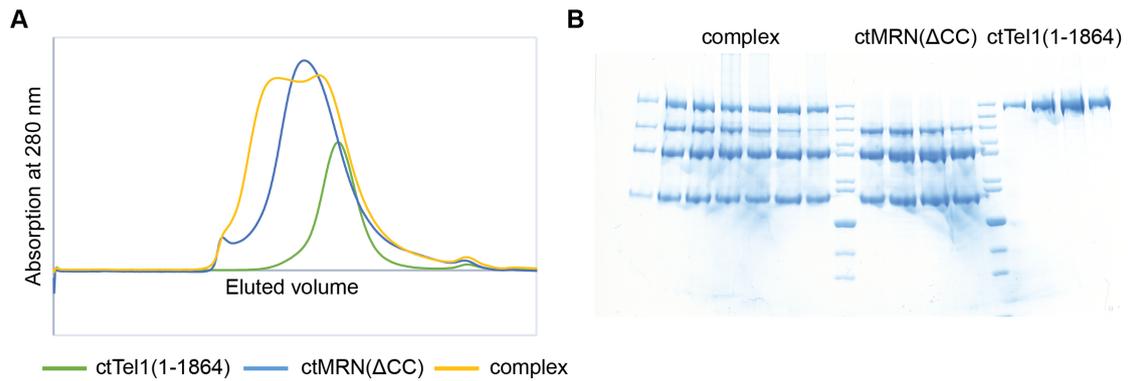
To test whether ctTel1 can bind to ctMRN, and whether it can also bind to ctMR, a pull down assay was performed. As it was not possible to obtain stoichiometric and stable full-length ctMRN, a construct of the ctMRN complex lacking the coiled-coils of Rad50 was used (ctMRN( $\Delta$ CC)). This complex furthermore had a tendency to stick unspecifically to most chromatography resins, making pull-down experiments problematic. However, ctMRN( $\Delta$ CC) was found not to stick to anti-MBP-nanobody-beads (ChromoTek, Martinsried), which were thus used for pull-down experiments.

As a bait, MBP-ctTel1(1-1864) was used, immobilised to anti-MBP beads. Immobilised MBP-ctTel1(1-1864) was incubated with ctMRN( $\Delta$ CC) or ctMR( $\Delta$ CC). It was observed that MBP-ctTel1(1-1864) can retain ctMRN( $\Delta$ CC) but not ctMR( $\Delta$ CC), suggesting that binding is mainly through ctNbs1, which is in accordance with previous studies (Figure 3.5)

However, this does not exclude the possibility that ctTel1 could interact with the coiled-coils of ctRad50 in an intact complex. Alternatively, MR could also interact with ctTel1 via the FATKIN domain which is lacking in ctTel1(1-1864).

### 3.2.4 Analytical gel filtration

As previous experiments showed that ctMRN( $\Delta$ CC) and ctTel1(1-1864) interact, gel filtration assays were performed to assemble a complex of ctMRN( $\Delta$ CC) and ctTel1(1-1864) and to assess the stability of the interaction. CtTel1(1-1864) and ctMRN( $\Delta$ CC) were incubated on ice for 30 min prior to injection on a Superose 6 5/150 analytical gel filtration column. CtTel1(1-1864) and



**Figure 3.6: Analytical gel filtration of ctTel1(1-1864) and ctMRN( $\Delta$ CC).** **A.** Chromatograms of the analytical gel filtration of CtTel1(1-1864), ctMRN( $\Delta$ CC) and the complex. **B.** Coomassie-stained SDS-PAGE gel of the fractions under the peak.

ctMRN( $\Delta$ CC) were analysed separately as well. Furthermore, the experiment was also performed in the presence of ATP, as the ATP-bound form of Rad50 is required for ATM activation.

CtTel1(1-1864) elutes from the column as a sharp peak without any shoulder. ctMRN( $\Delta$ CC) however forms two overlapping peaks. Analysis on SDS-PAGE gel shows that this is due to an under-representation of ctNbs1, leading to a peak consisting of ctMRN( $\Delta$ CC) and a peak consisting of ctMR( $\Delta$ CC)(Figure 3.6).

When ctTel1(1-1864) and ctMRN( $\Delta$ CC) are mixed, there is a significant shift in the first peak. Analysis of the peak fractions shows that ctTel1(1-1864) elutes earlier in the presence of ctMRN( $\Delta$ CC), indicating that ctMRN( $\Delta$ CC) and ctTel1 form a stable complex. However, there is also a clear second peak, consisting of ctTel1(1-1864) and ctMR( $\Delta$ CC). It is not possible to say whether there is any interaction taking place, or whether ctMR( $\Delta$ CC) and ctTel1(1-1864) elute at the same time because their hydrodynamic radius is similar. Addition of ATP gave rise to the same gel filtration profiles (data not shown), suggesting that ATP does not have an effect on the stability of the complex tested. However, in the constructs used the coiled-coils were removed, which have been shown to be essential for Rad50 function, also in its role in activating ATM. It may be that further interactions or stabilisation of an intact ATM-MRN complex may take place in the presence of ATP.

### 3.3 Kinase activity of ctTel1

#### 3.3.1 Initial characterisation kinase activity

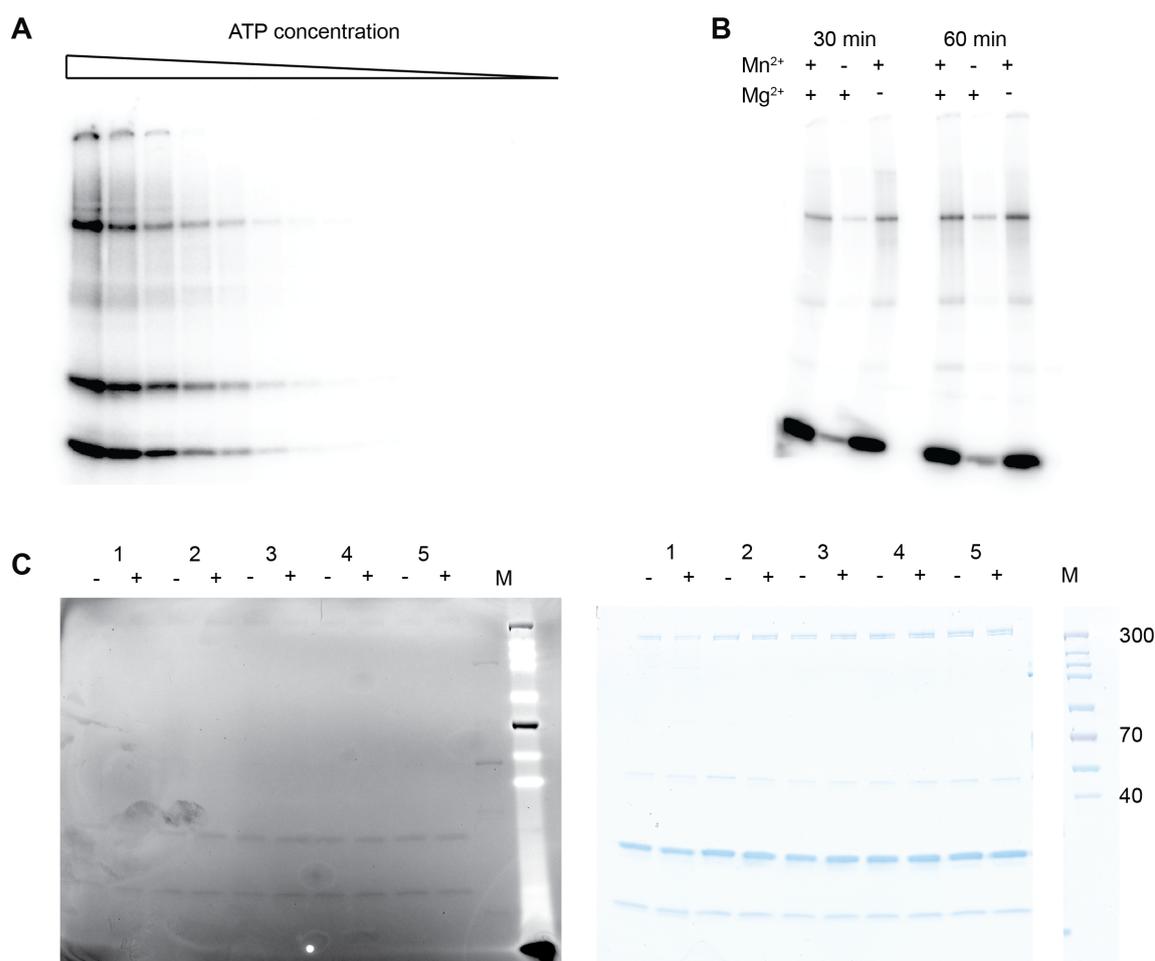
ATM is a highly promiscuous kinase, phosphorylating hundreds of substrates in the cell upon its activation. ATM targets proteins with an S/TQ motif, including other kinases such as CHK1 and CHK2. Other important targets are histone H2AX and MRN itself. Through this wave of phosphorylation, ATM orchestrates the DNA damage response by modulating the chromatin, recruitment and stimulation of DNA repair factors and through control of the cell cycle. Studies on hsATM suggested that MRN and DNA are necessary to stimulate ATM activity. Recently this has also been demonstrated for scTel1. Furthermore, many other factors have been described that stimulate or modulate ATM activity, such as BRCA1.<sup>89</sup>

In order to test basal ATM activity and its response to stimulating factors, an *in vitro* kinase assay was established. As a substrate, a GST-tagged ctH2AX histone tail was used, as H2AX is one of the prime targets of ATM *in vivo*. The kinase assay was done with FLAG-purified ctTel1.

The reactions were run on an SDS-PAGE gel. In order to visualise the phosphorylation of the substrate, the experiment was either performed with  $\gamma$ -P32-ATP and the incorporation of the radioisotope was visualised using autoradiography, or the gel was stained using a phosphostain (ProQ Diamond phosphoprotein gel stain, Thermo Fischer), in case no radioisotopes were used.

Initially, ATP was titrated to find the optimal concentrations that are needed to observe phosphorylation, starting from 1 mM ATP and 0.1 mM  $\gamma$ -P32-ATP. The reactions were run for 60 min. Interestingly, more substrates were phosphorylated than only GST-ctH2AX(120-133), also ctTel1 itself (upper band), the nanobody and a contaminant that was always observed in ctTel1 elutions with FLAG-peptide. Furthermore, no activating factors, such as MRN or DNA were added, suggesting that either ATM has some basal kinase activity, or there is a contaminating kinase responsible for the observed activity (Figure 3.7A).

It was unclear whether ATM prefers  $Mg^{2+}$  or  $Mn^{2+}$ , but it has been reported that described that hsATM prefers  $Mg^{2+}$  for its role in DNA damage signalling, and  $Mn^{2+}$  for redox signalling (Paull paper). Therefore, ctTel1 was incubated with either  $Mg^{2+}$ ,  $Mn^{2+}$  or both. Reactions were analysed after 30 min and after 60 min. CtTel1 is more active in the presence of  $Mn^{2+}$  than in the presence of  $Mg^{2+}$ . In other PIKKs, such as mTOR,  $Mg^{2+}$  was reported to be the physiological metal ion for catalysis, suggesting that  $Mn^{2+}$  despite its higher activity is perhaps not physiologically relevant (Figure 3.7B).



**Figure 3.7: Kinase assays with endogenous ctTel1.** **A.** Titration of different concentrations of (radiolabelled) ATP. Incorporation of radiolabelled ATP was visualised using autoradiography. **B.** Metal dependency of ctTel1. Reactions were run in the presence of  $Mn^{2+}$ ,  $Mg^{2+}$  or both, for 30 or 60 min. Incorporation of radiolabelled ATP was visualised by autoradiography. **C.** Phospho- and Coomassie-stained gels of the reactions of a kinase assay run in presence of inhibitors. 1: wortmannin, 2: KU55933, 3: KU60019, 4: Caffeine, 5: Torin-2. The upper band corresponds to ctTel1, the band at 50 kDa is a contaminant, the first band below 40 kDa is the GST-H2AX(120-133), the lowest band corresponds to the nanobody used for purification.

### 3.3.2 Kinase inhibitors

Due to its apical role in DNA damage signalling, ATM is considered to be an interesting drug target. Inhibition of ATM could lead to the accumulation of DNA damage in cancer cells, thereby enhancing the effect of other cytostatics that induce DNA damage in fast dividing cells.<sup>136</sup> As the active sites of ATM kinases are extremely conserved, some inhibitors that were specifically developed to inhibited hsATM could also be tested on ctTel1. As there are nevertheless minor differences in the active sites, more general PIKK inhibitors were also tested. Furthermore, it was

also important to establish whether the kinase activity observed in previous assays was specific for ctTel1 or caused by contaminating kinases.

Therefore, the kinase assay as described above was repeated, but without radiolabelled ATP. Prior to the addition of ATP to start the reaction, the reaction mixture was incubated with either wortmannin (a potent PIKK inhibitor), KU55933, KU60019 (specific hsATM inhibitors), caffeine (a PIKK inhibitor) and Torin-2 (specific for mTOR, ATM and ATR).<sup>218–222</sup> After adding ATP, the reactions were incubated for 60 min and then analysed on SDS-PAGE gel using both phosphostain and Coomassie.

Weak kinase activity was observed which was not inhibited by the presence of any of the kinase inhibitors. This suggests that the activities observed previously were due to contaminating kinases. Most likely, this is because it was not possible to perform more purification steps due to the very limited amounts of endogenous ctTel1 (Figure 3.7C).

## 3.4 Cryo-electron microscopy

### 3.4.1 Structure determination of ctTel1

Despite the fact that ATM has been studied for over two decades, many details of its cellular function and its activation mechanism remain elusive. This is largely due to the fact that is very difficult to obtain large quantities of full-length ATM for assays and structural biology. In recent years, however, developments in cryo-EM, often referred to as the "resolution revolution", enabled the structure determination of proteins and complexes to resolutions that allowed for building atomic models.<sup>223</sup> The advantage of cryo-EM over crystallography is that tiny quantities of protein are sufficient for grid preparation.

This made the structure determination of full-length endogenous ctTel1 in principle a feasible aim. Initial attempts at preparing negative stain grids were successful. It was, however, difficult to obtain reproducible cryo-EM grids of the ctTel1. The protein would not move into free ice, strictly requiring carbon support. To obtain sufficient particles for cryo-EM data acquisition and data processing, it was necessary to preincubate the protein sample on the grid for at least 45 sec. Even then, the ice quality of the grids was very poor. It was possible to obtain an initial dataset yielding a 7.8 Å map of ctTel1, but it was not possible to reproduce or improve this grid. Curiously, changing from a simple plasma cleaning chamber to a glow discharger yielded reproducible grids with good ice quality and a uniform particle distribution. Combined with the higher concentrations of ctTel1 obtained using proteolytic cleavage enabled the preparation of grids on which large datasets could be collected and the elucidation of the ctTel1 structure.

Preliminary test datasets were collected of ctTel1 in complex with different nanobodies to see which one yielded a more uniform map. For the nanobodies D4 and F7, the overall structure appeared to be

the same. However, nanobody F7 stabilised the very N-terminus more than nanobody D4. Therefore, we focussed on the F7-ctTel1 complex for high resolution structure determination. We collected three different datasets which were merged to enable thorough sorting of different conformations whilst simultaneously allowing for high resolution structure determination (Figure 3.9).

From 863,937 particles we reconstructed a 3D structure of a ctTel1 dimer in different conformations (Figure 3.8). The resolution of the FATKIN was overall 2.8 Å. At this resolution, it was possible to build an atomic model for the entire FATKIN domain, starting from a homology model based on the polyalanine model of hsATM (PDB: 5NP0).<sup>208</sup>

The N-terminal Spiral domain displayed a high degree of flexibility, but using masked 3D classification and refinement for the  $\alpha$ -solenoid of one protomer, two distinct classes were identified showing a closed conformation and a more open conformation. In the closed conformation, the  $\alpha$ -solenoid touches the FAT domain, whereas in the open conformation this interaction is broken. Using smaller masks and more rounds of refinement, a resolution of 3.4 Å was obtained for the closed protomer and a 3.6 Å for the open conformation. The closed conformation was then used to build an atomic model of the  $\alpha$ -solenoid *de novo*. The open conformation, despite its lower resolution, contained at the hinge a part that was better resolved than in the closed conformation. Otherwise, the model for the open conformation was obtained by a rigid body fit of the Spiral into the density.

Due to the higher resolution, the particle subset yielding the closed conformation was subjected to another round of masked 3D-classification, this time with a mask for the  $\alpha$ -solenoid of the other protomer. This yielded a class in which both  $\alpha$ -solenoids were closed. Refinement without symmetry lead to a map of a closed dimer at 4.0 Å and with C2 symmetry to 3.7 Å. Visual inspection of the map revealed that the N-terminus still displayed some flexibility, which resulted in some parts of the maps being patchy and unconnected. Therefore, this particle subset was sorted again using a masked 3D classification, this time with a mask encompassing both  $\alpha$ -solenoids. Refinement with C2 symmetry of the most prominent class yielded again a 3.7 Å map, in which all parts were equally well resolved. This map was then used to build an overall dimeric model of ctTel1, in which the merged structures of the closed  $\alpha$ -solenoid and the FATKIN domain were fitted as rigid bodies into the density.

Overall, we were able to build a side chain model for 90 % of the polypeptide (2652 out of 2944 residues), except for a number of flexible loops and insertions, providing the first near-complete structure of any ATM kinase (Figure 3.9, Table 3.1). Although at the moment of writing several ATM structures have been published, atomic models have only been built for the FATKIN domain, as for all other structures the N-terminal  $\alpha$ -solenoid displayed a high degree of flexibility, allowing maximally for the fitting of a polyalanine chain.

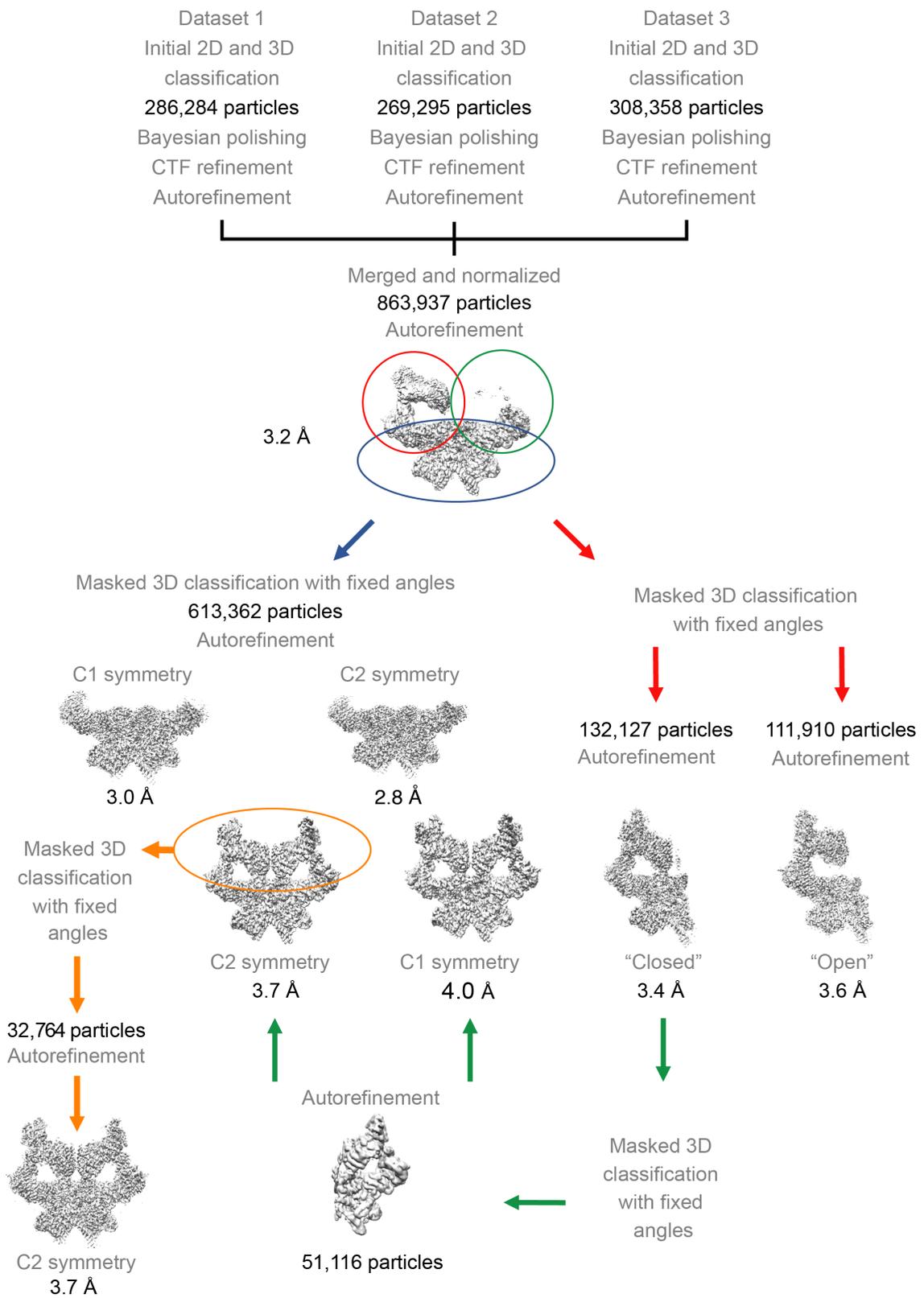
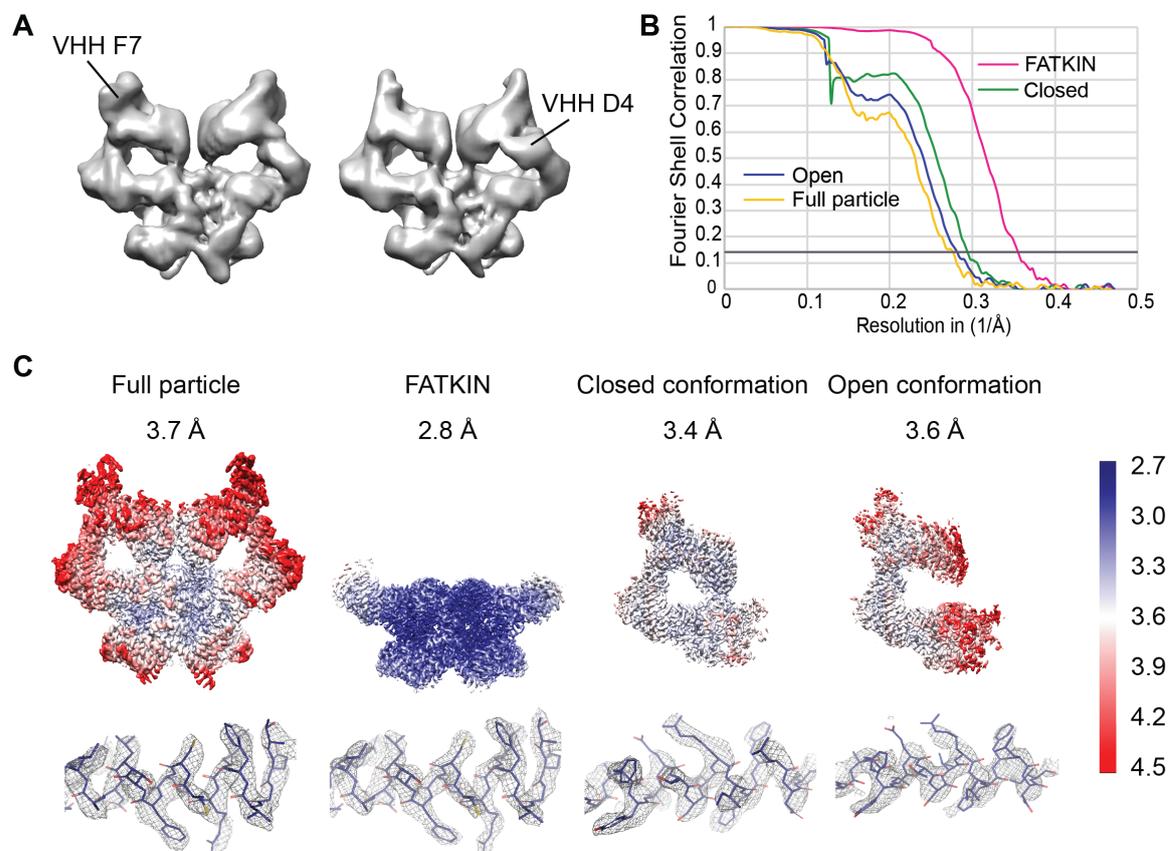


Figure 3.8: Data processing overview.



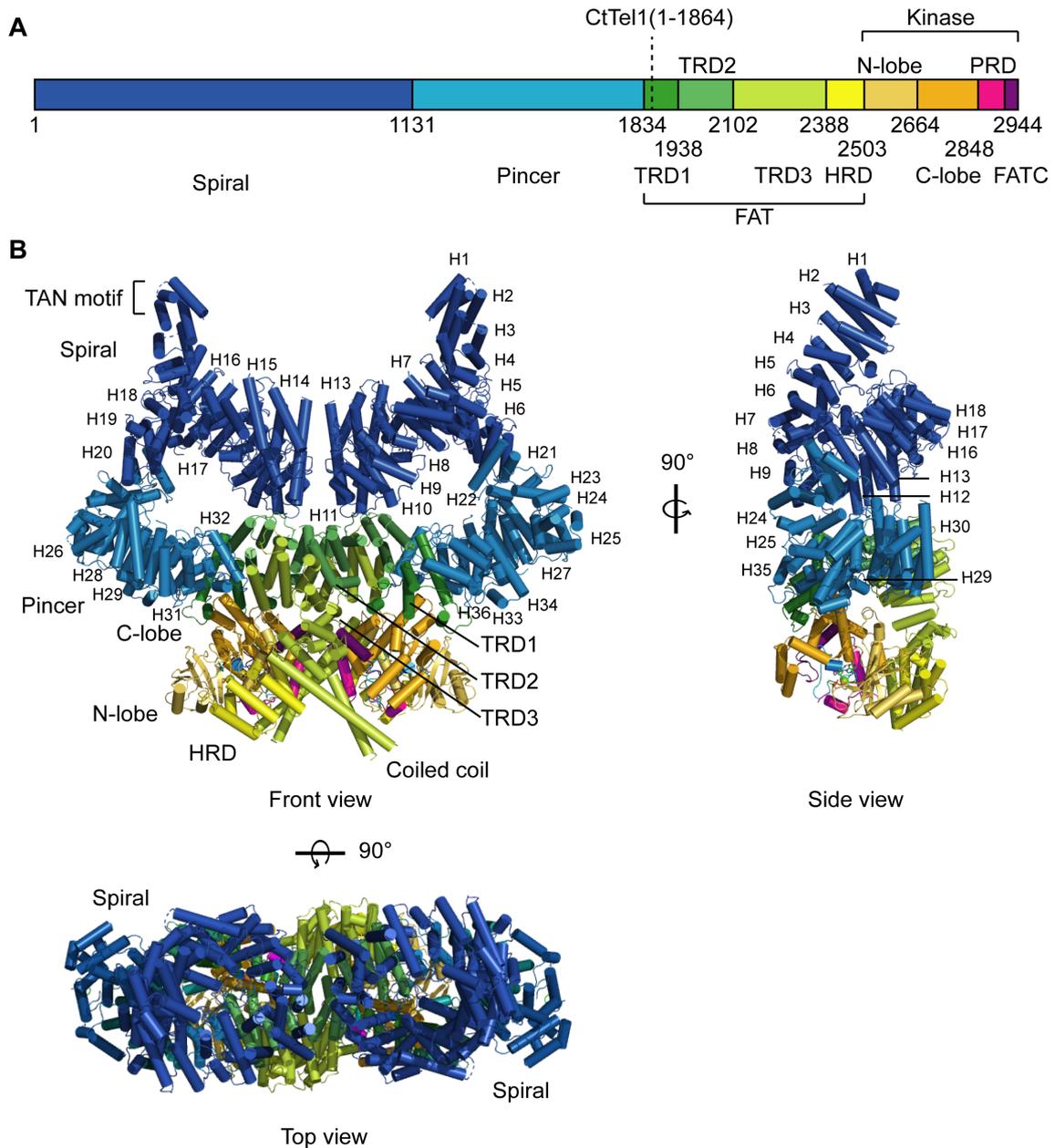
**Figure 3.9: Overview of the resolutions and flexibility within ctTel1.** **A.** Low-resolution maps of ctTel1 with nanobody F7 and D4 bound, respectively. The presence of the nanobody does not alter the structure. **B.** Fourier Shell correlation curves of the different cryo-EM maps. **C.** Cryo-EM maps coloured according to local resolutions, calculated in Relion 3.0. Different maps were calculated for individual domains of ctTel1 due to the high degree of flexibility in the  $\alpha$ -solenoid domain. Examples of local regions show side-chain density.

**Table 3.1: Cryo-EM data collection, 3D reconstruction, model building and refinement statistics.**

	FATKIN dimer	Closed conf.	Open conf.	Complete dimer
<b>Data collection and processing</b>				
Magnification	130 000	130 000	130 000	130 000
Voltage (kV)	300	300	300	300
Electron exposure (e/Å <sup>2</sup> )	1.34	1.34	1.34	1.34
Defocus range (µm)	-1.2 to -3.5	-1.2 to -3.5	-1.2 to -3.5	-1.2 to -3.5
Pixel size (Å)	1.059	1.059	1.059	1.059
Symmetry imposed	C2	C1	C1	C2
Initial particle images (no.)	863937	863937	863937	863937
Final particle images (no.)	613262	132127	111910	32764
Map resolution (Å)	2.82	3.45	3.52	3.65
FSC threshold	0.143	0.143	0.143	0.143
<b>Refinement</b>				
Model resolution (Å)	2.83	3.48	3.62	3.80
FSC threshold	0.5	0.5	0.5	0.5
Map sharpening B factor (Å <sup>2</sup> )	-93	-94	-92	-81
<b>Model composition</b>				
Non-hydrogen atoms	22348	20930	21041	41860
Protein residues	2824	2638	2652	5276
Ligands				
Mg <sup>2+</sup>	2	1	1	2
ATP <sub>γ</sub> S	2	1	1	2
<b>B factors (Å<sup>2</sup>)</b>				
Protein	43.8	43.3	52.78	
Ligand				
ATP <sub>γ</sub> S + Mg <sup>2+</sup>	40.1	32.0	40.92	
<b>R.m.s. deviations</b>				
Bond lengths (Å)	0.009	0.011	0.011	0.011
Bond angles (°)	0.67	0.80	0.82	0.80
<b>Validation</b>				
MolProbity score				
Clashscore	5.01	8.45	12.43	8.85
Poor rotamers (%)	3.63	5.39	6.46	5.39
<b>Ramachandran plot</b>				
Favored (%)	95.77	92.49	91.77	92.45
Allowed (%)	4.23	7.43	8.15	7.47
Disallowed (%)	0.0	0.08	0.08	0.08

### 3.4.2 Overall architecture of ctTel1

As described for other ATM kinase structures.<sup>188,189,208,224–227</sup>, also ctTel1 forms a dimer with a shape reminiscent of a butterfly. At the N-terminus, there is a Spiral domain, which is followed by a Pincer domain. These two domains combined form an  $\alpha$ -solenoid, which is entirely  $\alpha$ -helical. Towards the C-terminus, the  $\alpha$ -solenoid is followed by a FAT domain and a canonical bilobal kinase domain at the C-terminus. The C-terminus contains a number of characteristic insertions, such as the PRD and the FATC at the very C-terminus (Figure 3.10).



**Figure 3.10: The cryo-EM structure of the ctTel1 dimer in a closed conformation. A.** Domain organisation of ctTel1 indicated in different colours. The C-terminus of tTel1(1-1864) is indicated with a dashed line. **B.** Front and top views of the ctTel1 dimer, and a side view of one protomer. The HEAT repeats are labelled sequentially from N-terminus to the C-terminus. The colours correspond to the colours used in the domain overview.

The overall arrangement of domains in ctTel1 is the same as in scTel1 and hsATM, showing the high degree of conservation throughout evolution (see alignment in Appendix). This may not be so surprising for the FAT and kinase domains, of which the sequence is not only highly conserved

between ATM kinase orthologs, and also among PIKKs in general. Interestingly, despite the lower sequence conservation of the N-terminal  $\alpha$ -solenoid, the fold is nevertheless highly similar.

CtTel1 forms a dimer, of which most of the interface is formed by the FAT domain, which has been described for other ATM orthologs too. In total, ctTel1 has 3,500  $\text{\AA}^2$  buried surface area, as calculated using PISA.<sup>228</sup> For ctTel1 we observe a second, transient dimer interface between the two  $\alpha$ -solenoids of the two protomers. As described in more detail in the next section, the Spiral moves from a closed conformation to an open conformation relative to the FATKIN, with the hinge being formed by the Pincer. In the open conformation, the dimer interface is only formed by the FATKIN domain, whereas in the closed conformation Spiral domains can interact. Interestingly, the Spirals appear to move independently of each other, giving rise to either symmetric open and closed conformations, or asymmetric conformations in which one protomer is in a more open or closed conformation than the other.

Other studies have suggested that conformational changes in the N-terminus are coupled to conformation changes in the kinase domain, suggesting a potential mechanism for kinase activation.<sup>227</sup> When overlaying the ctTel1 structures of the open and the closed conformation, we do not observe any significant changes in the kinase domain. ATM has been described to form monomers upon activation. We never observed any monomers on the micrographs. Both findings support that ctTel1 is in a stable dimeric and autoinhibited state. More robust conformational transitions induced by, for example the MRN complex or DNA, are required for potential monomerization and activation.

### 3.4.3 Structure of the N-terminal $\alpha$ -solenoid

HEAT repeats and similar helix-turn-helix motifs, such as armadillo repeats, typically play a role in mediating protein-protein interactions or interactions with other macromolecules such as DNA. As the Spiral and Pincer are also entirely  $\alpha$ -helical, this domain of the protein is suggested to be a hub for interactions. Several studies provided both genetic and biochemical evidence that especially the Spiral functions as a hub for effector proteins of the DNA repair machinery, such as Nbs1 and potentially also Mre11 and Rad50, and signalling substrates of the DNA damage response, such as p53.<sup>60,88,144,229</sup>

In order to provide detailed structural insight in the N-terminal domain, we determined its structure to a resolution that allowed *de novo* model building. The  $\alpha$ -solenoid of ctTel1 is flexible, as it can move around a hinge from an open to a closed conformation relative to the FATKIN. This flexibility was also observed in other structural studies of ATM kinases and limited the resolution.

In order to solve a structure of the  $\alpha$ -solenoid, we made use of 3D classification with a mask for only the  $\alpha$ -solenoid. This resulted in a number of maps showing a closed conformation and multiple open conformations, which varied from slightly open to wide open. The densities with the highest resolution information, which was one in the closed state and one in an intermediate open state,

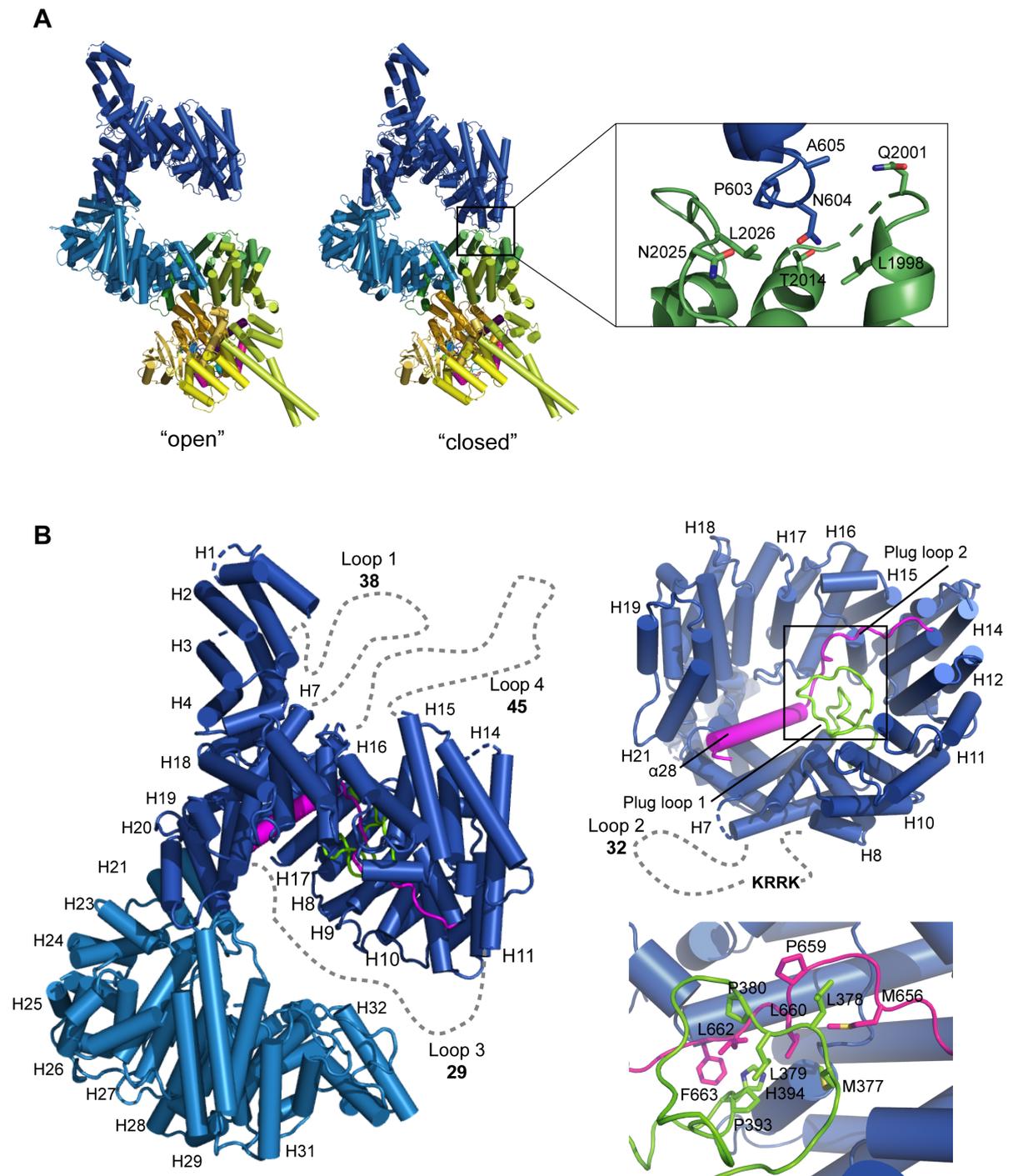
were refined using focused refinement methods. First, we started with a mask encompassing the Spiral and the Pincer. Then we restarted the refinement from the last step with a mask for only the Spiral. This resulted in densities with a resolution of 3.7 Å for the open conformation and 3.4 Å for the closed conformation. At these resolutions, side chain densities can be discerned which enabled us to build an atomic model *de novo* for this part of the protein, in both conformations. In contrast to the kinase domain, no homology model could be used for the  $\alpha$ -solenoid. However, as it is entirely  $\alpha$ -helical, model building could be guided by secondary structure predications. Furthermore, characteristic sequences with very bulky helices, such as multiple aromatic residues in close proximity, also made it possible to find the correct register of the helices.

The ctTel1 Spiral domain consists out of the HEAT repeats H1 to H22 and a number of helical and loop insertions of various lengths. For the sake of simplicity, we named all  $\alpha$ -helical repeats HEAT repeats, although some of the repeats are interspersed with longer loops than is typical and sometimes also contain very short  $\alpha$ -helices, which are not representative of canonical HEAT repeats (Figure 3.11).

The very N-terminus consists of H1 to H4 and forms a flexible but conserved protrusion. Although  $\alpha$ -helical density is visible for H1 and H2, there is no side chain information due to the flexibility. A tentative side chain model was built in these helices based on secondary structure predictions. The first two helices and their connecting loop (H1) contain the highly conserved Tel1/ATM N-terminal motif (TAN motif).<sup>230</sup> This motif is present in ATMs from yeast to human and is essential for telomere maintenance and the DNA damage response. It is conceivable that this motif serves as one of the interaction sites for other macromolecules, such as other proteins or DNA.

H3 and H4 sit at the bottom of the protrusion and interact with H17 and H18 through hydrophilic interactions. This interaction stabilises the highly bent conformation of the Spiral. Another site where the  $\alpha$ -solenoid is strongly bent is at H11-H13. In the closed conformation, these helices in the two protomers interact with each other, and also with the TRD2 domain of the FAT domain within one protomer. This small interface consists out of a mixture of hydrophobic and hydrophilic residues. Neither is it highly conserved, which suggests that the interaction is rather unspecific.

In ctTel1, the entire Spiral except for H1 to H4 appears to be in a rigid conformation and also moves as a rigid body from a closed to more open conformations. It was therefore possible to build the open conformation, which has a slightly worse resolution than the closed conformation, by rigid body fitting.



**Figure 3.11: Structures of the N-terminal  $\alpha$ -solenoid.** **A.** The Spiral domain can move from a closed to an open conformation around a hinge formed by the N-terminus of the Pincer domain. In the closed conformation, there is a small interface between the Spiral and TRD2 of the FAT domain. **B.** Details of the fold of the Spiral and the Pincer. The Spiral has several long, unstructured loops, which are indicated with a dotted line. Their length is shown in numbers. The interior of the Spiral is stabilised by two Plug loops, which form a structured core via hydrophobic interactions.

Apart from the interaction between H3-H4 with H17-H18, the Spiral is further stabilised by two loops in the core of the domain that form together a "Plug". The two Plug loops originate from the HEAT repeats forming the sides of the Spiral and fill the cavity that is formed by the arms of the  $\alpha$ -solenoid. The Plug is highly structured. The first Plug loop is a 40-residue insertion between helices  $\alpha$ 16 and  $\alpha$ 17, which form a HEAT-repeat-like motif. Two patches in this loop, 367-374 and 389-400, are conserved. The insertion is longer in *C. thermophilum* than in ATM orthologs from higher eukaryotes. Plug loop 1 interacts tightly through hydrophobic interactions with Plug loop 2, which is a long insertion (650-715) between H12 and H13. The first, hydrophobic part of the Plug loop 2 interacts with Plug loop 1, forming the core of the Spiral. The loop then reaches across the Spiral domain and ends in a helix ( $\alpha$ 28) that sits between the interface of the N-terminal H4 and C-terminal H20-H22 elements of the Spiral. The third and final part of the Plug loop is a long hydrophilic insert that connects to H13. It spans 60 Å on the interior side the gap underneath the plug, and is unstructured. Its sequence is poorly conserved.

Apart from the Plug loops, the Spiral contains an additional number of relatively long loop insertions for which we did not observe any density. This reflects the dynamic nature of these loops. The first 32-residue unstructured loop (417-447) is located on the outside of the Spiral, connecting alpha helix 17 to H8. Interestingly, this loop contains a highly conserved positively charged patch (KRRK in *C. thermophilum*, KRKK in humans), suggesting that this loop may be involved in interactions with DNA or negatively charged patches on proteins.

The Spiral is connected to the FATKIN via the Pincer. The N-terminal part of the Pincer is forming the hinge around which the Spiral can move as a rigid body. Due to the inherent flexibility of such a hinge, this part of the Pincer is less resolved in our structure. It was not possible to see side chains for a number of  $\alpha$ -helices, and for some  $\alpha$ -helices it was hard to discern any clear density.

As we observed direct binding of the ctNbs1 C-terminus to the  $\alpha$ -solenoid of CtTel1(1-1864), we also sought to obtain a structure of ctTel1 with an Nbs1 peptide bound. Therefore we added the synthetic Nbs1 peptide in a 1000-fold excess over ctTel1 before grid preparation. However, when comparing the dataset with Nbs1 to a dataset without Nbs1, we could not find any additional density that could reliably account for a peptide bound to ctTel1 (data not shown). This could be on the one hand because the peptide is too small to give clear density. Alternatively, ctNbs1 could bind in one of the more flexible parts of the  $\alpha$ -solenoid, such as close the hinge.

#### 3.4.4 FATKIN domain

All PIKKs have a FAT domain which wraps around the kinase domain. Both domains are highly conserved and are often denoted as FATKIN as they form one structural unit. It is likely that the FAT contains key roles in regulating the kinase activity by transferring conformational changes induced by allosteric binding to, for example, the  $\alpha$ -solenoid (ref). The ATM/Tel1 FATKIN is highly conserved both in terms of its sequence and architecture. As in other ATM structures described to

date, the ctTell1 FAT domain is also entirely alpha helical and consists of three tetratricopeptide domains, denoted as TRD1 (1834-1937), TRD2 (1938-2101), TRD3 (2102-2387) and one HEAT repeat domain, HRD (2388-2502). TRD3 contains a coiled-coil extension (2285-2387), which is in close proximity to the kinase domain and therefore also believed to play a direct role in regulating kinase activity.

The kinase domain has a canonical bilobal kinase architecture, consisting of an N-terminal lobe and a C-terminal lobe which form a cleft in which the active site is located. The coiled-coil extension of TRD3 directly interacts with the LBE in the C-lobe via residues 2681-2715. The LBE is a conserved element, which in mTOR binds to LST8. The C-terminal further contains the PRD (residues 2848-2910) and FATC domain (2911-2944), elements which are conserved amongst all PIKKs.

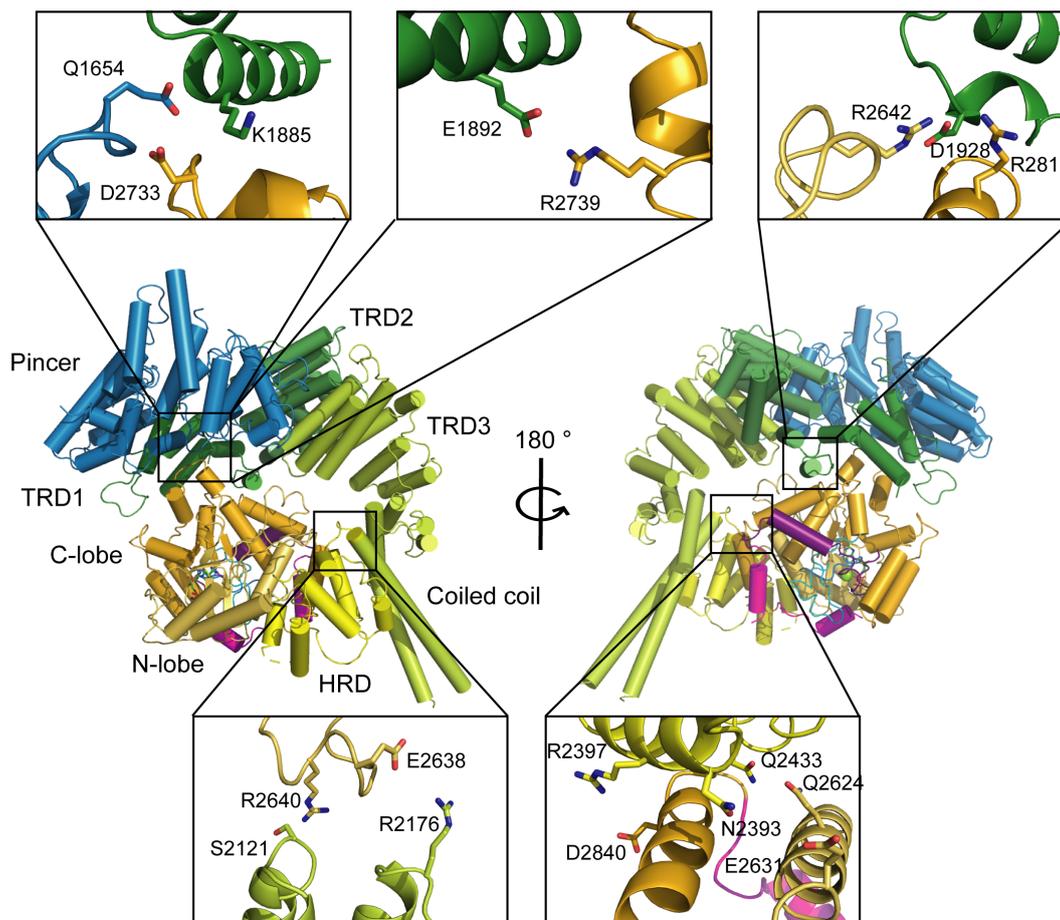
The FAT domain encircles the kinase domain, but only interacts directly with the kinase through a number of distinct interactions, mainly consisting of hydrogen bonds and salt bridges. These interactions are astonishingly conserved throughout evolution, and some are identical from yeast to human. Point mutations in these salt bridges have been associated with A-T and some forms of cancer. This reflects their importance in the stabilisation of the interface between FAT and kinase domain (Figure 3.12).<sup>208,227</sup>

TRD1 and the Pincer interact directly with the C-lobe of the kinase domain via Glu1654 (Pincer)-Lys1885 (TRD1)-Asp2833 (C-lobe). TRD1 also interacts with the kinase domain via the highly conserved salt bridge Glu1892-Arg2739, which has been described in other ATM structural studies as well, and Arg2641-Asp1928-Arg2819.

The coiled-coil within TRD3 and the HRD also interact directly with the kinase domain. A salt bridge is formed by Glu2638-Arg2176. Additionally, there is a hydrogen bond between Ser2121-Arg2640.

The HRD connects the TRD domains of the FAT to the N-lobe of the kinase domain. Due to its proximity, a more extensive surface exists between HRD and the kinase, with highly conserved salt bridges and other hydrophilic interactions formed by Arg2397-Asp2840, Ans2393-Glu2631 and Gln2433-Gln2624.

Through this network of interactions, the FATKIN forms one structural unit.



**Figure 3.12: Overview of the interactions between the kinase domain and the Pincer/FAT domain.** Two rotated views of the Pincer and the FATKIN domains of one protomer are shown. The interacting residues are pictures in boxes, with the interaction sites mapped onto the structures.

### 3.4.5 Dimer interface

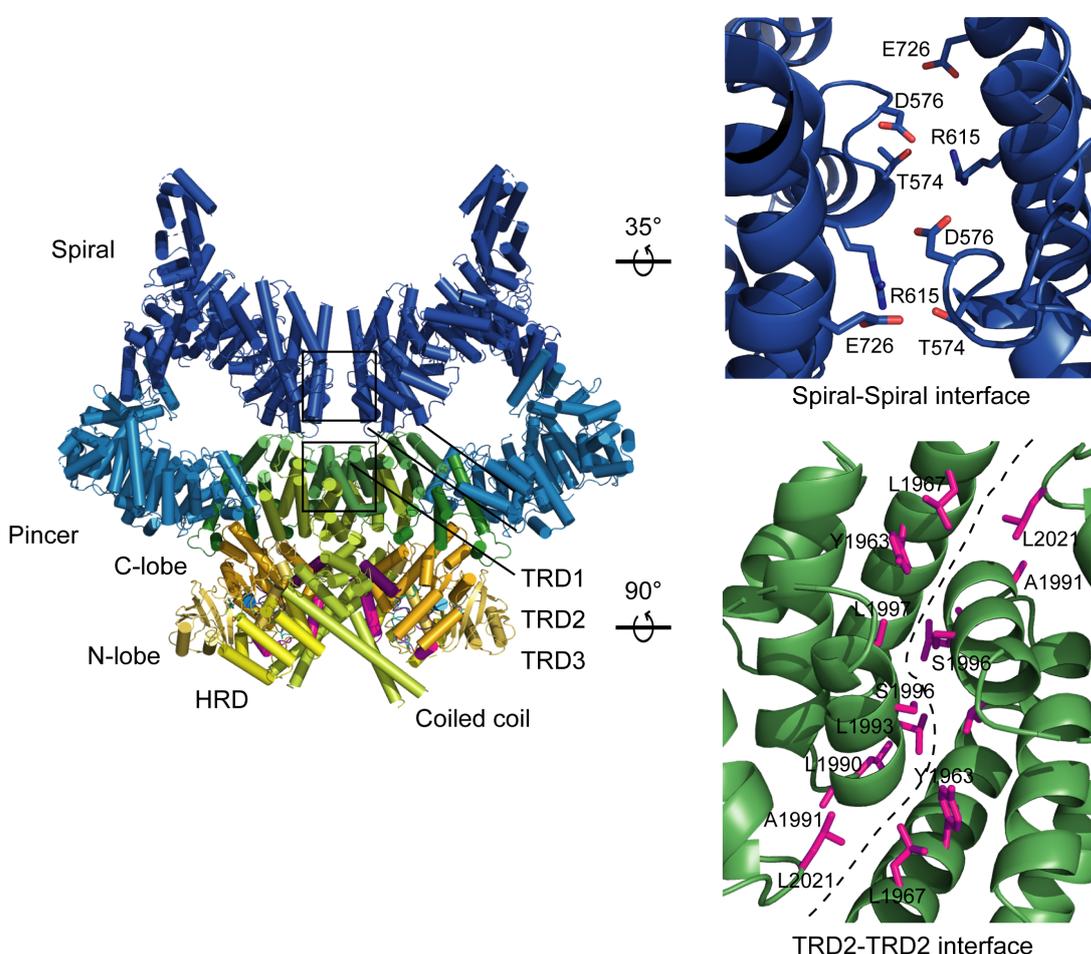
ATM resides in the nucleus as an inactive dimer, and has been reported to monomerise upon activation, although this monomerisation theory of activation is controversial.<sup>46,188</sup> During the structural analysis of ctTel1, only dimers were observed. This is in accordance with the model that ATM is a dimer in absence of any activators. To clarify whether monomerisation would in principle be possible, we analysed the nature of the dimer interface of ctTel1 (Figure 3.13).

CtTel1 has multiple dimer interfaces. Starting from the N-terminus, the Spirals of each protomer form a transient hydrophilic dimer interface, consisting mainly of ion pairs, when the two protomers are in a symmetric closed conformation. Interactions such as between Asp576 and Arg615 from the opposing Spiral domains are in sufficiently close proximity to stabilise the transient interaction.

The major dimer interface is formed by TRD2-TRD2 interactions. The molecular details of this interface were not described in other structural studies of ATM kinases, presumably due to the limited resolution, but our high resolution allowed us to resolve this interface to a higher degree of

detail. The TRD2-TRD2 interface consists mainly out of leucines that form a hydrophobic interface together with Tyr1963. Additionally, Ser1996 of opposing protomers may form a hydrogen bond to stabilise the interface further. The overall interface is 870 Å<sup>2</sup> (calculated using PISA), suggesting that a dimer-to-monomer transition would require a substantial amount of free energy.

Other layers of the dimerisation interface are formed by interactions between TRD2 and TRD3 and by interactions between TRD3 and the kinase domain. The TRD2-TRD3 interface is smaller than the TRD2-TRD2 interface and consists out of a mixture of hydrophobic and hydrophilic contacts. Two polar interactions between TRD2 and TRD3 of the opposing protomer, Gln1962 (TRD2)-Thr 2186 (TRD3) and Ser 1965 (TRD2)-Asp 2190 (TRD3) form a direct connection between these domains. Another function of the interactions between the TRD2-TRD3 may be to further stabilise the TRD2-TRD2 interface.



**Figure 3.13: Details of the dimer interfaces of ctTel1.** The Spiral domains interact in the closed conformations via polar interactions and salt bridges. The TRD2-TRD2 interface is extensive and consists predominantly out of hydrophobic residues. The coiled-coil of one protomer is in close proximity to the FATC, PRD and LBE within the kinase domain of the opposing protomer.

Numerous interactions, mainly hydrophilic, exist between the kinase domain and TRD3 of the FAT domain of the opposing protomer. The coiled-coil within the FAT domain is in close proximity to

the active site of the other protomer, forming another part of the dimer interface. The coiled-coil could form direct interactions with the LBE helix bundle (which is the binding site of LST8 in mTOR). The main interactions are formed by Gln 2354 (coiled-coil)-Thr 2700 (LBE) and Gln 2361 (coiled-coil)-Asn 2702 (LBE).

Furthermore the activation loop within the active site is also in very close proximity to the coiled-coil, forming another potential interaction site. Another set of interactions might be formed between the flexible loop in the PRD and the coiled-coil. Although the PRD insert is not visible, based on the locations of the PRD helices it is in close proximity to the coiled-coil. The insert is glycine rich and contains numerous negatively charged residues, which may interact with a positively charged region (2342-RNRYQSH-2348) in the coiled-coil of the other protomer (see also Figure 3.16).

The FATC domain of one protomer can also directly interact with TRD3 of the other protomer, mainly via hydrophilic interactions. TRD3 forms an additional dimer interface with FATC in the kinase C-lobe, by forming a salt bridge between Arg 2260 (TRD3)-Asp 2922 (FATC).

To summarise, ctTel1 has an extensive dimer interface consisting of multiple layers. The core of the dimer interface is hydrophobic, whereas the other layers are predominantly hydrophilic. This suggests that it is unlikely that ctTel1 forms monomers as part of its activation, but there is more flexibility in the remainder of the molecule. The direct interactions between the TRD3, especially via its coiled-coil, indicate that conformational changes in one protomer may have a direct effect on the kinase domain of the other protomer, providing an initial insight in the elements required for activation.

### 3.4.6 Kinase domain and active site

In order to stabilise ctTel1 and to identify how a nucleotide is bound in the active site, we solved a structure of ctTel1 in complex with ATP $\gamma$ S. It was possible to resolve the kinase domain of ctTel1 to a resolution of 2.8 Å, which allowed for the building of a reliable atomic model of the active site with the ATP $\gamma$ S moiety (Figure 3.14).

As the kinase domain is the most conserved part of ctTel1, it is not only very similar to other ATM kinases but also to the kinase domains of other PIKKs due to a high degree of sequence and structural conservation. The active site is formed by a cleft between the N-lobe and the C-lobe. In contrast to other PIKKs, ctTel1 and other ATM kinases do not have an FRB four-helical bundle in the N-lobe. In the C-lobe, there is an LBE, which in mTOR binds to LST8.<sup>149</sup> As in the other structures reported to date, the activation loop and the catalytic loop, which are located in the C-lobe and are required to enable catalysis, are structured. The activation loop wraps around the catalytic loop and keeps it structured through interactions with the backbone of the catalytic loop, such as between Asp2794 in the activation loop and the backbone of Asp2760 in the catalytic loop. The activation loop itself is structured mainly through hydrophobic interactions with the

FATC domain, which lies buried in the N-lobe. In other known kinase activation mechanisms, the activation mechanism often entails a change of these loop elements from unstructured to structured to enable catalysis to take place in the active site.<sup>231</sup> Thus, the structured active site of ctTel1 is indicative of a catalytically proficient, active kinase.<sup>149</sup>

ATP $\gamma$ S is bound together with a Mg<sup>2+</sup> ion in the active site cleft. The adenine moiety lies in a hydrophobic patch of the active site, which is formed by the side chains of numerous hydrophobic residues from both lobes of the kinase, including Val2664, Ile2778 and Leu2667 (Figure 3.14B) and also stacks against the Phe2663 side chain with one side of the purine. Furthermore, the adenine moiety is coordinated by hydrogen bonds of its NH2 and N6 with the backbone carbonyl of Glu2662 and the backbone nitrogen of Val2664.

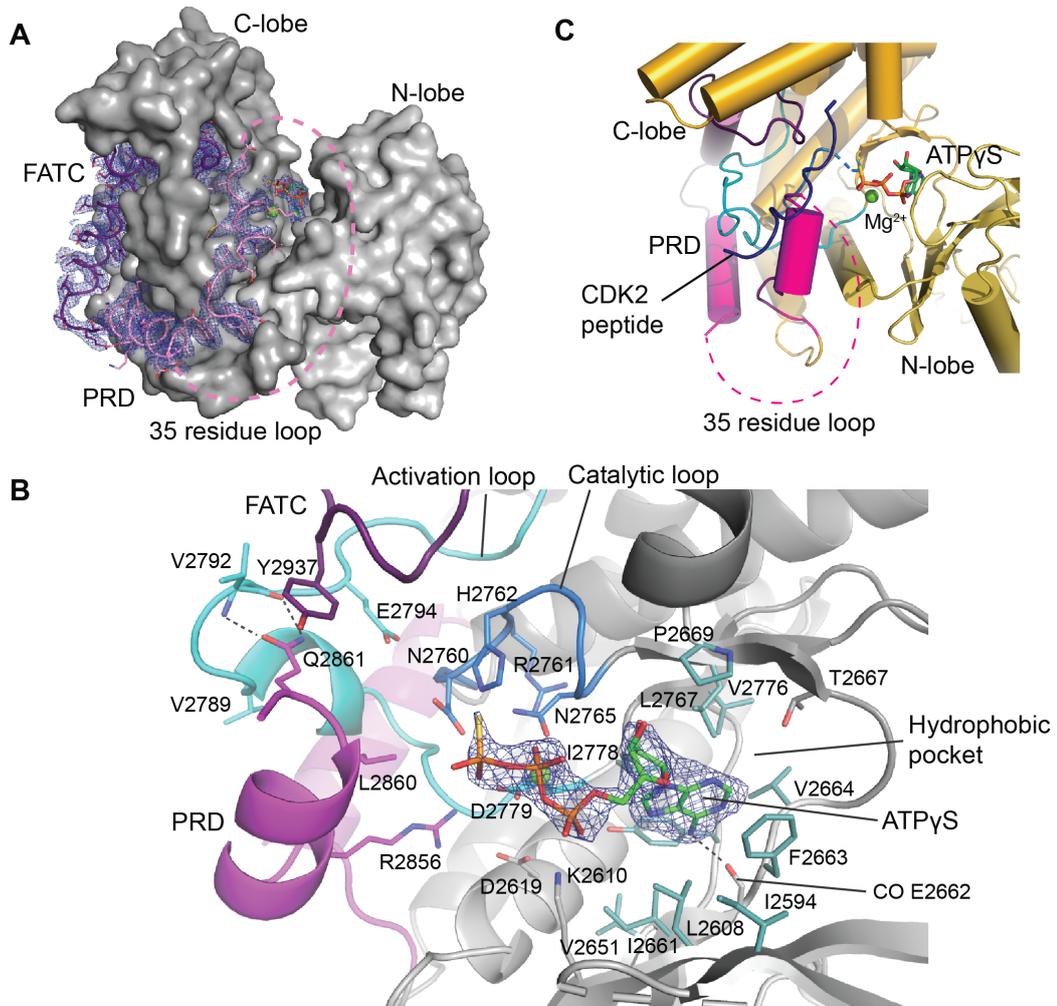
The 2-hydroxyl group of the sugar moiety is coordinated by the backbone carbonyl of His2764 through a hydrogen bond. The triphosphates are coordinated by a Mg<sup>2+</sup> ion, which is bound either directly to the side chains of Asn2765 and Asp2779 which are in the activation loop, or indirectly through a water molecule, for which we do not see any clear density. Lys2610 in the C-lobe, which is highly conserved, also can make a direct contact to the  $\alpha$ -phosphate of ATP $\gamma$ S, which reflects its role in binding ATP and catalysis. As the catalytic loop is well structured through interactions with the activation loop and FATC, the key catalytic residue Asp2760 is in the appropriate position for orienting any bound peptide substrate for nucleophilic attack. His2762 could also interact with the hydroxyl group of the substrate to stabilise the transition state.

The active site is partially capped by a lid formed by the P-loop (2585-IASGVSAPK-2593), typically a conserved feature in kinases, but which is not glycine-rich in ctTel1. Another highly conserved element of the ctTel1 kinase domain is the PRD, which consists of two  $\alpha$ -helices connected by a long flexible loop (2863-2896) of which the sequence is not highly conserved, apart from a number of negatively charged residues that may interact with the coiled-coil of TRD3.

Due to its flexibility, we could not build this loop, but there was density for the majority of the helices making up the PRD, except for the residues that connect directly to the flexible loop. The helices of the PRD directly interact with the activation and catalytic loops, providing a way by which the PRD can regulate the kinase activity. Tyr2848 and Trp2850 in the PRD contact Phe2784 and Met2786 in the activation loop through hydrophobic interactions. Additionally, a salt bridge is formed by Glu2785 in the activation loop and Lys2909 in the PRD. The PRD is stabilised by the insertion of its Arg2856 into a negatively charged pocket consisting of Asp2619, Asp2760 and Asp2779. The positively charged nature of this interaction is highly conserved in ATM orthologs, as some possess a lysine instead of an arginine. At the C-terminal end of the first PRD helix, Leu2860 interacts with a hydrophobic pocket in the active site. Gln2861 stacks with Tyr2937 in the FATC domain and forms a polar interaction with the backbone of Val2792 in the activation loop.

Due to its position, the PRD appears to regulate ctTel1 activity by sterically blocking access to the active site. Intriguingly, Leu2860-Gln2861 are located closely (less than 10 Å) to the P $\gamma$ of

ATP $\gamma$ S. As ATM phosphorylates proteins with an S/TQ motif, Gln2861 might (partially) block the substrate specificity pocket, while Leu2860 itself cannot be phosphorylated and prohibits catalysis. The PRD therefore may function as a pseudo-substrate.



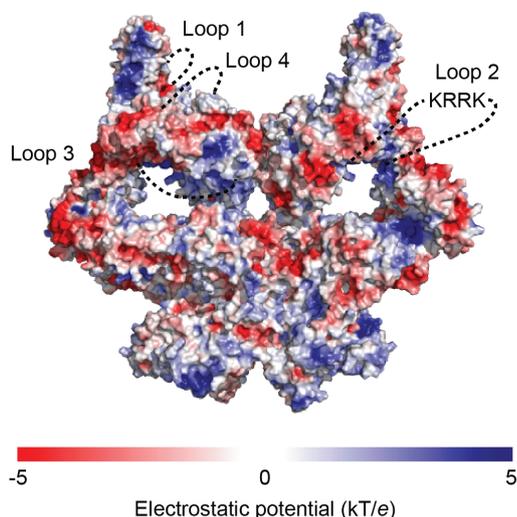
**Figure 3.14: Overview of the kinase domain and the active site** **A.** Overview of the kinase domain. The N-lobe is coloured light-grey, the C-lobe dark-grey, the PRD and its flexible loop pink and the FATC purple. The density for the PRD and FATC domains is displayed in deep blue. The active site lies in a cleft formed by the N- and C-lobes. Access to the active site is blocked by the N-terminal helix of the PRD. **B.** Detailed view of the active site. Density is shown for ATP $\gamma$ S and Mg<sup>2+</sup>. The catalytic loop is coloured blue and the activation loop cyan. The residues forming the hydrophobic pocket are depicted in light teal. All other elements are coloured as in A. **C.** The active site of ctTel1 modelled with a peptide substrate of CDK2 (PDB: 3QWH) depicted in navy blue. The active sites of the two kinases were superimposed. The substrate peptide would clash with the PRD of ctTel1, suggesting that the PRD inactivates ctTel1 by blocking substrate access through steric hindrance.

In an earlier structural study done on human ATM, the crystal structure of CDK2 with a peptide substrate (PDB: 3QWH) was overlaid with apo-hsATM. The PRD clearly clashed with the peptide substrate, suggesting that the PRD needs to move away as a part of the activation mechanism. We made an overlay of CDK2 with ctTel1 with ATP $\gamma$ S bound and observed that also in the nucleotide-bound state the PRD of ctTel1 clashes with the substrate. This indicates that nucleotide binding alone is not sufficient for ATM activation (Figure 3.14C).

Based on the structure of the active site, ctTel1 displays all the hallmarks of a catalytically proficient kinase. The loops required for catalysis are structured and the nucleotide is also correctly oriented. However, the PRD forms a pseudosubstrate blocking the active site. In order to activate ctTel1, the PRD blockage needs to be lifted, probably due to conformational changes induced by activators such as MRN or DNA.

### 3.4.7 Potential DNA binding sites

Several studies have shown that ATM has DNA binding properties and that its kinase activity is stimulated by DNA. As we confirmed that ctTel1(1-1864) indeed can bind to DNA with a high affinity, we set out to map the potential DNA binding sites in the  $\alpha$ -solenoid. The full-length structure, including the Spiral-Pincer domains which have DNA binding properties, enabled us to calculate the electrostatic potential at the solvent-accessible surface using PyMOL and identify the location of any positively-charged patches that could contribute to DNA binding.



**Figure 3.15: Surface map of the electrostatic potential of ctTel1.** The electrostatic potential at the solvent-accessible surface reveals a number of positively charged patches, which may be involved in DNA binding. Locations of the unstructured loops are indicated with a dotted line. Loop 2 contains a highly conserved KRRK motif, which may also play a role in DNA binding.

CtTel1 indeed has a number of positively charged patches. One is located close to the N-terminal TAN motif, which would due to its role in DNA repair be a logical site for DNA binding. The inside of the groove formed by the Spiral, Pincer and FAT domain contains two potential patches, one located halfway the Spiral domain, close to the site where the two protomers touch each other in the closed conformation. Another patch lies close to the hinge formed by the Pincer around which the Spiral domain pivots.

Several long loops could not be built due to their inherent flexibility. Therefore, we analysed a multiple sequence alignment for any conserved positively-charged patches in these regions. Interestingly, there is a highly conserved KRRK motif in the large loop insertion between H7 and H8. Due to the high degree of conservation it is tempting to assume it is playing a role in DNA binding.

Taken together, ctTel1 has several positively-charged patches that could account for its DNA binding properties. DNA binding could also be mediated by multiple patches simultaneously, e.g. the TAN motif and the conserved KRRK motif in a flexible loop, which are in close proximity. If multiple patches across the protein surface are required for DNA binding, this may explain why ctTel1 binds longer dsDNA with higher affinity.

### 3.4.8 Mapping of disease mutations

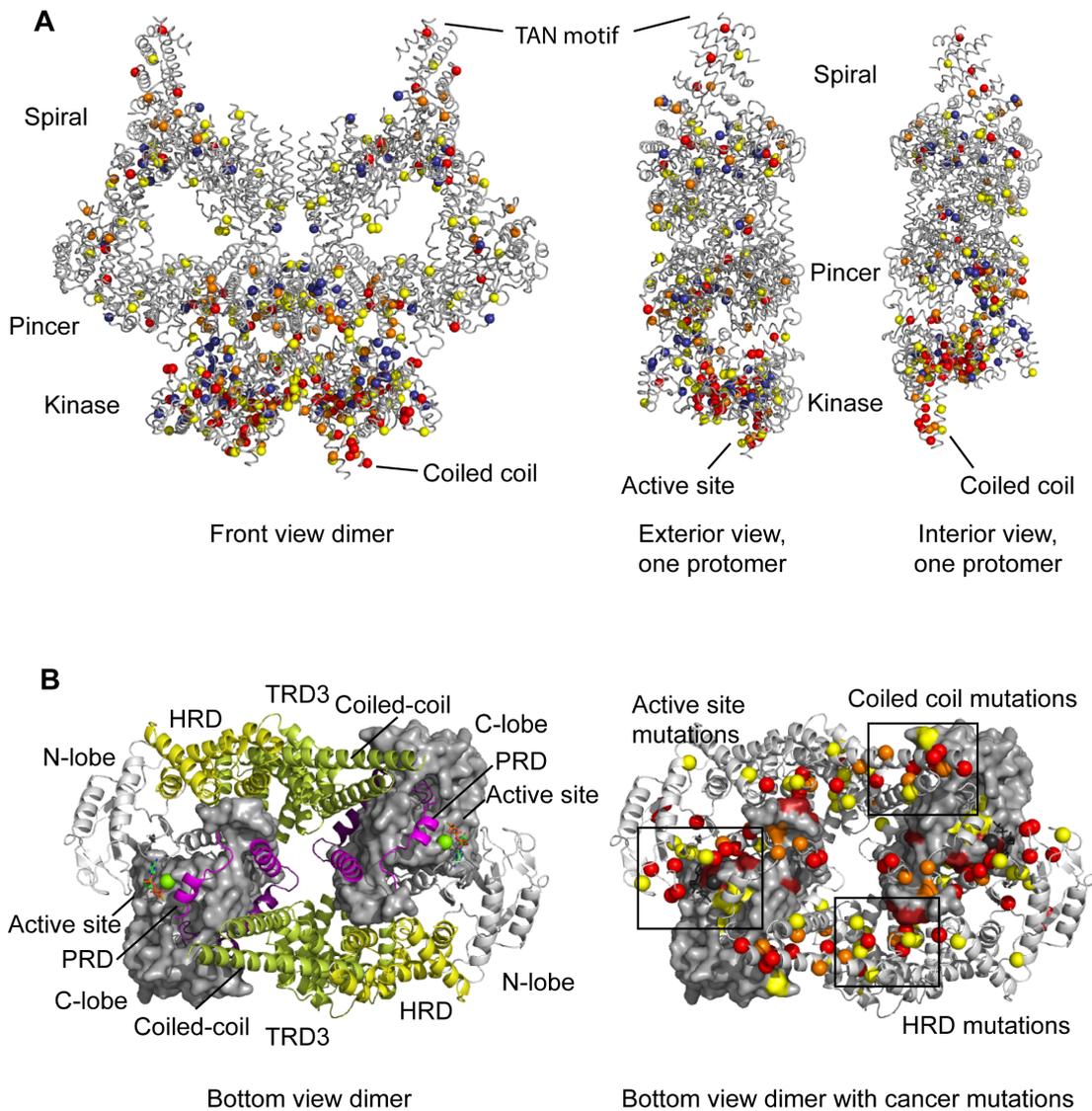
Our structure is the first (near-)complete atomic model of any ATM kinase to date. The high sequence similarity between ctTel1 and hsATM, especially in the FATKIN domain, made it possible to map somatic cancer mutations from the COSMIC database and germline A-T mutations, as they are listed in the UniProt database.<sup>232,233</sup> Included in the sequence alignment (see appendix) are cancer mutation sites that were reported at least three times and A-T mutations.

Mutations, both cancer mutations and A-T mutations, occur throughout the entire protein, but there are higher clusters consisting of more frequent mutations in the kinase domain and the surrounding FAT domain. Many mutations are found in the active sites and its direct surroundings. It is evident that any mutations here could have a detrimental effect on the catalysis, hampering DNA damage signalling. Unsurprisingly, the catalytic residues mentioned in the previous section are indeed frequently mutated.

Another cluster of mutations is found in the PRD. Arg3008 in hsATM (Arg2901) has been reported to be mutated in more than 53 cases. Although this residue does not interact with anything in the resting state of hsATM/ctTel1, it may be the case that this residue is important in an active conformation or in binding to substrates. This observation again reflects the importance of the PRD in the activation of ATM kinases.

Two other domains that contain clusters of cancer mutations are the HRD and the coiled-coils of TRD3. The interface between HRD and the kinase domain is frequently mutated, indicating that this domain also plays a role in regulation of kinase activity. The coiled-coils contain two clusters

of mutations. As it is likely that the coiled-coil directly interacts with the PRD and other parts of the kinase domain, it is likely that mutations in this area also have a direct effect on ATM activity. The coiled-coils have been suggested to play a critical role in the transition from an inactive to an active state (Baretic 2017). As the HRD is also directly interacting with the coiled-coils, it is possible that further conformational changes in the coiled-coil are mediated through the HRD to the kinase domain.



**Figure 3.16: Mapping of A-T and cancer mutations.** **A.** Overview of germline A-T mutations, as reported on UniProt, and somatic cancer mutations, as reported in the COSMIC database. The positions of the mutations are depicted as spheres. A-T mutations are depicted in navy blue. Cancer mutations that occur three times in the COSMIC database are depicted in yellow, four times in orange and five times or more in red. **B.** Two bottom views of the ctTell dimer, with and without cancer mutations. The left panel shows the C-lobe as a surface, with the other domains depicted as cartoons. The colours scheme is the same as in previous figures. In the right panel, the same colouring as in A. is used.

In the Spiral, of which the sequence is less conserved, mutations occur less frequently than in the FATKIN domain, although it contains a number of relatively frequently mutated residues. In the very N-terminus, where the TAN motif is located, which plays a role in telomerase maintenance and DNA damage signalling. Arg23 in hsATM (Arg29 in ctTel1, which is also highly conserved amongst other orthologs) within this motif is indeed frequently mutated, reflecting its role in the signalling function of ATM. Further mutations found in the Spiral could also impair stability of the protein. For example, Arg377 in hsATM (77 times reported), Gln400 in ctTel1) is found at the connection of one of the plug loops and H7, where it appears to play a stabilising role in the packing of the HEAT repeats. Any mutation here might destabilise the Spiral. The unstructured loop connecting H15 to H16 also contains several frequent cancer mutations sites, including Phe858 in hsATM, which is reported 19 times. As this loop is located on the top of the Spiral, in close proximity to the N-terminus, which is known to play a role in telomerase maintenance and DNA damage signalling, it could be that mutations in this loop interfere negatively with regulating proteins or DNA.

The A-T mutation sites follow a similar distribution pattern as the cancer mutations, and frequently overlap. Interestingly, one cluster A-T of mutations is found, and only one frequent (16 times reported, Arg2722) cancer mutation in two helices and a loop in the C-lobe of the kinase facing the FAT domain. As this area surrounds the FATC domain, mutations here might also destabilise the active site indirectly by destabilising the FATC and the loops that are structured by this element.

In conclusion, most mutations appear to be in places that either destabilise the structure, impair activation or directly interfere with catalysis. Residues that are directly interacting with the kinase domain and elements of the active site, such as the coiled-coil of TRD3, presumably play a larger role in ATM activation.

# Chapter 4

## Discussion

### 4.1 The recombinant expression of ctTel1

Soon after the discovery of the ATM gene in the disease A-T, it became clear that ATM is a highly important protein in our cells, as ATM orchestrates the cellular response towards DNA damage. This explains why A-T patients are prone to developing tumours and why ATM is also frequently mutated in many forms of cancer.<sup>89,131</sup>

Despite its importance, the structure and many aspects of the biochemistry of ATM remained elusive for many years. One of the main reasons for this knowledge gap is that it is difficult to obtain ATM in large quantities due to its enormous size. Although systems to express human ATM have been described, ATM and its yeast homologs such as spTel1 and scTel1 can only be obtained by tagging or (over)expressing it in its native host, e.g. hsATM can be expressed in HEK cells, whereas tagged scTel1 can be purified from *S. cerevisiae*. Still the quantities were not sufficient for crystallography.<sup>112,177,225,226</sup>

In recent years *C. thermophilum* became popular, especially amongst structural biologists, as the increased thermostability of its proteins makes them more suited for structural analysis. Therefore, previous work in the lab focussed on obtaining a more stable ATM kinase ortholog from *C. thermophilum* in larger quantities for crystallography. As ATM is highly conserved amongst all eukaryotes, an ATM ortholog was identified in *C. thermophilum* based on homology of the FATKIN domain, which was then named ctTel1 (work by C. Linke-Winnebeck).

Unfortunately, it was not possible to express ctTel1 in any recombinant system tested, including *E. coli*, *S. cerevisiae*, insect cells and mammalian cell lines, usually yielding nothing or at most insoluble protein. In *E. coli*, however, ctTel1 was degraded to distinct bands, which were analysed by LC-MS/MS finger printing. The largest fragment consisted of the first 35 HEAT-repeats of ctTel1 but lacked the kinase domain. This construct (ctTel1(1-1864)) could be overexpressed in

*E. coli* in large quantities. Due to the large amounts this construct could be used for biochemical assays and crystallisation trials. Crystallography, however, was unsuccessful. In order to stabilise the flexible HEAT-repeats, antibodies and nanobodies were produced against ctTel1(1-1864). Using the nanobodies, it was still not possible to obtain crystals, but it enabled purification of the endogenous full-length ctTel1. However, this required very large amounts of ctTel1 culture for small protein yields, and furthermore it was not possible to create mutants. Although cloning methods for *C. thermophilum* have been described, the cloning is inefficient and the protein yields are low.

It was then attempted to express ctTel1 with a codon-optimised kinase domain, as codon optimisation can improve the protein quality and the expression levels. This was again unsuccessful, as ctTel1 was degraded immediately. This raised again the question why full-length ctTel1 is resistant to recombinant expression while the entire HEAT repeat domain can be easily expressed in large quantities.

One explanation is that PIKKs depend on a dedicated chaperone system for proper folding and maintenance. Cdc37 and Hsp90 are required for the folding of regular, canonical kinases. PIKKs do not depend on Cdc37, but instead depend on the chaperone Hsp90 and the Tel2-Tti1-Tti2 (TTT) complex. TTT and Hsp90 do not interact directly with each other, and details on the interaction between the TTT complex and PIKKs also remain to be elucidated, although there is evidence that its Tel2 subunit directly interacts with the HEAT repeat domains of PIKKs.<sup>234-236</sup> The interaction between TTT and Hsp90 depends on another protein complex, R2TP, which consists of AAA-ATPase Rvb1 and Rvb2, and the proteins Tah1 and Pih1.<sup>234</sup> This complex, together with Hsp90, interacts with newly synthesized proteins to assemble larger protein complexes, such as for example RNA polymerase II.<sup>237</sup> R2TP does not interact with PIKKs directly, but indirectly via the TTT complex.<sup>234,238</sup> Recently, the R2TP complex was renamed PAQosome (Particle Arrangement of Quaternary structure), as it has been found that this complex, in combination with other factors such as prefoldins, plays a central role in folding and organising the proteome in complexes for many different clients, including PIKKs. For PIKKs it seems to be the case that there are different maturation pathways per family member.<sup>234,239</sup>

It has been proposed that the folding of a large solenoid might depend on chaperones to fold helical repeats that are distant from each other at the primary sequence.<sup>240</sup> The folding of the HEAT repeats appeared not to be the problem, however, as these were overexpressed in large quantities in *E. coli*, despite its large size (210 kDa). Furthermore the protein interacts with DNA and MRN, suggesting it is correctly folded. This suggests that the problem rather lies in the intricate fold of the FATKIN domain.

As is now evident from our structure and the other structures published, the FATKIN domain has a complex fold. The FAT domain wraps around the kinase domain, interacting via a number of highly conserved salt bridges. In order to keep the FAT domain open to accommodate the kinase domain, it is clear that chaperones are required. Furthermore, in ATM there is a large hydrophobic interface (TRD2), which forms the core of dimer interface. Chaperones might play a role here in avoiding

aggregation. The kinase domain itself also has a complex fold, consisting out of several structured loops that are stabilised by the FATC at the very C-terminus, which itself is highly hydrophobic and folds back into the C-lobe. Furthermore, there is a large unstructured insert in the PRD, and several elements in the kinase domain interact with the TRD3 of the opposing protomer.

It is likely that the TTT complexes and the PAQosome, despite being conserved elements amongst eukaryotes, in the recombinant expression systems tested are too different from their orthologs in *Chaetomium* that they cannot fold ctTel1 properly. Perhaps ctTel1 too different from the endogenous ATM found in these expression systems to be recognised by these folding proteins. Tel2, for example, indeed has low sequence homology.<sup>234,240</sup> Although it would have been interesting to coexpress these chaperones with ctTel1, it was not possible to find orthologs for all proteins in *C. thermophilum*. In the future, further characterisation of such biogenesis pathways may be interesting for proteins which are difficult to express.

## 4.2 Binding activities of ctTel1: interaction with DNA and MRN

The large quantities of ctTel1(1-1864) enabled us to perform biochemical and biophysical assays were are not possible endogenous ctTel1, such as fluorescence anisotropy and gel filtration assays. As HEAT repeats are frequently implied in protein-protein and protein-DNA interactions, ctTel1(1-1864) is a suitable construct for these assays. In DNA-PK, for example, where Ku70/80 and DNA directly interact with the HEAT-repeats of DNA-PKcs.<sup>241</sup>

Using fluorescence anisotropy, we observed that ctTel1(1-1864) has DNA binding properties and displays a preference for longer, double stranded, blunt-ended DNA. We also observed that the truncated protein can directly interact with an Nbs1 peptide, corresponding to the last C-terminal residues, and with ctMRN( $\Delta$ CC), a construct that lacks the coiled-coils of ctRad50. The addition of ATP did not appear to make a difference for the interaction with ctTel1(1-1864). Furthermore, we observed that ctMR( $\Delta$ CC) does not interact with ctTel1(1-1864)

### 4.2.1 Interaction with MRN

An early link between ATM and MRN was discovered when it was found that in ATLD Mre11 is mutated. Patients with ATLD have symptoms that are similar to those observed in A-T patients, suggesting that Mre11 and ATM are linked. It was later discovered that MRN senses DSBs and recruits ATM to the lesion, explaining this observation.<sup>89,134</sup>

Several studies have shown that the final C-terminal residues of Nbs1 (or Xrs2 in yeast) are responsible for this interactions.<sup>60,88</sup> In early work, colocalization of ATM and Nbs1 at DSBs was observed.<sup>242</sup> Subsequent studies further narrowed down the binding domain of Nbs1 to the very

C-terminal residues. Using a 2-hybrid system in yeast, it was reported again that scTel1 associates with DSBs via Xrs2, and that this binding is mainly mediated via the Xrs2 C-terminus.<sup>137</sup>

Two studies in 2005 further confirmed Nbs1-ATM interaction. In a study performed on human ATM and Nbs1, it was demonstrated that the last 20 C-terminal residues of hsNbs1 directly interact with hsATM, and that this C-terminus is needed for recruitment of ATM by MRN to the DNA.<sup>60</sup> To map the binding of Nbs1 to ATM, nuclear pull-downs with GST-ATM fragments were performed. Nbs1 binding mapped to residues 248-522 and 1436-1770 in hsATM, although the GST-fragment in the latter case seemed very impure. With the structural knowledge we currently have, this either corresponds to a region at the bottom of the flexible N-terminus, at the top of the Spiral, or to a region close to the hinge formed by the Pincer.

Another study with *S. cerevisiae* two-hybrid assays and *Xenopus* egg extracts defined an FxF/Y motif, preceded by an acid patch, in the carboxyl tail, which is highly conserved and required for the interaction. It was found that the interaction with Xrs2 is essential for recruitment of ATM to sites of DNA damage.<sup>137</sup> The last 50 residues are required, both in *Xenopus* and in yeast. For the two-hybrid analysis to map the binding site of Nbs1 on ATM, the 49 predicted HEAT repeats were divided into smaller parts. Repeats 17-18 and 21-22 interacted strongly with Nbs1, which led to the conclusion that Nbs1 has two ATM binding sites.<sup>88</sup> In our structure, these repeats are located around the flexible hinge of ctTel1, suggesting that Nbs1 binding might modulate the conformation in this region.

Based on sequence alignment, a homolog of Nbs1, ctNbs1, was identified in *C. thermophilum*, which also has the characteristic highly conserved FxF/Y motif and preceding acidic patch (SEEESED-DELKFRFRGRRR), similar to the motifs found in scXrs2 (GDGDDDDDDGPKFTFKRRKG) and in hsNbs1 (KEESLADDLFRYNPYLKRRR). Pull-down experiments performed with GST-ctNbs1 fragments indeed showed that the C-terminus of ctNbs1 binds to ctTel1(1-1864) (work by C. Linke-Winnebeck).

To measure the affinity of this interaction, fluorescence anisotropy measurements were performed using a 6-FAM labelled ctNbs1 peptide. Under the experiment conditions the  $K_D = 764 \text{ nM} \pm 219 \text{ nM}$ , suggesting a medium-strength binding. The same peptide, but without a label, was used to prepare cryo-EM grids in order to map the Nbs1 binding site. Although there appeared to be some additional density in the HEAT repeats on top of the Spiral, the density was too weak to unambiguously fit the peptide sequence, or see which residues are exactly involved in the binding. Neither were any conformational changes observed (data not shown). Alternatively, the peptide could interact with the hinge region, as earlier studies suggested. Here the resolution is worse due to the higher degree of flexibility, making it impossible to find a small peptide.

A stable interaction with ATM for recruitment to the DSB may require the entire MRN complex. Several studies have indeed found that, although the localisation to DSBs seems to be mainly dependent on Nbs1 interaction.<sup>242,243</sup> MR is required for activation, especially with Rad50 in an

ATP-bound state, even in the absence of Nbs1.<sup>112,244</sup> Several studies have suggested that MR directly interacts with ATM. Using assays in *S. pombe* it was shown that MR still can bind to overexpressed Tel1 in absence of Nbs1, suggesting there must be some low-affinity interaction between MR and Tel1.<sup>178</sup> Similar findings were reported for purified human MRN and ATM, showing that MR also binds ATM directly. It was found that this interaction is dependent on Rad50, not on Mre11.<sup>112</sup> In another recent study, pull-downs were performed with the individual components of the yeast MRN complex and Tel1. It was found that Tel1 interacts with all subunits, but in contrast to other studies, the interaction with Mre11 is stronger than with Rad50. In the same study, MRN and different combinations of heterodimeric complex, such as MR, RN, and MN were tested in their ability to stimulate Tel1 activity. Surprisingly, MN did not stimulate Tel1 activity while MR and RN did, suggesting not only that the activation might depend on a direct interaction between Rad50 and Tel1.<sup>111</sup>

To test the interaction with MRN, we made use of a ctMRN construct in which the coiled-coils of Rad50 had been truncated (ctMRN( $\Delta$ CC)). At the time the experiments were performed, it was not possible to work with intact, full-length ctTel1, as the complex obtained was not stoichiometric, with ctRad50 binding strongly to chromatographic resins. The sample without coiled-coils was biochemically better behaved, and therefore it was used for initial studies.

In pull-down experiments performed with MBP-ctTel(1-1864), it was observed that ctTel1(1-1864) interacts with ctMRN( $\Delta$ CC), but not with ctMR( $\Delta$ CC) which suggests that the interaction with MRN mainly depends on Nbs1. In gel filtration assays it was observed that ctTel(1-1864) and ctMRN( $\Delta$ CC) coelute as a complex, as indicated by a significant shift in the peak on the chromatogram, suggesting a strong interaction. When ATP was added when mixing the proteins and to the running buffer, no differences were observed. As in ctMRN( $\Delta$ CC) ctNbs1 is underrepresented, there was a second peak corresponding to ctMR( $\Delta$ CC). It was unclear whether this peak coelutes with ctTel1(1-1864) because also separately they elute at the same volume.

Although the Nbs1-binding is in accordance with earlier work demonstrating that Nbs1 interacts with ATM, other studies have shown that there are also direct interactions between MR and ATM. Although it is possible that we did not observe direct interactions between MR and ATM because of low-affinity binding, another explanation would be the lack of the Rad50 coiled-coils. The coiled-coils of Rad50 are essential for the proper functioning of the MRN complex, including its role DNA damage signalling through ATM.<sup>129</sup> Mutations and truncations in the coiled-coil domain, and the Zn<sup>2+</sup>-hook abolish the functions of the MRN complex.<sup>100–102,129,245</sup>

Furthermore, several studies have also reported that the ATP-bound state of Rad50, but not ATP hydrolysis, is required to activate ATM and maintain its binding to the DSB, both in yeast and in human, suggesting that the structural features of MR are also important for the activation mechanism.<sup>100,111,112,129,244,246</sup> Several structures of MR have been reported showing indeed large conformation changes based on ATP-binding state.<sup>94,116</sup> The ATP-bound state of Rad50 can be maintained for a relatively long time as the rate of hydrolysis is low.<sup>108,246,247</sup> However, it is likely

that also a proper ATP-bound state depends on the presence of intact coiled-coils. This may explain why addition of ATP had no effect on the chromatograms obtained by gel filtration.

Furthermore, the entire FATKIN domain was missing in ctTel1(1-1864). Any potential interactions between MRN and ATM in this region would also be missed when using truncated constructs. Although the finding that ctTel1 interacts with the C-terminus of ctNbs1 is in accordance with literature, intact proteins should be used for future biochemical and structural characterisation in order to find explanations for the observations reported in other studies.

#### 4.2.2 Interaction with DNA

Several studies over the years have reported that ATM has DNA binding properties. This is not surprising as MRN recruits ATM directly to the lesion. Any DNA binding properties of ATM may, together with MRN, stabilise its binding to DNA. Therefore, the DNA binding properties of ctTel1(1-1864) were investigated using fluorescence anisotropy. Work previously done in our lab and experiments in this thesis indeed show that ctTel1(1-1864) has DNA binding properties, and displays a preference for double-stranded blunt-ended DNA (work by C. Linke-Winnebeck). In our fluorescence anisotropy studies, we found a length-dependency, with a ds60mer binding more strongly ( $K_D = 50$  nM) than a ds30mer, a ds25mer or a ds20mer.

DNA binding properties were observed before, and interestingly, most studies also report a length-dependency, with ATM preferring longer DNA for binding and kinase stimulation.

An early study with purified hsATM showed that ATM activity is stimulated by DNA. ATM also was shown to prefer to bind to DNA ends using AFM, although a third of the population of ATM molecules binds internally to the DNA. The binding appeared not to be sequence specific, but was in accordance what we found highly length dependent.<sup>182</sup> In another study on hsATM it was found that both MRN and DNA are required for ATM activation. Again, the DNA sequence did not seem to matter, but maximum stimulation was only observed in the presence of 1-2 kb DNA.<sup>112</sup>

Several studies on the DNA binding properties of ATM were performed using *Xenopus* extracts. One report claimed that ATM can be activated by DNA breaks alone in the absence of MRN, in large excess of DNA.<sup>180</sup> In another study in which more physiological amounts of DNA added to *Xenopus* extract, a length dependency was observed, with a minimal length required of 200 bp for ATM activation. Another finding was that full ATM activation requires binding of ATM to the regions flanking the DNA ends, but not to the ends themselves (You 2007)

More recently, it was demonstrated for purified yeast Tel1 that both MRN and DNA can independently stimulate Tel1 activity, and together they have a strong synergistic effect. For minimal activation, a ds80mer was required, but the optimal DNA substrate is 2 kb linearised, nucleosome-free plasmid DNA. In line with this finding, there is no requirement for double-stranded ends, although this is in contrast with the strict requirement for ends of hsATM.<sup>111,112</sup> DNA stimulatory

effect *in vitro* were also reported for hsATM, also showing more stimulation in the presence of longer double stranded DNA.<sup>189</sup>

The length-dependency observed in other studies involves much longer DNAs than we tested. Partially, the length of DNA we tested limited due to the requirements of fluorescence anisotropy (if the distance between the binding site and the fluorophore is so long that the fluorophore can move independently, it is not possible to measure anisotropy). Other studies also reported a preference for double-stranded DNA, which we also observed. Curiously, some studies reported that ATM prefers to bind to internal regions on the DNA and not to ends.<sup>111,182</sup> Depending on the mode in which ctTel1 binds DNA, a ds60mer might be too short to allow for internal binding.

Based on our atomic model, we calculated a surface charge map (PyMOL), which indeed reveals a number of positively charged patches that might play a role in DNA binding. Interestingly, the unstructured loop between H7 and H8 contains a highly conserved RKKR motif, which is a promising site for potential DNA binding. Another candidate motif for DNA binding is the TAN motif, which lies exposed at the N-terminus. The TAN motif is also conserved amongst ATMs in all eukaryotes, and is involved in telomere maintenance and DNA damage signalling.<sup>230</sup> Arg29 in ctTel1 is universally conserved and its equivalent in hsATM is frequently mutated in cancer. Additionally, there are a number of positively charged patches on the  $\alpha$ -solenoid, which also may contribute to DNA binding.

One explanation for the length dependency we observe is that multiple patches might be involved in stable DNA binding. If this is the case, it is more likely that a longer oligonucleotide binds with all the patches, than a short oligonucleotide. An interesting future experiment would be to mutate these patches and test whether these mutations interfere with DNA binding and DNA-dependent stimulation of kinase activity.

This would however still not fully explain the length dependency on DNA longer than 200bp or more. Additionally, it was observed that scTel1 requires nucleosome-free DNA for optimal activation.<sup>111</sup> The physiological meaning of this is unclear, as DSBs and ATM activation leads to condensation of the chromatin to silence transcription.<sup>248</sup>, while other studies report a depletion of nucleosomes in HR.<sup>249</sup> The dependence on dsDNA is more explicable, as ATM "sees" dsDNA when it recruited. However, during resection, ssDNA is formed, which is coated by RPA. This stimulates ATR instead of ATM, which might lead to a downregulation of ATM activity and enhance ATR activity for later steps of HR, providing a way to modulate ATM activity. Alternatively, longer DNA fragments could allow more ATM molecules to bind, which would lead also lead to an increase in stimulation of kinase activity.<sup>42,89</sup>

Taken together, our studies on the interaction with MRN and DNA show that ctTel1 can bind to both, which is in accordance with literature. For a better understanding of the system, however, the assays need to be repeated with full-length proteins.

## 4.3 Structural biology

Although ATM has been studied for over twenty years now, many aspects of its biochemistry are not well understood. This is largely due to lack of structural data. The identification of ctTel1(1-1864) and the generation of nanobodies against this construct enabled the purification of full-length endogenous ctTel1 from *C. thermophilum*. Although the yields of these purifications are small, it was sufficient for structure determination by cryo-EM.

To date, our ctTel1 structural model is the first ATM structure of which an atomic model could be built for the entire molecule, except for a number of flexible loops.

### 4.3.1 Comparison to other ATM kinase structures

Over the last years, numerous ATM structures have been solved, to ever increasing resolutions. Initial low-resolution structures revealed that ATM has a dimeric structures with a striking butterfly-like architecture, in which the dimerisation is mediated by the FATKIN domain, but not by the N-terminal  $\alpha$ -solenoid.<sup>224–226</sup> The higher resolution structures that have been resolved more recently allow us to have a closer look at the molecular details of especially the FATKIN domain and the active site.<sup>188,189,208,227</sup>

Overall, the structure described in this thesis is in perfect agreement with other structures, due to high degree of conservation in especially the FATKIN domain. It is however the first structure to describe an atomic model for the  $\alpha$ -solenoid domain.

#### 4.3.1.1 Comparison $\alpha$ -solenoid

The presence of the nanobodies has been beneficial for stabilisation of the N-terminal Spiral domain. Different nanobodies were found to stabilise different parts of the N-terminus. Nanobody D4, for example, stabilises the HEAT repeat on the apical side of the Spiral, whereas nanobody F7 stabilises the very N-terminus. As the latter nanobody seemed to stabilise a larger part of ctTel1, it was also used for sample preparation for a large cryo-EM data acquisition. This enabled us, after sorting out the different conformations of the Spiral domain, to build nearly the entire polypeptide of the N-terminal Spiral, except for a number of unstructured loops.

Although the FATKIN domain appeared to be rather rigid in our structure and in other structures reported, the N-terminal  $\alpha$ -solenoid displayed more flexibility, with the Spiral moving from an open to a closed conformation around a hinge formed by the Pincer. Movements within the N-terminal HEAT-repeats forming an  $\alpha$ -solenoid domain have been described in the other ATM/Tel1 structures.<sup>208,227</sup> So far, ctTel1 is the only structure in which interactions between the Spiral domains

of opposing protomers have been described. In all other structures, the  $\alpha$ -solenoid was not resolved sufficiently to see similar interactions.

In the scTell1 structure, conformational changes are described to be coupled to rearrangements in the active site, especially in the PRD which blocks access to the active site.<sup>227</sup> As is described below in greater detail, the PRD blocks the access to the active site by functioning as a pseudosubstrate. Although we also see an open and closed conformation of the  $\alpha$ -solenoid, we do not see any coupled changes in the active site in our structure. Neither were similar motions observed in another recent scTell1 structure.<sup>188</sup>

In the first hsATM structure, two types of dimers are reported, one with symmetrical related FATKINs, which were referred to as closed dimers, and one with an asymmetric conformation, referred to as a more open conformation. The resolution of the open conformation, however is rather limited, to 11.5 Å. In this "open" conformation, however, the  $\alpha$ -solenoid seems to be more closed whereas the kinase domain is more open. In the asymmetric form a ratchet-like movement is observed between the two FATKINs, twisting one protomer by 30° with respect to the other, leading to a disordered PRD, suggesting a more active state.<sup>208</sup> Similar twisting motions, however have not been observed in our ctTell1 structure.

In a more recent scTell1 structure with AMP-PNP bound in the active site, motions were observed as well in the  $\alpha$ -solenoid, but these motions were not transferred in the active site, similarly to what we observe.<sup>188</sup> It is not clear whether the active site is more rigid due to the presence of a nucleotide, and whether the coupled conformational changes observed in the apo-scTell1 structure reflect a more active state that could be further stabilised by MRN or DNA. Both scTell1 and hsATM have been demonstrated to have a basal activity. However, it is likely that for robust ATM activity, activators such as MRN are needed to induce more robust conformational changes. ATM and Tell1 are indeed more active in the presence of DNA and MRN.<sup>111,112,208</sup>

#### 4.3.1.2 FATKIN domain

The FAT domain is highly conserved amongst all PIKKS and amongst ATM kinases, they are almost identical. The salt bridges that connect the FAT domain to the kinase domain are conserved from yeast to human, and are frequently mutated in cancer, suggesting that these residues are critical for the stability of ATM.

An important feature of the FAT domain of ATM is that it forms the dimer interface. ATM is reported to form monomers as part of its activation mechanism.<sup>46</sup> However, in all higher resolution scTell1 structures and our ctTell1 structure the TRD2-TRD2 interface, which forms the core of the dimerisation interface, is very hydrophobic.<sup>188,227</sup> Monomerisation would be highly energetically unfavourable, suggesting that it is unlikely that this is part of its activation mechanism. The other interfaces are predominantly hydrophilic, with extensive contacts existing between the kinase

domain and TRD3, especially between its coiled-coil and the C-lobe. Based on this, it seems more likely that the activation mechanism of ATM involves motions around a hydrophobic which could release the hydrophilic interfaces and open the active site. Such a turning motion has been described for hsATM, where one protomer turns, yielding an asymmetric dimer.

Surprisingly, a 7.8 Å structure of a hsATM monomer has been reported, which was shown to be more active, as the restraints imposed by the presence of the other protomer have been lifted, opening up the kinase domain. The monomers were suggested to be induced by the exposure to DNA during lysis and by leaving out DNase out of the lysis buffer. However, in our ctTel1 preparations, no DNase was added and no monomers were observed. Thus it is unclear whether the observed monomers resulted from activation.

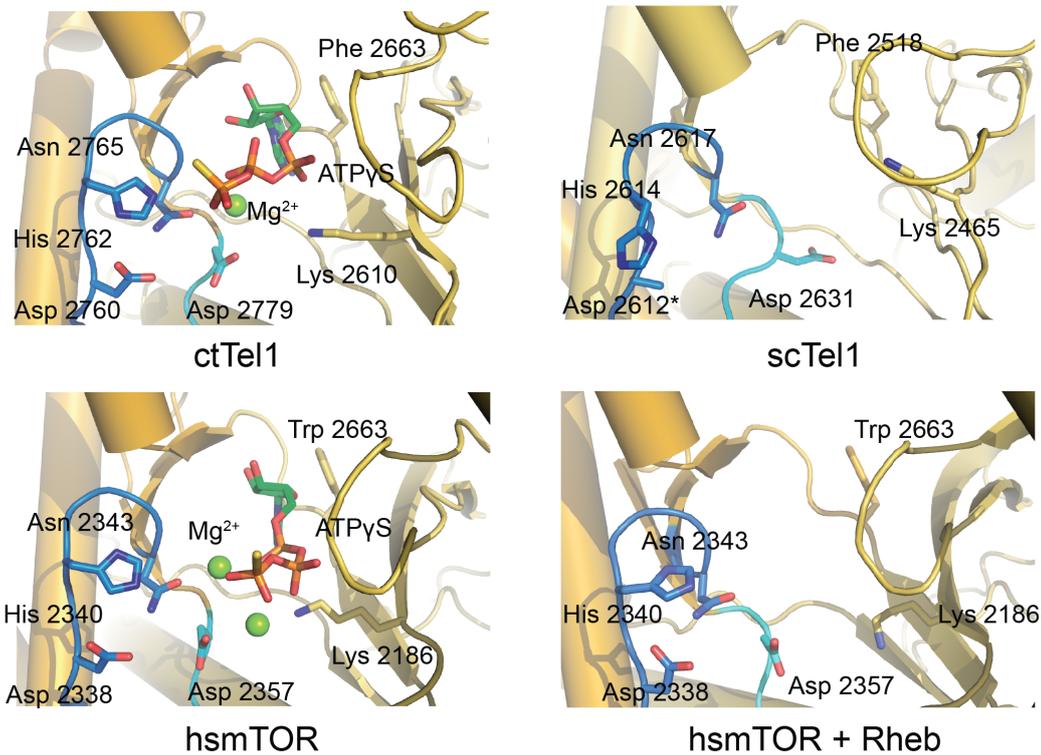
#### 4.3.1.3 Kinase domain

The kinase domain of the different ATMs is extremely conserved, both in sequence and in structure. All elements that are characteristic for PIKKs are present, such as the catalytic and activation loop, a hydrophobic FATC domain and the PRD, which block substrate access to the active site.

In the active site, there are minor differences between the apo-state and the nucleotide bound state. In all PIKKs the catalytic elements, such as the catalytic loop and the activation loop, are structured through interactions with the FATC domain. Therefore, the residues required for catalysis are also in a correct position. In the nucleotide-bound state, they are oriented such that they can coordinate the ATP analogs. Similar differences have been observed for mTOR as well when comparing a nucleotide-bound state to an apo-state (Figure 4.1).

There are conflicting studies on the number of Mg<sup>2+</sup> ions in the active site. For ctTel1, density has only been observed for one Mg<sup>2+</sup>, whereas for mTOR two Mg<sup>2+</sup> have been observed, at rather different positions in the active. For scTel1 there was faint density for a second Mg<sup>2+</sup> at a similar position as has been described for mTOR, suggesting that ATMs might have a second Mg<sup>2+</sup> as well. Other studies on CDK2 have suggested that the number of Mg<sup>2+</sup> ions depends on the stage of the catalysis, with a second Mg<sup>2+</sup> ion binding as an activating Mg<sup>2+</sup> together with a substrate.<sup>149,188,250</sup>

The PRD, which is conserved amongst all PIKKs, functions in ATM as a pseudosubstrate which blocks the entrance to the active site by sterically hindering peptide binding. The two helices that make up the PRD (also frequently denoted as Kα9 and Kα10) are highly conserved amongst ATMs, whereas the linker loop (often denoted as the PRD-insert) connecting the helices is not conserved in length and amino acid composition. In ctTel1, there is an LQ motif, which reminiscent of the S/TQ phosphorylation motif, in close proximity to the γ-phosphate. It is tempting to speculate that the Q binds to a substrate binding pocket, thereby anchoring the PRD. However, the LQ motif is not conserved among all ATMs, with scTel1 being an exception, suggesting the LQ is not essential for PRD function.



**Figure 4.1: Comparison of the active sites of ctTel1, scTel1 and hsmTOR.** Nucleotide-bound ctTel1: 6SKY, nucleotide-free scTel1: 6JXC, hsmTOR: 4JSV and nucleotide-free hsmTOR in complex with Rheb: 6BCU. The active sites ctTel1, scTel1 and hsmTOR are highly conserved. Binding of ATP $\gamma$ S induces the same orientation of the active site residues to coordinate the nucleotide. The Rheb-bound mTORCs was reported to activate mTORC1 by closing the P-loop covering the active site more. However, the residues are not oriented accordingly and the P-loop is in the same conformation as in the ATP $\gamma$ S-bound state.

In several structural studies on ATM kinases, the PRD was reported to become more disordered, either in response to movements in the  $\alpha$ -solenoid, ratchet movements giving rise to asymmetric dimers, or monomerisation. However, many of the maps these structures were based on were at rather low resolution, in which it is difficult to see whether the helices are disordered or whether the density in this region is poor due to low resolution.

The PRD has been described to be more ordered in the nucleotide bound state, when compared to the apo-state.<sup>188,227</sup> As has been described before, the PRD makes several contacts with the activation loop and other parts of the kinase domain. Although it does not appear to make any direct connections to the ATP analogue, the presence of ATP puts restraints on the active site, and could thus stabilise the PRD indirectly via the activation loop.

The PRD flexible insert is believed to interact directly with the coiled-coil of TRD3, and other elements in the active site, such as FATC and the activation loop also interact with TRD3. Any

shift of TRD3 away from the kinase domain would release the binding to these elements, which could then become less ordered. While the FATC lies embedded in the C-lobe and the activation loop, the PRD has fewer interactions with the kinase domain. This could in principle enable it to become more disordered. The PRD is more ordered in the nucleotide bound state, when compared to the apo-state.<sup>188,227</sup> As has been described before, the PRD makes several contacts with the activation loop and other parts of the kinase domain. Although it does not appear to make any direct connections to the ATP analogue, the presence of ATP restrains the active site, and could stabilise the PRD indirectly via the activation loop.<sup>251</sup>

### 4.3.2 Comparison to other PIKKs

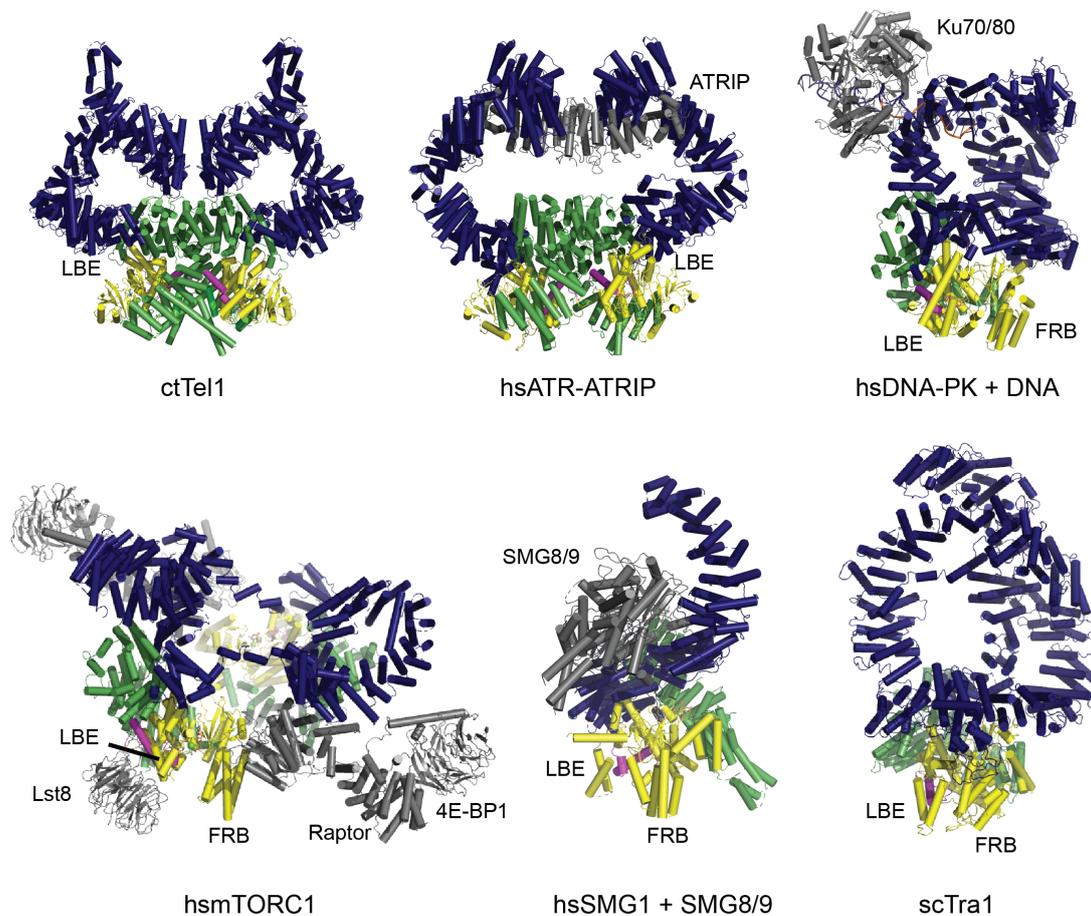
To date, no structures have been reported of ATM kinases in complex with activators such as MRN or DNA. Comparing ATM to other PIKKs might shed more light on the activation mechanism and binding mode of modulators. Structures of all PIKKs have been reported now, most of which have been solved by cryo-EM. In addition, of several PIKKs higher order complexes have been reported.

Although the  $\alpha$ -solenoids are highly divergent in sequence and structure, all have a similar function as a integration hub for activators and modulators. The large variation in these HEAT-repeat domains also reflects the diverse roles that PIKKs play in the cell. In contrast, the FATKIN domain and especially the kinase domain are highly conserved. Any differences in these domains also reflect diverging substrate specificities. In this section some structures of PIKKs that have recently been solved are compared to the ctTel1 structure in order to gain more insight in common themes (Figure 4.2 and 4.3).

#### 4.3.2.1 Comparison to ATR

ATR/Mec1 is also involved in DNA damage signalling. It is the PIKK which is most similar to ATM in terms of sequence similarity and architecture. ATR also forms a dimer, and the  $\alpha$ -solenoid folds in a comparable manner. Together with ATRIP (Ddc2 in yeast) it forms one structural unit, that dimerises in a similar fashion as ATM. Structures have been solved for both ATR-ATRIP and Mec1-Ddc2 at 4.7 Å and 3.9 Å respectively.<sup>252,253</sup>

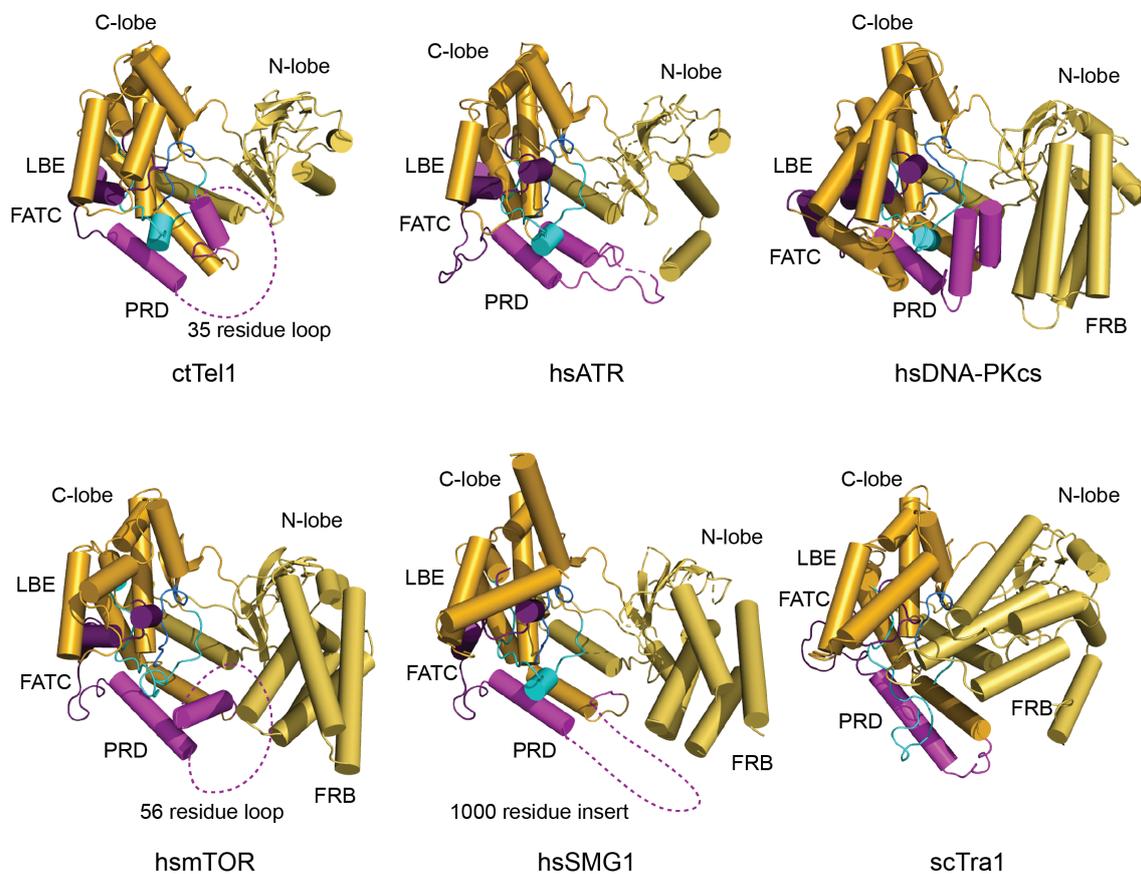
Whereas in ATM the dimer interface is dominated by interactions between the FATKIN domains of neighbouring protomers, the dimer interface of ATR/Mec1 is more distributed along the structure, and ATRIP/Ddc2 also plays a role in dimerisation at the  $\alpha$ -solenoid close to the N-terminus. However, the exact nature of the ATRIP/Ddc2 dimer is unclear as the different structures reported are contradictory. A crystal structure has been reported that suggests that ATRIP forms a coiled-coil, but such a coiled-coil is not observed in the ATR-ATRIP structure. The limited resolution of the ATRIP/Ddc2 densities did not allow for modelling the entire ATRIP molecule.



**Figure 4.2: Overview of the structures of all PIKKs.** hsATR:ATRIP: 5YZ0, hsDNA-PK+DNA: 5Y3R, hsmTORC1: 6BCX, hsSMG1 + SMG8/9: 6SYT, scTra1: 5OJS. The domains are coloured according to the same scheme in the introduction. The HEAT repeat domain is navy blue, the FAT domain green, the kinase domain yellow and the FATC domain magenta. Any additional binding proteins are depicted in grey. The  $\alpha$ -solenoid domains are both in terms of structure and sequence the most divergent domain, whereas the FAT and especially the kinase domains show more similarities.

The protomers are reported to have distinct conformations due to inherent flexibility. Similar asymmetry we observe for ctTel1 too in the  $\alpha$ -solenoid domain. The residues involved in dimerisation are highly conserved. Interestingly, the TRD2-TRD2 interface of ATR is also hydrophobic, comparable to ATM, although in contrast to ATM, ATR has not been reported to monomerise, despite having a similar architecture.<sup>252,253</sup>

The kinase domain of ATR is highly similar to the kinase domain of ATM. Both ATM and ATR lack the FRB four-helix bundle located close the entrance to the active site in all other PIKKs. The FRB is suggested to play a role in the regulation of mTOR in which is was discovered as the binding site of the mTOR inhibitor rapamycin.<sup>149</sup> It might have a similar regulatory role in other PIKKs in



**Figure 4.3: Overview of the kinase domains of all PIKKs.** hsATR: 5YZ0, hsDNA-PKcs: 5Y3R, hsmTORC1: 6BCX, hsSMG1: 6SYT, scTra1: 5OJS. The domains are coloured according to the same scheme as in results section xyz. The kinase domains are highly conserved in terms of architecture, although there are a number of striking differences. ATM does not have the FRB four helix bundle. Furthermore, the PRD, which functions as a pseudosubstrate in ATM, and perhaps also in DNA-PKcs, does not block the active site cleft in all PIKKs, as is evident in hsATR. In hsmTOR and hsSMG1 the unstructured insert in the PRD may play a larger role in regulation the activity than the helices. The activation loop of scTra1, which is longer than in other PIKKs, lies on the PRD, which is not blocking an active site. Instead, it is blocked by the FRB and LBE, reflecting its loss of kinase activity.

limiting their activity. It is therefore also tempting to speculate that the substrate promiscuity of ATM and ATR might in part be due to the lack of this domain. The active sites of ATM and ATR are also highly conserved, showing the same architecture and orientation of the catalytic residues. In the structure of Mec1, however, part of the activation loop is not visible due to a lack of density. The catalytic residues are also positioned in a way that is not optimal for catalysis. Whether this reflects an inactive state is not clear, however, it was suggested that an in/out transition of the catalytic DFG motif could form an alternative activation mechanism.<sup>251</sup>

The PRD in ATR is shorter than the PRD in ctTel1, with a loop just long enough to connect the  $\alpha$ -helices. When comparing the kinase domains, it is obvious that the PRD of ATR does not block the active site access as it does in ATM, and the activation loop is exposed. In ATM, on the other hand, the activation loop is encircled by the PRD. There is however no steric hindrance when modelling a Chk1 peptide onto ATR. While the PRD of ATR does not regulate activity by blocking the active site, it has been shown to be the interaction site of TopBP1, which has been proposed to activate ATR directly via the activation loop.<sup>147,252,253</sup>

#### 4.3.2.2 Comparison to DNA-PKcs

DNA-PKcs is the other PIKK involved DNA damage signalling and mainly plays a role in responding to DSBs that are repaired by NHEJ. Together with the Ku70/80 dimer it forms DNA-PK. Of DNA-PKcs and DNA-PK several structures have been reported in the last decade. Astonishingly, two crystal structures have been solved, an apo-crystal structure of 6.6 Å, and a structure with the C-terminal tail of Ku80 bound at 4.3 Å. Furthermore, cryo-EM structures have been reported of DNA-PKcs and C-terminal tail of Ku80 and a structure of a DNA-PK complex including the Ku70/Ku80 dimer and DNA.<sup>153,241,254,255</sup>

DNA-PKcs folds into an enormous, concave ring-like structure, with the FATKIN domain forming a head that sits on top of the circular  $\alpha$ -solenoid domain. The fold of the HEAT repeats is rather different than when comparing it to ATM kinases, but the FAT and kinase domain are highly conserved.

Comparable to ATM, of which the Nbs1 C-terminus has been mapped to the HEAT repeats, the C-terminus of Ku80 was also mapped to bind in an equivalent region in DNA-PKcs.<sup>60</sup> Two structures have been reported with only the C-terminal tail of Ku80 present. Curiously, these structures are contradictory, as the crystal structure shows segments of the C-terminal tail spread over the HEAT repeat domain.<sup>254</sup>, whereas an EM structure does not show these segments, but a rather weak density (9 Å) sitting on the  $\alpha$ -solenoid in a flexible region close the FATKIN. The latter study fits better with existing NMR data, according to which the Ku80 C-terminus form globular core of six helices. Interestingly, for ATM the Nbs1 binding sites have been mapped as well to the flexible hinge region which is relatively close the FATKIN domain.<sup>255</sup>

A later study resolved the structure of complete DNA-PK, with Ku70, Ku80 and DNA. The binding between DNA-PKcs and Ku70/80 is weak in the absence of DNA, and vice versa the affinity for DNA is also weak in the absence of Ku70/80. Similar as has been described for ATM and MRN, binding of Ku70/80 stimulates the kinase more strongly in presence of DNA. The actual DNA interaction site, which is located in the  $\alpha$ -solenoid, is very limited and unspecific, allowing for a wide range of different DNA sequences as substrates. This is another example of how PIKKs are allosterically regulated and activated. It is conceivable that there is a similar mode of binding for ATM, with DNA interaction mediated by MRN, but also involving direct binding to ATM.

For ATM, the location of the DNA binding site is not known, but DNA binding properties have been demonstrated. On our ctTel1 map, there are multiple patches that could be involved in DNA binding, but as in DNA-PKcs, they do not involve many residues.<sup>241</sup>

The DNA duplex in DNA-PK lies in flexible ridge between two HEAT repeat segments. ATM also has such ridges where the  $\alpha$ -solenoid takes a sharp turn, such as the hinge which Nbs1 binding has been mapped. It could be that such interactions, whether with DNA or regulatory proteins, stabilise a conformation in the HEAT repeats which are subsequently allosterically transferred to the kinase domain.

The FATKIN and active site are highly conserved. In DNA-PKcs a conserved metal is observed in the active site. Comparable to mTOR there is an FRB like domain, which is not present in ATM. Where in ATM the PRD consists of two helices connected by a flexible loop, the PRD in DNA-PKcs forms a helical hairpin of about 30 residues occluding direct access to the active site as it stacks between the catalytic loop and the FRB. Allosteric binding of Ku70/80 and DNA has been described to induces an inward motion of the HEAT repeats, which in the kinase domain leads to a slightly larger entry groove for the substrate. The PRD does not appear to move, however, leaving unclear what changes DNA-PK is requires to become activated instead.<sup>241,254</sup>

#### 4.3.2.3 Comparison to mTOR

In terms of structural insight into the activation mechanisms of PIKKs, the different structures that have been solved of (m)TOR and the different complexes in which it exists provide the most complete picture. In 2013, a crystal structure was reported of the FATKIN domain of human mTOR in complex with LST8, and also in complex with ATP analogs and inhibitors. This was the first high resolution structure of any PIKK. A few years later structures were reported that included the  $\alpha$ -solenoid, also in complex with its regulators Raptor, Rictor and the activator Rheb.<sup>149,256-260</sup>

In contrast to ATM, mTOR does not dimerise via its FATKIN domain, but via its  $\alpha$ -solenoid, forming a symmetric dimer, with the FATKIN domains and LST8 at opposing ends of the structure. In the middle of the structure is a large cavity. The  $\alpha$ -solenoid forms a large spiral-like structure, where regulating proteins such as Raptor, Rictor and Rheb bind.<sup>256,259-261</sup>

The FATKIN of mTOR is so highly conserved that it fits into low resolution ATM maps, reflecting the degree of structural conservation among PIKKs. The main differences are the absence of a comparable coiled-coil in TRD3, while mTOR has a prominent FRB domain, which forms a four-helix bundle in the N-lobe. The FRB of mTOR is the rapamycin-binding location, which inhibits its activity. On the other side of the substrate access cleft LST8 is bound to the LBE (LST8 Binding Element) in the C-lobe. All PIKKs including ATM have an LBE domain, but they do not bind to LST8, and neither is the role of LBE in the regulation of kinase activity clear. In the ctTel1

structure however, the LBE interacts directly with the coiled-coils of TRD3 of the other protomer, suggesting that the LBE could play a role in the activation of ATM.<sup>149</sup>

Between the FRB and the LBE with LST8 lies the mTOR PRD, which has a long insertion of 56 residues. As for ATM, the mTOR PRD also plays an important role in regulating the activity of mTOR, potentially by blocking access to the active site. Mutations in the PRD have indeed been identified to make mTOR more active. Although the  $\alpha$ -helix of mTOR is not in exactly the same position as the  $\alpha$ -helix of ctTel1, it might be that the insert blocks substrate access.<sup>149</sup>

The active site is highly conserved, which is particularly striking when comparing the ATP $\gamma$ S bound state to our ctTel1 structure. The mTOR active site contains two Mg<sup>2+</sup> ions. While in the ctTel1 structure the Mg<sup>2+</sup> ion is in a different position, we did not observe convincing density for second Mg<sup>2+</sup> ion (Figure 4.1).

A cryo-EM structure of mTOR in complex with its activator Rheb provides the most complete picture so far of PIKK activation. Binding of Rheb to the mTOR  $\alpha$ -solenoid induces conformational changes that cause the  $\alpha$ -solenoid to move inwards to form new interactions between the N-terminal part of the  $\alpha$ -solenoid and the FAT domain. These new interactions then lead to conformational changes in the FAT domain, which occur mainly at the flexible hinges. A large, 30 ° angle, twist between TRD1 and TRD2 in FAT domain was described to lead to further closure of the active site by the N-lobe and a rearrangement of catalytic residues to optimal position the ATP. The PRD helices do not appear to move significantly, however, but they also do not form a pseudosubstrate as in ctTel1 or other ATM kinases.<sup>261</sup>

Conformational changes induced by MRN and DNA in the ATM  $\alpha$ -solenoid could have comparable effects on the FAT domain and the kinase domain. Especially the turn of the FAT domain at the TRD2 residue in mTOR is intriguing, as a similar ratchet movement has been suggested for hsATM, which would leave the hydrophobic TRD2-TRD2 interface intact. Such movements would break the contacts between TRD3 and the kinase domain, opening the active site without requiring energetically unfavourable monomerisation.<sup>251,261</sup>

#### 4.3.2.4 Comparison to SMG1

Whereas ATM and ATR are involved in many phosphorylation events, SMG1 only phosphorylates UPF1, which is a key event in NMD. It forms a complex with SMG8 and SMG9 which regulate its kinase activity. Very recently several structures have been reported of human SMG1 in complex with SMG8 and 9, and a structure of apo-SMG1, of which the coordinates have not been released at the moment of writing this thesis.<sup>262,263</sup>

As seen in all other PIKKs, the  $\alpha$ -solenoid of SMG1 forms also a Spiral like structure which functions as a hub for interactions with its regulators SMG8 and 9 and also the FATKIN is highly conserved.

Interestingly, SMG1 has a longer  $\alpha$ -helical bundle in the TRD3 at the same location as ATM kinases have their coiled-coil insertion, although its function within a SMG1 monomer is not clear.

Like mTOR there is an FRB-like domain in the N-lobe of the kinase. The active site is also similar due to its high level of conservation. Unexpectedly, the PRD of SMG1 contains an extremely large, unstructured insert, spanning over a thousand residues, and end in the FATC domain, which is structured as in all other PIKKs. The function of this insert is not clear but presumably also plays a regulatory function, comparable to the other PRDs.

From the comparison of the apo-SMG1 structures and the SMG1-complex the main differences in the  $\alpha$ -solenoid are in the N-terminal HEAT repeats, which adopt a more curved conformation upon SMG8/9 binding. SMG8/9 inhibits SMG1 kinase activity, which seems to be via an allosteric mechanism, but it is from the available structures not clear how.

Another unusual feature of SMG1 is the presence of InsP<sub>6</sub> (inositol hexaphosphate) in a conserved positively charged pocket, where it seems to have a role in stabilising a cavity between the FAT domain and the kinase domain. It was suggested that an unaccounted density in mTOR might also correspond to InsP<sub>6</sub>, suggesting this is common for more PIKKs. In ctTel1, however, there is a positively charged cavity present as well, but no density for InsP<sub>6</sub> was observed, even at the high resolutions we obtained. The positively charged residues lining the cavity also do not point inwards, suggesting furthermore nothing is bound. Whether the presence of InsP<sub>6</sub> is only a feature of some PIKKs, or whether it depends on the organism, the purification method or the activation state remains to be investigated.<sup>262</sup>

#### 4.3.2.5 Comparison to Tra1

Tra1/TRAPP is an interesting member of the PIKK family as it is the only pseudokinase. Tra1 functions as a scaffold in numerous complexes, such as SAGA and NuA4, which are involved in the regulation of transcription.<sup>156,157</sup>

In terms of its general architecture, yeast Tra1 is strikingly similar to DNA-PKcs. This is even more curious, as DNA-PKcs, an active kinase, only exist in higher eukaryotes. Several structures have been described of Tra1, both of the apo-protein and of the protein within higher order complexes. Tra1 functions entirely as a scaffold for different proteins. From mapping mutations and structural studies of higher order complexes, Tra1 has multiple non-competing binding sites for different proteins contributing to its role as an integration hub. This may be true for the  $\alpha$ -solenoids of other PIKKs as well. Of mTOR it has been well established it functions in the context of different complexes.<sup>146,156,157</sup> This is comparable to the many modulators of ATM activity that have been described, some of which might bind simultaneously to ATM to fine-tune its kinase activity.<sup>89</sup>

Despite being a pseudokinase, Tra1 still retains all of the characteristics of PIKKs, albeit it with modifications that reflect its inactivity. The active site within the kinase domain is conserved,

but the catalytic DFG motif is not present anymore. The activation loop contains an 18-residue insertion not present in other PIKKs, which lies on the PRD on the outside of the kinase domain. Furthermore, the FRB four helix bundle contacts the LBE in the C-lobe, occluding the substrate binding site completely. Tra1 also contains a conserved FATC domain, in which mutations are often found to interfere with Tra1 functioning. As Tra1 is not an active kinase, it is likely that the FATC is not important for kinase activity or substrate recognition per se, but mainly fulfils a role in maintaining structural integrity of the kinase domain.<sup>156,157</sup>

### 4.3.3 Summary

In recent years, the improvements in cryo-EM have allowed for the unravelling of structures of all PIKKs. Although their  $\alpha$ -solenoid domains are highly divergent in architecture, their FATKIN domains are conserved. For all PIKKs, the  $\alpha$ -solenoid functions as a hub for regulating proteins. Allosteric binding of these regulators can induce changes in the  $\alpha$ -solenoid, which are mediated through the FAT domain toward the kinase domain, as has been described for mTORC1 in complex with Rheb. Motions in the FAT domain enable further closing of the N-lobe and rearrangement of the catalytic residues for optimal ATP binding.<sup>261</sup>

The kinase domains of all PIKKs are highly conserved, with only Tra1 diverging more due to the fact it is catalytically dead.<sup>156</sup> ATM and ATR do not contain the FRB insertion in the N-lobe, which might enable them to bind to an enormous variety of potential substrates. ATM contains an LBE, which in mTOR interacts with the regulatory protein LST8, and in ATM interacts with the coiled-coil of TRD3, suggesting that the LBE in ATM also has a regulatory function. The PRD, an element which is present in all PIKKs, differs significantly in structure and position. In ATM it functions as a pseudosubstrate, while in other kinases its regulatory role is less clearly defined. All these differences are likely to be a reflection of the different roles and corresponding substrate specificities the PIKKs have in the cell.

## 4.4 A model for ATM activation

The comparison between the different ATM kinase structures and other PIKK structures published to date allow us to speculate on the activation mechanism of ctTel1. Furthermore, the mapping of the somatic cancer and germline A-T mutations also sheds light on which elements are critical for ATM/Tel1 activity.

In earlier work on human ATM it was reported that its activation mechanism involves phosphorylation of Ser1981 and subsequent monomerization, which releases the autoinhibited state.<sup>46</sup> Both concepts are controversial, as the autophosphorylation site is not conserved and not required in mouse ATM.<sup>186,187</sup> Furthermore, this serine appears to be located in a flexible loop on the outside

of the molecule, and is not located closely to the dimerisation interface.<sup>188,208</sup> Monomerization is unlikely due to the large hydrophobic dimer interface, as is observed in our ctTel1 structure and in a recent scTel1 structure.<sup>188</sup>

Our structure reveals the autoinhibitory circuitry of ctTel1, which shows that the active site is in an optimal conformation for catalysis, but is rendered inactive by the PRD, which blocks the active site as a pseudosubstrate. The PRD directly interacts with the activation loop, and also potentially interacts with the coiled coil in the FAT domain of the other protomer via its unstructured insertion. Any conformational changes within the FAT domain, especially in TRD3, could directly lead to movements in the kinase domain that would open the active site.

As was described for mTORC1, the binding of Rheb to the  $\alpha$ -solenoid induces a conformational change which is translated through large twists in the FAT domain to multiple rearrangements in the active site. The N-lobe and P-loop close and catalytic residues are positioned more optimally to bind the ATP. It is conceivable that similar motions are induced by the binding of MRN and DNA. Although we and others observed motions in the  $\alpha$ -solenoid, these were not unambiguously coupled to changes in the active site. This suggests that the movements that are induced by MRN and DNA must be larger. However, in apo-scTel1, motions in the  $\alpha$ -solenoid were coupled to a more disordered PRD. In a study on human ATM a ratchet movement within the FAT domain was observed, which also led to a more disordered PRD. Such findings suggest that MRN and ATM could alternatively stabilise or induce such motions as part of the activation mechanism.<sup>227,261</sup>

The ratchet movement described for human ATM is reminiscent of the twisting motion the mTORC1 FAT induced by Rheb. Any conformational changes in the FAT domain have a direct effect on the coiled coil, which is directly connected to the PRD. Indeed, most disease mutations are also clustered in the active site, the coiled coil and the region between the coiled coil and the active site. Where the mutations in the active site are detrimental for catalysis, mutations in the other regions potentially interfere with the activation process.<sup>208,261</sup>

When a double strand break occurs, more factors that are associated with the break modulate ATM activity. One of the first steps in the activation of ATM is the acetylation of Lys3016 in hsATM (which is Lys2909) by Tip60/Kat50. This lysine is located in the PRD, where it forms a salt bridge with Glu2785 in the activation loop (in *C. thermophilum*). This salt bridge would be broken upon acetylation, thereby contributing to the release of the PRD and increasing the levels of ATM activity.<sup>89,264,265</sup>

To summarise, ATM activation depends on MRN and DNA, which are likely induce conformational changes via the FAT domain through allosteric binding to the  $\alpha$ -solenoid, which positions the active site residues in an optimal conformation for catalysis and releases the PRD pseudosubstrate to enable phosphorylation of a broad range of substrates.

## 4.5 Conclusion and outlook

In this thesis a structure of ctTel1, a fungal ATM ortholog, is presented, providing the first complete atomic model of any ATM kinase to date, with only a small number of flexible loops missing.

The kinase domain was solved to a resolution of 2.8 Å and an overall structure including the HEAT repeat domain to 3.7 Å. The structure reveals the architecture of the  $\alpha$ -solenoid, which is involved in binding to MRN and DNA. Using a truncated ctTel1 construct, we could demonstrate that this domain interacts with MRN, and particularly via Nbs1, and with DNA.

Interestingly, ctTel1 displays a preference for longer DNA, binding a ds60mer with a higher affinity than shorter oligonucleotides. By calculating the electrostatic surface of ctTel1 and multiple sequence alignment we could indeed identify a number of conserved positively charged patches, which may be involved in DNA binding.

The high resolution of our structure allowed for analysis of the dimer interfaces. Interestingly, this revealed a hydrophobic dimer interface, suggesting monomerisation is unlikely to be an aspect of the activation mechanism, as this would be highly energetically unfavourable.

The kinase domain is extremely conserved when compared to other PIKKs. In the active site of the kinase the loops required for catalysis are structured, coordinating an ATP $\gamma$ S molecule and a Mg<sup>2+</sup> ion. This finding suggested that ATM is in a catalytic proficient state. However, access to the active site is physically blocked by the PRD, which appears to function as a pseudosubstrate. This also suggests that part of the activation mechanism entails its removal in response to allosteric binding of activators.

By mapping cancer mutations and A-T mutations, certain clusters of mutations appeared, suggesting that these are found in elements which are important for the activation. Apart from mutations in the kinase domain itself, where any mutations has a detrimental effect on catalysis, many mutations were also found in interface between the kinase domain and the FAT domain, especially in its coiled-coil, which directly interacts with conserved elements lining the active site of the opposing protomer.

The ctTel1 structure thus reveals the autoinhibitory circuitry and provides the first step towards understanding the activation mechanism of this important kinase. However, important questions remain: what conformational changes take place in ATM upon activation? How is the PRD released from the active site? Does it form a monomer, or would the FAT domain undergo a twist, thereby releasing the constraints on the kinase domain?

In order to understand the activation mechanism, it will be necessary to obtain structures of ctTel1, or another ATM ortholog, in complex with activators, such as DNA and MRN. The reconstitution of higher order complexes will require careful analysis of the biochemical requirements, such as the

ATP-state of Rad50, the presence of DNA and the nature of the oligonucleotide, and possibly even post-translational modifications, such as phosphorylation.

Another way to approach the question would be to generate mutants and test the activity of ATM. It would, for example, be highly interesting to mutate the positively charged patches, such as the TAN motif and the conserved KRRK motif, to confirm whether these are involved in DNA binding, and whether ATM would still be stimulated when these are mutated. Similarly, analysis of the residues in the coiled-coil and along the kinase interface could yield insight in which elements play a role in ATM activation.

To summarise, this work provides a structural framework for future studies on the activation mechanism. As ATM is one of the apical kinases in the DNA damage response, and also is frequently mutated in cancer, it will continue to be a highly important protein to be studied, both out of scientific interest as from a therapeutic perspective.

# Bibliography

- [1] Jan H J Hoeijmakers. DNA Damage, Aging, and Cancer. *The New England Journal of Medicine*, 361(15):1475–1485, 2009.
- [2] Thomas S Dexheimer. DNA Repair Pathways and Mechanisms. In L.A. Mathews, S.M. Cabarcas, and E.M. Hurt, editors, *DNA Repair of Cancer Stem Cells*, pages 19–32. Springer Netherlands, Dordrecht, 2013.
- [3] Jean Cadet, Markus Berger, and Jean-Luc Ravanat. Oxidative Damage to DNA: Formation, Measurement, and Biological Significance. *Reviews of Physiology, Biochemistry and Pharmacology*, 131:39–40, 1997.
- [4] Julian Stingele, Roberto Bellelli, and Simon J. Boulton. Mechanisms of DNA-protein crosslink repair. *Nature Reviews Molecular Cell Biology*, 18:563–573, 2017.
- [5] Tomas Lindahl and Barbro Nyberg. Rate of Depurination of Native Deoxyribonucleic Acid. *Biochemistry*, 11(19):3610–3618, 1972.
- [6] Scott D McCulloch and Thomas A Kunkel. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Research*, 18(1):148–161, 2008.
- [7] Jean-Luc Ravanat, Thierry Douki, and Jean Cadet. Direct and indirect effects of UV radiation on DNA and its components. *Journal of Photochemistry and Photobiology*, 63:88–102, 2001.
- [8] Stephen P Jackson and Jiri Bartek. The DNA-damage response in human biology and disease. *Nature*, 461(7267):1071–1078, 2009.
- [9] Josef Jiricny. The multifaceted mismatch-repair system. *Nature Reviews Molecular Cell Biology*, 7:335–346, 2006.
- [10] Sheila S David, Valerie L O Shea, and Sucharita Kundu. Base-excision repair of oxidative DNA damage. *Nature*, 447:941–950, 2007.
- [11] Hannes Lans, Jan H J Hoeijmakers, Wim Vermeulen, and Jurgen A Marteijn. The DNA damage response to transcription stress. *Nature Reviews Molecular Cell Biology*, 20:766–784, 2019.
- [12] Raphael Ceccaldi, Beatrice Rondinelli, and Alan D D’Andrea. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in Cell Biology*, 26(1):52–64, 2016.
- [13] Myron F Goodman and Roger Woodgate. Translesion DNA Polymerases. *Cold Spring Harbor Perspectives in Biology*, 5:1–20, 2013.

- [14] Isabel Lam and Scott Keeney. Mechanism and Regulation of Meiotic Recombination Initiation. *Cold Spring Harbor Perspectives in Biology*, 7:1–24, 2015.
- [15] Susumu Tonegawa. Somatic generation of antibody diversity. *Nature*, 302:575–581, 1983.
- [16] Haihui Lu, Klaus Schwarz, and Michael R Lieber. Extent to which hairpin opening by the Artemis: DNA-PKcs complex can contribute to junctional diversity in V(D)J recombination. *Nucleic Acids Research*, 35(20):6917–6923, 2007.
- [17] Eric J Gapud, Baek-Seung Lee, Grace K Mahowald, H Bassing, and Barry P Sleckman. Repair of Chromosomal RAG-Mediated DNA Breaks by Mutant RAG Proteins Lacking Phosphatidylinositol 3-Like Kinase Consensus Phosphorylation Sites. *The Journal of Immunology*, 187:1826–1834, 2011.
- [18] Zhenming Xu, Hong Zan, Egest J Pone, Thach Mai, and Paolo Casali. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nature Reviews Immunology*, 12:517–531, 2012.
- [19] James E Haber. Mating-Type Genes and MAT Switching in *Saccharomyces cerevisiae*. *Genetics*, 191(May):33–64, 2012.
- [20] Aisha H Syeda, Michelle Hawkins, and Peter McGlynn. Recombination and Replication. *Cold Spring Harbor Perspectives in Biology*, pages 1–14, 2014.
- [21] Vincenzo Costanzo, Kirsten Robertson, Marina Bibikova, Edward Kim, Domenico Grieco, Max Gottesman, Dana Carroll, and Jean Gautier. Mre11 Protein Complex Prevents Double-Strand Break Accumulation during Chromosomal DNA Replication. *Molecular Cell*, 8(1):137–147, 2001.
- [22] Andrei Kuzminov. DNA replication meets genetic exchange: Chromosomal damage and its repair by homologous recombination. *Proceedings of the National Academy of Sciences*, 98(15):8461–8468, 2001.
- [23] Michelle K Zeman and Karlene A Cimprich. Causes and consequences of replication stress. *Nature Cell Biology*, 16(1):2–9, 2014.
- [24] David Cortez. Preventing replication fork collapse to maintain genome integrity. *DNA Repair*, 32:149–157, 2015.
- [25] Gulam Waris and Haseeb Ahsan. Reactive oxygen species: role in the development of cancer and various chronic diseases. *Journal of Carcinogenesis*, 5(15):1–8, 2006.
- [26] J F Ward. The Complexity of DNA Damage : Relevance to Biological Consequences. *International Journal of Radiation Biology*, 66(5):427–432, 1994.
- [27] Anthony P Breen and John A Murphy. Reactions of oxyl radicals with DNA. *Free Radical Biology and Medicine*, 18(6):1033–1077, 1995.
- [28] Betsy M Sutherland, Paula V Bennett, Olga Sidorkina, and Jacques Laval. Clustered DNA damages induced in isolated DNA and in human cells by low doses of ionizing radiation. *Proceedings of the National Academy of Sciences*, 97(1):103–108, 2000.
- [29] David S. Goodsell. The Molecular Perspective: Ultraviolet Light and Pyrimidine Dimers. *Stem Cells*, 19(4):348–349, 2001.

- [30] Marila Cordeiro-Stone, Alexander M Makhov, Liubov S Zaritskaya, and Jack D Griffith. Analysis of DNA Replication Forks Encountering a Pyrimidine Dimer in the Template to the Leading Strand. *Journal of Molecular Biology*, 289:1207–1218, 1999.
- [31] Daniel A Koster, Komaraiah Palle, Elisa S M Bot, Mary-Ann Bjornsti, and Nynke H Dekker. Antitumour drugs impede DNA uncoiling by topoisomerase I. *Nature*, 448(July):213–217, 2007.
- [32] Jingyang Chen and JoAnne Stubbe. Bleomycins: Towards better therapeutics. *Nature Reviews Cancer*, 5:102–112, 2005.
- [33] Andrew J Wyrobek, Thomas E Schmid, and Francesco Marchetti. Relative Susceptibilities of Male Germ Cells to Genetic Defects Induced by Cancer Chemotherapies. *Journal of the National Cancer Institute*, 34:31–35, 2005.
- [34] Dong Wang and Stephen J Lippard. Cellular processing of platinum anticancer drugs. *Nature Reviews Drug Discovery*, 4:307–320, 2005.
- [35] Alberto Ciccia and Stephen J Elledge. The DNA Damage Response: Making it safe to play with knives. *Molecular Cell*, 40(2):179–204, 2010.
- [36] Larry H. Thompson. Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: The molecular choreography. *Mutation Research*, 751(2):158–246, 2012.
- [37] Ralph Scully, Arvind Panday, Rajula Elango, and Nicholas A Willis. DNA double-strand break pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology*, 20:698–714, 2019.
- [38] Andrés Aguilera and Belén Gómez-González. Genome instability: a mechanistic view of its causes and consequences. *Nature Reviews Genetics*, 9(March 2008):204–217, 2008.
- [39] Anuja Mehta and James E Haber. Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair. *Cold Spring Harbor Perspectives in Biology*, 6:1–17, 2014.
- [40] Wolf-Dietrich Heyer. Regulation of Recombination and Genomic Maintenance. *Cold Spring Harbor Perspectives in Biology*, 7:1–22, 2015.
- [41] Stephen C Kowalczykowski. An Overview of the Molecular Mechanisms of Recombinational DNA Repair. *Cold Spring Harbor Perspectives in Biology*, 7:1–36, 2015.
- [42] Andrew N Blackford and Stephen P Jackson. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Molecular Cell*, 66(6):801–817, 2017.
- [43] Travis H. Stracker and John H J Petrini. The MRE11 complex: Starting from the ends. *Nature Reviews Molecular Cell Biology*, 12:90–103, 2011.
- [44] Shuhei Matsuoka, Bryan A. Ballif, Agata Smogorzewska, E. Robert McDonald, Kristen E. Hurov, Ji Luo, Corey E. Bakalarski, Zhenming Zhao, Nicole Solimini, Yaniv Lerenthal, Yosef Shiloh, Steven P. Gygi, and Stephen J. Elledge. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, 316:1160–1166, 2007.
- [45] Yosef Shiloh and Yael Ziv. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nature Reviews Molecular Cell Biology*, 14(4):197–210, 2013.

- [46] Christopher J Bakkenist and Michael B Kastan. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature*, 421:499–506, 2003.
- [47] Alessandro A Sartori, Claudia Lukas, Julia Coates, Martin Mistrik, Shuang Fu, Jiri Bartek, Richard Baer, Jiri Lukas, and Stephen P Jackson. Human CtIP promotes DNA end resection. *Nature*, 450:509–514, 2007.
- [48] Samuel F Bunting, Elsa Callen, Nancy Wong, Hua-Tang Chen, Federica Polato, Amanda Gunn, Anne Bothmer, Samuel F Bunting, Elsa Calle, Niklas Feldhahn, Oscar Fernandez-Capetillo, Liu Cao, Xiaoling Xu, Chu-Xia Deng, Toren Finkel, Michel Nussenzweig, Jeremy M Stark, and André Nussenzweig. 53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks. *Cell*, 141:243–254, 2010.
- [49] Branden Wolner, Stephen Van Komen, Patrick Sung, and Craig L Peterson. Recruitment of the Recombinational Repair Machinery to a DNA Double-Strand Break in Yeast. *Molecular Cell*, 12:221–232, 2003.
- [50] Aura Carreira, Jovencio Hilario, Ichiro Amitani, Ronald J Baskin, Mahmud K K Shivji, Ashok R Venkitaraman, and Stephen C Kowalczykowski. The BRC Repeats of BRCA2 Modulate the DNA-Binding Selectivity of RAD51. *Cell*, 136(6):1032–1043, 2009.
- [51] Jie Liu, Ludovic Renault, Xavier Veaute, Francis Fabre, Henning Stahlberg, and Wolf-Dietrich Heyer. Rad51 paralogues Rad55–Rad57 balance the antirecombinase Srs2 in Rad51 filament formation. *Nature*, 479:245–248, 2011.
- [52] Patrick Sung and Hannah Klein. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nature Reviews Molecular Cell Biology*, 7:739–750, 2006.
- [53] Michael J McIlwraith, Alexandra Vaisman, Yilun Liu, Ellen Fanning, Roger Woodgate, and Stephen C West. Human DNA Polymerase  $\eta$  Promotes DNA Synthesis from Strand Invasion Intermediates of Homologous Recombination. *Molecular Cell*, 20:783–792, 2005.
- [54] William Douglass Wright, Shanaya Shital Shah, and Wolf-Dietrich Heyer. Homologous recombination and the repair of DNA double-strand breaks. *Journal of Biological Chemistry*, 293(27):10524–10535, 2018.
- [55] Haley D M Wyatt and Stephen C West. Holliday Junction Resolvases. *Cold Spring Harbor Perspectives in Biology*, 6:1–30, 2014.
- [56] Mireille Betermier, Pascale Bertrand, and Bernard S Lopez. Is Non-Homologous End-Joining Really an Inherently Error-Prone Process? *PLoS Genetics*, 10(1):1–9, 2014.
- [57] John R. Walker, Richard A. Corpina, and Jonathan Goldberg. Structure of the Ku heterodimer bound to dna and its implications for double-strand break repair. *Nature*, 412:607–614, 2001.
- [58] Pierre-Olivier Mari, Bogdan I Florea, Stephan P Persengiev, Nicole S Verkaik, Hennie T Bruggenwirth, Mauro Modesti, Giuseppina Giglia-Mari, Karel Bezstarosti, Jeroen A A Demmers, Theo M Luiders, Adriaan B Houtsmuller, and Dik C Van Gent. Dynamic assembly of end-joining complexes requires interaction between Ku70 / 80 and XRCC4. *Proceedings of the National Academy of Sciences*, 103(49):18597–18602, 2006.

- [59] Michael R Lieber. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annual Review of Biochemistry*, 79:181–211, 2010.
- [60] Jacob Falck, Julia Coates, and Stephen P Jackson. Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature*, 434:605–611, 2005.
- [61] Tanya M Gottlieb and Stephen P Jackson. The DNA-Dependent Protein Kinase: for DNA Ends and Association with Ku Antigen. *Cell*, 72:131–142, 1993.
- [62] Lisa G Defazio, Rachel M Stansel, Jack D Grif, and Gilbert Chu. Synapsis of DNA ends by DNA-dependent protein kinase. *The EMBO Journal*, 21(12), 2002.
- [63] Katheryn Meek, Pauline Douglas, Xiaoping Cui, Qi Ding, and Susan P Lees-Miller. Trans Auto-phosphorylation at DNA-Dependent Protein Kinase’s Two Major Autophosphorylation Site Clusters Facilitates End Processing but Not End Joining. *Molecular and Cellular Biology*, 27(10):3881–3890, 2007.
- [64] Naoya Uematsu, Eric Weterings, Ken-ichi Yano, Keiko Morotomi-yano, Burkhard Jakob, Gisela Taucher-scholz, Pierre-olivier Mari, Dik C Van Gent, Benjamin P C Chen, and David J Chen. Auto-phosphorylation of DNA-PKcs regulates its dynamics at DNA double-strand breaks. *Journal of Cell Biology*, 177(2):219–229, 2007.
- [65] Penny Jeggo and Peter O’Neill. The Greek Goddess, Artemis, reveals the secrets of her cleavage. *DNA Repair*, 1:771–777, 2002.
- [66] Yunmei Ma, Ulrich Pannicke, Klaus Schwarz, and Michael R Lieber. Hairpin Opening and Overhang Processing by an Artemis / DNA-Dependent Protein Kinase Complex in Nonhomologous End Joining and V(D)J Recombination. *Cell*, 108:781–794, 2002.
- [67] Claire Chappell, Les A Hanakahi, Feridoun Karimi-Busheri, Michael Weinfeld, and Stephen C West. Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. *The EMBO Journal*, 21(11):2827–2832, 2002.
- [68] Kiran N Mahajan, Lisa Gangi-Peterson, David H Sorscher, Jingsong Wang, Karen N Gathy, Nupam P Mahajan, Westley H Reeves, and Beverly S Mitchell. Association of terminal deoxynucleotidyl transferase with Ku. *Proceedings of the National Academy of Sciences*, 96(24):13926–13931, 1999.
- [69] Kiran N Mahajan, Stephanie A Nick Mcelhinny, Beverly S Mitchell, Dale A Ramsden, Chapel Hill, and North Carolina. Association of DNA Polymerase mu (polmu) with Ku and Ligase IV: Role for pol in End-Joining Double-Strand Break Repair. *Molecular Cell Biology*, 22(14):5194–5202, 2002.
- [70] Jean-Pascal Capp, Pascale Bertrand, Audrey Laroche-Clary, Philippe Pourquier, Bernard S Lopez, Christophe Cazaux, and Yvan Canitrot. The DNA polymerase lambda is required for the repair of non-compatible DNA double strand breaks by NHEJ in mammalian cells. *Nucleic Acids Research*, 34(10):2998–3007, 2006.
- [71] Brandi L Mahaney, Katheryn Meek, and Susan P Lees-miller. Repair of ionizing radiation-induced DNA double strand breaks by non-homologous end-joining. *Biochemical Journal*, 417(3):639–650, 2009.

- [72] Chun J Tsai, Sunny A Kim, and Gilbert Chu. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. *Proceedings of the National Academy of Sciences*, 104(19):7851–7856, 2007.
- [73] Jiafeng Gu, Haihui Lu, Albert G Tsai, Klaus Schwarz, and Michael R Lieber. Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence. *Nucleic Acids Research*, 35(17):5755–5762, 2007.
- [74] Catherine T Yan, Cristian Boboila, Ellen Kris Souza, Sonia Franco, Thomas R Hickernell, Michael Murphy, Sunil Gumaste, Mark Geyer, Ali A Zarrin, John P Manis, Klaus Rajewsky, and Frederick W Alt. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature*, 449:478–483, 2007.
- [75] Nicole S Verkaik, Rebecca E E Esveldt-van Lange, Diana Van Heemst, Hennie T Bruggenwirth, Jan H J Hoeijmakers, Malgorzata Z Zdzienicka, and Dik Van Gent. Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. *European Journal of Immunology*, 32:701–709, 2002.
- [76] James E Haber. Alternative endings. *Proceedings of the National Academy of Sciences*, 105(2):405–406, 2008.
- [77] Shuren Liao, Margaret Tamaro, and Hong Yan. The structure of ends determines the pathway choice and Mre11 nuclease dependency of DNA double-strand break repair. *Nucleic Acids Research*, 44(12):5689–5701, 2016.
- [78] Nicole Hustedt and Daniel Durocher. The control of DNA repair by the cell cycle. *Nature Cell Biology*, 19(1):1–9, 2017.
- [79] Meena Shrivastav, Leyma P De Haro, and Jac A Nickoloff. Regulation of DNA double-strand break repair pathway choice. *Cell Research*, 18:134–147, 2008.
- [80] Kyoko Nakamura, Wataru Sakai, Takuo Kawamoto, Ronan T Bree, Noel F Lowndes, Sunichi Takeda, and Yoshihito Taniguchi. Genetic dissection of vertebrate 53BP1: A major role in non-homologous end joining of DNA double strand breaks. *DNA Repair*, 5:741–749, 2006.
- [81] Michal Zimmermann, Francisca Lottersberger, Sara B Buonomo, Agnel Sfeir, and Titia De Lange. 53BP1 Regulates DSB Repair Using. *Science*, 339:700–705, 2013.
- [82] Cristina Escribano-Diaz and Daniel Durocher. DNA repair pathway choice—a PTIP of the hat to 53BP1. *EMBO reports*, 14(8):665–666, 2013.
- [83] Longchuan Chen, Christian J Nievera, Alan Yueh-luen Lee, and Xiaohua Wu. Cell Cycle-dependent Complex Formation of BRCA1 CtIP MRN Is Important for DNA Double-strand Break Repair. *Journal of Biological Chemistry*, 283(12):7713–7720, 2008.
- [84] Andrew Seeber, Michael Hauer, and Susan M Gasser. Nucleosome remodelers in double-strand break repair. *Current Opinion in Genetics & Development*, 23:174–184, 2013.
- [85] Yael Aylon, Batia Liefshitz, and Martin Kupiec. The CDK regulates repair of double-strand breaks by homologous recombination during the cell. *The EMBO Journal*, 23(24):4868–4875, 2004.

- [86] Grzegorz Ira, Achille Pelliccioli, Alitukiriza Balijja, Xuan Wang, Simona Fiorani, Walter Carotenuto, Giordano Liberi, Debra Bressan, Lihong Wan, Nancy M Hollingsworth, James E Haber, and Marco Foiani. DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. *Nature*, 431:1011–1017, 2004.
- [87] Nozomi Tomimatsu, Bipasha Mukherjee, Molly Catherine Hardebeck, Mariya Ilcheva, Cristel Vanessa Camacho, Janelle Louise Harris, Matthew Porteus, Bertrand Llorente, Kum Kum Khanna, and Sandeep Burma. Phosphorylation of EXO1 by CDKs 1 and 2 regulates DNA end resection and repair pathway choice. *Nature Communications*, 5:1–10, 2014.
- [88] Zhongsheng You, Charly Chahwan, Julie Bailis, Tony Hunter, and Paul Russell. ATM Activation and Its Recruitment to Damaged DNA Require Binding to the C Terminus of Nbs1. *Molecular and Cellular Biology*, 25(13):5363–5379, 2005.
- [89] Tanya T. Paull. Mechanisms of ATM Activation. *Annual Review of Biochemistry*, 84:711–738, 2015.
- [90] Christian B Schiller, Katja Lammens, Iliaria Guerini, Britta Coordes, Heidi Feldmann, Florian Schlauderer, Carolin Möckel, Alexandra Schele, Katja Strässer, Stephen P Jackson, and Karl-Peter Hopfner. Structure of Mre11–Nbs1 complex yields insights into ataxia-telangiectasia-like disease mutations and DNA damage signaling. *Nature Structural & Molecular Biology*, 19(7):693–700, 2012.
- [91] Christian Bernd Schiller, Florian Ulrich Seifert, Christian Linke-Winnebeck, and K.-P. Hopfner. Structural Studies of DNA End Detection and Resection in Homologous Recombination. *Cold Spring Harbor Perspectives in Biology*, 6:1–25, 2014.
- [92] Karl-Peter Hopfner, Annette Karcher, David S Shin, Lisa Craig, L Matthew Arthur, James P Carney, and John A Tainer. Structural Biology of Rad50 ATPase: ATP-Driven Conformational Control in DNA Double-Strand Break Repair and the ABC-ATPase Superfamily. *Cell*, 101:789–800, 2000.
- [93] Karl-Peter Hopfner, Annette Karcher, Lisa Craig, Tammy T Woo, James P Carney, and John A Tainer. Structural Biochemistry and Interaction Architecture of the DNA Double-Strand Break Repair Mre11 Nuclease and Rad50-ATPase. *Cell*, 105:473–485, 2001.
- [94] Katja Lammens, Derk J. Bemeleit, Carolin Mockel, Emanuel Clausing, Alexandra Schele, Sophia Hartung, Christian B Schiller, Maria Lucas, Christof Angermuller, Johannes Soding, Katja Strässer, and Karl-Peter Hopfner. The Mre11 : Rad50 Structure Shows an ATP-Dependent Molecular Clamp in DNA Double-Strand Break Repair. *Cell*, 145:54–66, 2011.
- [95] Gareth J Williams, R Scott Williams, Jessica S Williams, Gabriel Moncalian, Andrew S Arvai, Oliver Limbo, Grant Guenther, Soumita Sildas, Michal Hammel, Paul Russell, and John A Tainer. ABC ATPase signature helices in Rad50 link nucleotide state to Mre11 interface for DNA repair. *Nature Structural and Molecular Biology*, 18(4):423–432, 2011.
- [96] Young Bong Park, Jina Chae, Young Chang Kim, and Yunje Cho. Crystal Structure of Human Mre11: Understanding Tumorigenic Mutations. *Structure*, 19(11):1591–1602, 2011.
- [97] Florian Ulrich Seifert, Katja Lammens, and Karl Peter Hopfner. Structure of the catalytic domain of Mre11 from *Chaetomium thermophilum*. *Acta Crystallographica Section F*, 71:752–757, 2015.

- [98] Martijn De Jager, M Trujillo, Kelly, Patrick Sung, Karl-Peter Hopfner, James P Carney, John A. Tainer, John C Connelly, David R F Leach, Roland Kanaar, and Claire Wyman. Differential Arrangements of Conserved Building Blocks among Homologs of the Rad50/Mre11 DNA Repair Protein Complex. *Journal of Molecular Biology*, 339(4):937–949, 2004.
- [99] Karl-Peter Hopfner, Lisa Craig, Gabriel Moncalian, Robert A Zinkel, Takehiko Usui, Barbara A L Owen, Annette Karcher, Brendan Henderson, Jean-Luc Bodmer, Cynthia T McMurray, James P Carney, John H J Petrini, and John A Tainer. The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature*, 418:562–566, 2002.
- [100] Young Bong Park, Marcel Hohl, Michal Padjasek, Eunyoung Jeong, Kyeong Sik Jin, Artur Krezel, John H.J. Petrini, and Yunje Cho. Eukaryotic Rad50 functions as a rod-shaped dimer. *Nature Structural and Molecular Biology*, 24(3):248–257, 2017.
- [101] Marcel Hohl, Tomasz Kochanczyk, Cristina Tous, Andres Aguilera, Artur Krezel, and John H J Petrini. Interdependence of the Rad50 Hook and Globular Domain Functions. *Molecular Cell*, 57:479–491, 2015.
- [102] Jed J W Wiltzius, Marcel Hohl, James C Fleming, and John H J Petrini. The Rad50 hook domain is a critical determinant of Mre11 complex functions. *Nature Structural and Molecular Biology*, 12(5):403–407, 2005.
- [103] Janette Lloyd, J Ross Chapman, Julie A Clapperton, Lesley F Haire, Edgar Hartsuiker, Jiejun Li, Antony M Carr, Stephen P Jackson, and Stephen J Smerdon. A Supramodular FHA/BRCT-Repeat Architecture Mediates Nbs1 Adaptor Function in Response to DNA Damage. *Cell*, 139:100–111, 2009.
- [104] R Scott Williams, Gerald E Dodson, Oliver Limbo, Yoshiki Yamada, Jessica S Williams, Grant Guenther, Scott Classen, J N Mark Glover, Hiroshi Iwasaki, Paul Russell, and John A Tainer. Nbs1 Flexibly Tethers Ctp1 and Mre11- Rad50 to Coordinate DNA Double-Strand Break Processing and Repair. *Cell*, 139:87–99, 2009.
- [105] J Ross Chapman and Stephen P Jackson. Phospho-dependent interactions between NBS1 and MDC1 mediate chromatin retention of the MRN complex at sites of DNA damage. *EMBO reports*, 9(8):795–801, 2008.
- [106] Chao Xu, Liming Wu, Gaofeng Cui, Maria Victoria Botuyan, Junjie Chen, and Georges Mer. Structure of a Second BRCT Domain Identified in the Nijmegen Breakage Syndrome Protein Nbs1 and its Function in an MDC1-Dependent Localization of Nbs1 to DNA Damage Sites. *Journal of Molecular Biology*, 381(2):361–372, 2008.
- [107] A M I Desai-Mehta, Karen M Cerosaletti, and Patrick Concannon. Distinct Functional Domains of Nibrin Mediate Mre11 Binding , Focus Formation , and Nuclear Localization. *Molecular and Cellular Biology*, 21(6):2184–2191, 2001.
- [108] Martijn De Jager, Claire Wyman, Dik C Van Gent, and Roland Kanaar. DNA end-binding specificity of human Rad50/Mre11 is influenced by ATP. *Nucleic Acids Research*, 30(20):4425–4431, 2002.
- [109] Rajashree A Deshpande, Ji-Hoon Lee, and Tanya T Paull. Rad50 ATPase activity is regulated by DNA ends and requires coordination of both active sites. *Nucleic Acids Research*, 45(9):5255–5268, 2017.

- [110] Kelly M Trujillo, Dong Hyun, Ling Chen, Stephen Van Komen, Alan Tomkinson, and Patrick Sung. Yeast Xrs2 Binds DNA and Helps Target Rad50 and Mre11 to DNA Ends. *Journal of Biological Chemistry*, 278(49):48957–48964, 2003.
- [111] Sarem Hailemariam, Sandeep Kumar, and Peter M Burgers. Activation of Tel1 ATM kinase requires Rad50 ATPase and long nucleosome-free DNA , but no DNA ends. *Journal of Biological Chemistry*, 294:10120–10130, 2019.
- [112] Ji-Hoon Lee and Tanya T Paull. ATM Activation by DNA Double-Strand Breaks Through the Mre11-Rad50-Nbs1 Complex. *Science*, 308(April):551–554, 2005.
- [113] Yaqi Liu, Sihyun Sung, Youngran Kim, Fuyang Li, Gwanghyun Gwon, Aera Jo, Ae-Kyoung Kim, Taeyoon Kim, Ok-kyu Song, Sang Eun Lee, and Yunje Cho. ATP-dependent DNA binding, unwinding, and resection by the Mre11/Rad50 complex. *The EMBO Journal*, 35(7):743–758, 2016.
- [114] Anna Rojowska, Katja Lammens, Florian U Seifert, Carolin Drenberger, Heidi Feldmann, and Karl-Peter Hopfner. Structure of the Rad50 DNA double strand break repair protein in complex with DNA. *The EMBO Journal*, 33:2847–2859, 2014.
- [115] Florian Ulrich Seifert, Katja Lammens, Gabriele Stoehr, Brigitte Kessler, and Karl-Peter Hopfner. Structural mechanism of ATP-dependent DNA binding and DNA end bridging by eukaryotic Rad50. *The EMBO journal*, 35(7):759–72, 2016.
- [116] Carolin Mockel, Katja Lammens, Alexandra Schele, and Karl-peter Hopfner. ATP driven structural changes of the bacterial Mre11:Rad50 catalytic head complex. *Nucleic Acids Research*, 40(2):914–927, 2012.
- [117] Lisa Kaeshammer, Jan-Hinnerk Saathoff, Katja Lammens, Aaron Alt, Brigitte Kessler, and Karl-Peter Hopfner. Mechanism of DNA End Sensing and Processing by the Mre11-Rad50 Complex. *Molecular Cell*, 76:382–394, 2019.
- [118] Jing Zhuang, Guochun Jiang, Henning Willers, and Fen Xia. Exonuclease Function of Human Mre11 Promotes Deletional Nonhomologous End Joining. *The Journal of Biological Chemistry*, 284(44):30565–30573, 2009.
- [119] Anyong Xie, Amy Kwok, and Ralph Scully. Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nature Structural and Molecular Biology*, 16(8):814–819, 2009.
- [120] Logan R Myler, Ignacio F Gallardo, Michael M Soniat, Rajashree A Deshpande, Xenia B Gonzalez, Yoori Kim, Tanya T Paull, and Ilya J Finkelstein. Single-Molecule Imaging Reveals How Mre11-Rad50-Nbs1 Initiates DNA Break Repair. *Molecular Cell*, 67:891–898, 2017.
- [121] Giordano Reginato, Elda Cannavo, and Petr Cejka. Physiological protein blocks direct the Mre11-Rad50-Xrs2 and Sae2 nuclease complex to initiate DNA end resection. *Genes and Development*, 31:2325–2330, 2017.
- [122] Elda Cannavo and Petr Cejka. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature*, 514:122–127, 2014.
- [123] Pablo Huertas, Felipe Corte, Alessandro A Sartori, and Stephen P Jackson. CDK targets Sae2 to control DNA-end resection and homologous recombination. *Nature*, 455(689-693), 2008.

- [124] Oliver Limbo, Charly Chahwan, Yoshiki Yamada, Robertus A M de Bruin, Curt Wittenberg, and Paul Russell. Ctp1 Is a Cell-Cycle-Regulated Protein that Functions with Mre11 Complex to Control Double-Strand Break Repair by Homologous Recombination. *Molecular Cell*, 28:134–146, 2007.
- [125] Hideki Takata, Yayoi Tanaka, and Akira Matsuura. Late S Phase-Specific Recruitment of Mre11 Complex Triggers Hierarchical Assembly of Telomere Replication Proteins in *Saccharomyces cerevisiae*. *Molecular Cell*, 17:573–583, 2005.
- [126] Hiroyuki Takai, Agata Smogorzewska, and Titia De Lange. DNA Damage Foci at Dysfunctional Telomeres. *Current Biology*, 13:1549–1556, 2003.
- [127] Jan Karlseder, Dominique Broccoli, Yumin Dai, Stephen Hardy, and Titia De Lange. p53- and ATM-Dependent Apoptosis Induced by Telomeres Lacking TRF2. *Science*, 283:1321–1326, 1999.
- [128] Masahiko Kobayashi, Naoyuki Hayashi, Minoru Takata, and Ken-ichi Yamamoto. NBS1 directly activates ATR independently of MRE11 and. *Genes to Cells*, 18:238–246, 2013.
- [129] Joon Lee and William G Dunphy. The Mre11-Rad50-Nbs1 (MRN) complex has a specific role in the activation of Chk1 in response to stalled replication forks. *Molecular Biology of the Cell*, 24:1343–1353, 2013.
- [130] Magtouf Gatei, Amanda W Kijas, Denis Biard, Thilo Dork, and Martin F Lavin. RAD50 phosphorylation promotes ATR downstream signaling and DNA restart following replication stress. *Human Molecular Genetics*, 23(16):4232–4248, 2014.
- [131] Kinneret Savitsky, Anat Bar-Shira, Shlomit Gilad, Galit Rotman, Yael Ziv, Lina Vanagaite, Danilo A. Tagle, Sara Smith, Tamar Uziel, Sharon Sfez, Maya Ashkenazi, Iris Pecker, Moshe Frydman, Reli Harnik, Sankhavaram R. Patanjali, Andrew Simmons, Gregory A. Clines, Adam Sartiel, Richard A. Gatti, Luciana Chessa, Ozden Sanal, Martin F. Lavin, N. G. J. Jaspers, A. Malcolm R. Taylor, Colin F. Arlett, Toru Miki, Sherman M. Weissman, Michael Lovett, Francis S. Collins, and Yosef Shiloh. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science*, 268:1749–1753, 1995.
- [132] Sian D Spacey, Richard A Gatti, and Gwyn Bebb. The Molecular Basis and Clinical Management of Ataxia Telangiectasia. *The Canadian Journal of Neurological Sciences*, 27:184–191, 2000.
- [133] Grant S Stewart, Richard S Maser, Tanja Stankovic, Debra A Bressan, Mark I Kaplan, Nikolaas G J Jaspers, Anja Raams, Philip J Byrd, John H J Petrini, and A Malcolm R Taylor. The DNA Double-Strand Break Repair Gene hMRE11 Is Mutated in Individuals with an Ataxia-Telangiectasia-like Disorder. *Cell*, 99:577–587, 1999.
- [134] A M R Taylor, A Groom, and P J Byrd. Ataxia-telangiectasia-like disorder (ATLD) — its clinical presentation and molecular basis. *DNA Repair*, 3:1219–1225, 2004.
- [135] Tamar Uziel, Yaniv Lerenthal, Lilach Moyal, Yair Andegeko, Leonid Mittelman, and Yosef Shiloh. Requirement of the MRN complex for ATM activation by DNA damage. *The EMBO Journal*, 22(20):5612–5621, 2003.
- [136] M. Choi, T. Kipps, and R. Kurzrock. ATM Mutations in Cancer: Therapeutic Implications. *Molecular Cancer Therapeutics*, 15(8):1781–1791, 2016.
- [137] Daisuke Nakada, Kunihiro Matsumoto, and Sugimoto. ATM-related Tel1 associates with double-strand breaks through an Xrs2-dependent mechanism. *Genes and Development*, 17:1957–1962, 2003.

- [138] Rolf J Craven, Patricia W Greenwell, Margaret Dominska, and Thomas D Petes. Regulation of Genome Stability by TEL1 and MEC1 , Yeast Homologs of the Mammalian ATM and ATR Genes. *Genetics*, 507:493–507, 2002.
- [139] Curtis T Keith and Stuart L Schreiber. PIK-Related Kinases: DNA Repair , Recombination, and Cell Cycle Checkpoints. *Science*, 270:50, 1995.
- [140] Domagoj Baretic and Roger L. Williams. PIKKs - the solenoid nest where partners and kinases meet. *Current Opinion in Structural Biology*, 29:134–142, 2014.
- [141] Seong-tae Kim, Dae-sik Lim, Christine E Canman, and Michael B Kastan. Substrate Specificities and Identification of Putative Substrates of ATM Kinase Family Members. *The Journal of Biological Chemistry*, 274(53):37538–37543, 1999. ISSN 1083-351X. doi: doi:10.1074/jbc.274.53.37538.
- [142] Ted O’Neill, Alison J. Dwyer, Yael Ziv, Doug W. Chan, Susan P. Lees-Miller, Robert H. Abraham, Jack H. Lai, David Hill, Yossi Shiloh, Lewis C. Cantley, and Gary A. Rathbun. Utilization of oriented peptide libraries to identify substrate motifs selected by ATM. *Journal of Biological Chemistry*, 275(30):22719–22727, 2000.
- [143] Jason Perry and Nancy Kleckner. The ATRs, ATMs, and TORs are giant HEAT repeat proteins. *Cell*, 112:151–155, 2003.
- [144] Miguel A Andrade, Carolina Perez-Iratxeta, and Chris P Ponting. Protein Repeats: Structures, Functions ,and Evolution. *Journal of Structural Biology*, 134:117–131, 2001.
- [145] Harri Lempiainen and Thanos D Halazonetis. Emerging common themes in regulation of PIKKs. *The EMBO Journal*, 28(20):3067–3073, 2009.
- [146] Stefan Imseng, Christopher HS Aylett, and Timm Maier. Architecture and activation of phosphatidylinositol 3-kinase related kinases. *Current Opinion in Structural Biology*, 49:177–189, 2018.
- [147] Daniel A Mordes and David Cortez. Activation of ATR and related PIKKs. *Cell Cycle*, 7(18):2809–2812, 2008.
- [148] Joungmok Kim and Kun-liang Guan. mTOR as a central hub of nutrient signalling and cell growth. *Nature Cell Biology*, 21:63–71, 2019.
- [149] Haijuan Yang, Derek G. Rudge, Joseph D. Koos, Bhamini Vaidialingam, Hyo J. Yang, and Nikola P. Pavletich. mTOR kinase structure, mechanism and regulation. *Nature*, 497:217–223, 2013.
- [150] Lee Zou and Stephen J Elledge. Sensing DNA Damage Through ATRIP Recognition of RPA-ssDNA Complexes. *Science*, 300:1542–1549, 2003.
- [151] Akiko Kumagai, Joon Lee, Hae Yong Yoo, and William G Dunphy. TopBP1 Activates the ATR-ATRIP Complex. *Cell*, 124:943–955, 2006.
- [152] Katharine O Hartley, David Gell, Graeme C M Smith, Hong Zhang, Nullin Divecha, Margery A Connelly, Arie Admon, Susan P Lees-Miller, Carl W Anderson, and Stephen P Jackson. DNA-Dependent Protein Kinase Catalytic Subunit: A Relative of Phosphatidylinositol 3-Kinase and the Ataxia Telangiectasia Gene Product. *Cell*, 82:849–856, 1995.
- [153] Bancinyane L. Sibanda, Dimitri Y. Chirgadze, and Tom L. Blundell. Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. *Nature*, 463:118–121, 2010.

- [154] Peggy P Hsu, Seong A Kang, Jonathan Rameseder, Yi Zhang, Kathleen A Ottina, Daniel Lim, Timothy R Peterson, Yongmun Choi, Nathanael S Gray, Michael B Yaffe, Jarrod A Marto, and David M Sabatini. The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science*, 332(6035):1317–1322, 2011.
- [155] James P B Lloyd. The evolution and diversity of the nonsense-mediated mRNA decay pathway. *F1000Research*, 7(1299):1–27, 2018.
- [156] Luis Miguel Díaz-Santín, Natasha Lukoyanova, Emir Aciyan, and Alan C M Cheung. Cryo-EM structure of the SAGA and NuA4 coactivator subunit Tra1 at 3.7 Å resolution. *eLife*, 6:1–20, 2017.
- [157] Haibo Wang, Christian Dienemann, Alexandra Stützer, Henning Urlaub, Alan C M Cheung, and Patrick Cramer. Structure of the transcription coactivator. *Nature*, 577:717–720, 2020.
- [158] Manuel Stucki, Julie A Clapperton, Duaa Mohammad, Michael B Yaffe, Stephen J Smerdon, and Stephen P Jackson. MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks. *Cell*, 123:1213–1226, 2005.
- [159] Zhenkun Lou, Katherine Minter-dykhhouse, Sonia Franco, Monica Gostissa, Melissa A Rivera, Arkady Celeste, John P Manis, Jan Van Deursen, Tanya T Paull, Frederick W Alt, and Junjie Chen. MDC1 Maintains Genomic Stability by Participating in the Amplification of ATM-Dependent DNA Damage Signals. *Molecular Cell*, 21:187–200, 2006.
- [160] Andrea Beucher, Julie Birraux, Leopoldine Tchouandong, Olivia Barton, Atsushi Shibata, Sandro Conrad, Aaron A Goodarzi, Andrea Krempler, Penny A Jeggo, and Markus Lobrich. ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. *The EMBO Journal*, 28(21):3413–3427, 2009.
- [161] Enriqueta Riballo, Martin Ku, Nicole Rief, Aidan Doherty, Graeme C M Smith, Caroline Reis, Kirsten Dahm, Andreas Fricke, Andrea Krempler, Antony R Parker, Stephen P Jackson, Andrew Gennery, Penny A Jeggo, and Markus Lobrich. A Pathway of Double-Strand Break Rejoining Dependent upon ATM, Artemis, and Proteins Locating to  $\gamma$ -H2AX Foci. *Molecular Cell*, 16:715–724, 2004.
- [162] Aaron A Goodarzi, Yaping Yu, Enriqueta Riballo, Pauline Douglas, A Walker, Ruiqiong Ye, Christine Ha, Caterina Marchetti, Nick Morrice, A Jeggo, and Susan P Lees-Miller. DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *The EMBO Journal*, 25(16):3880–3889, 2006.
- [163] Vipul Kumar, Frederick W Alt, and Valentyn Oksenysh. Functional Overlaps Between XLF and The ATM-dependent DNA Double Strand Break Response. *DNA Repair*, 16:11–22, 2014.
- [164] Sébastien Britton, Julia Coates, and Stephen P Jackson. A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *Journal of Biological Chemistry*, 202(3):579–595, 2013.
- [165] Pauline Chanut, Sébastien Britton, Julia Coates, Stephen P. Jackson, and Patrick Calsou. Coordinated nuclease activities counteract Ku at single-ended DNA double-strand breaks. *Nature Communications*, 7:1–11, 2016.
- [166] Michael J. Kruhlak, Arkady Celeste, Graham Dellaire, Oscar Fernandez-Capetillo, Waltraud G. Müller, James G. McNally, David P. Bazett-Jones, and André Nussenzweig. Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. *The Journal of Cell Biology*, 172(6): 823–834, 2006.

- [167] Andreas Kakarougkas, Amani Ismail, Anna L Chambers, Enriqueta Riballo, Alex D Herbert, Julia Ku, Penny A Jeggo, and Jessica A Downs. Requirement for PBAF in Transcriptional Repression and Repair at DNA Breaks in Actively Transcribed Regions of Chromatin. *Molecular Cell*, 55:723–732, 2014.
- [168] Matthias Altmeyer, Kai J Neelsen, Federico Teloni, Irina Pozdnyakova, Stefania Pellegrino, Merete Grøfte, Maj Britt Druedahl Rask, Werner Streicher, Stephanie Jungmichel, Michael Lund Nielsen, and Jiri Lukas. Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nature Communications*, 6(8088):1–12, 2015.
- [169] Pierre Caron, Jonathan Choudjaye, Thomas Clouaire, Felipe Corte, Virginie Daburon, Marion Aguirrebengoa, Thomas Mangeat, Jason S Iacovoni, Alejandro Alvarez-Quilon, Felipe Cortes-Ledesma, and Gaëlle Legube. Non-redundant Functions of ATM and DNA-PKcs in Response to DNA Double-Strand Breaks Article Non-redundant Functions of ATM and DNA-PKcs in Response to DNA Double-Strand Breaks. *Cell Reports*, 13:1598–1609, 2015.
- [170] Maria Tresini, Daniël O. Warmerdam, Petros Kolovos, Loes Snijder, Mischa G. Vrouwe, Jeroen A.A. Demmers, Wilfred F.J. Van Ijcken, Frank G. Grosveld, René H. Medema, Jan H.J. Hoeijmakers, Leon H.F. Mullenders, Wim Vermeulen, and Jurgen A. Marteijn. The core spliceosome as target and effector of non-canonical ATM signalling. *Nature*, 523:53–58, 2015.
- [171] Zhi Guo, Rajashree Deshpande, and Tanya T Paull. ATM activation in the presence of oxidative stress. *Cell Cycle*, 9(24):4805–4811, 2010.
- [172] Govind A. Shah and Clodagh C. O’Shea. Viral and Cellular Genomes Activate Distinct DNA Damage Responses. *Cell*, 162:987–1002, 2015.
- [173] Susan P Lees-Miller, Melissa C Long, M Alison Kilvert, Vivian Lam, Stephen A Rice, and Charlotte A Spencer. Attenuation of DNA-Dependent Protein Kinase Activity and Its Catalytic Subunit by the Herpes Simplex Virus Type 1 Transactivator ICP0. *Journal of Virology*, 70(11):7471–7477, 1996.
- [174] Andrew N Blackford, Rakesh N Patel, Natalie A Forrester, Kathrin Theil, Peter Groitl, Grant S Stewart, A. Malcolm R Taylor, M Morgan, Iain, Thomas Dobner, Roger J A Grand, and Andrew S Turnell. Adenovirus 12 E4orf6 inhibits ATR activation by promoting TOPBP1 degradation. *Proceedings of the National Academy of Sciences*, 107(27):12251–12256, 2010.
- [175] Travis H Stracker, Christian T Carson, and Matthew D Weitzman. Adenovirus oncoproteins inactivate the Mre11 Rad50-NBS1-DNA repair complex. *Nature*, 418:348–352, 2002.
- [176] Oscar Vadas, John E. Burke, Xuxiao Zhang, Alex Berndt, and Roger L. Williams. Structural basis for activation and inhibition of class I phosphoinositide 3-kinases. *Science Signaling*, 4(195):re2, 2011.
- [177] Ji-Hoon Lee and Tanya T Paull. Direct Activation of the ATM Protein Kinase by the Mre11 / Rad50 / Nbs1 Complex. *Science*, 304:93–96, 2004.
- [178] Oliver Limbo, Yoshiki Yamada, and Paul Russell. Mre11-Rad50-dependent activity of ATM/Tel1 at DNA breaks and telomeres in the absence of Nbs1. *Molecular Biology of the Cell*, 29(11):1389–1399, 2018.
- [179] Hartlerode Andrea J, Mary J Morgan, Yipin Wu, Jeffrey Buis, and David O Ferguson. Recruitment and activation of the ATM kinase in the absence of DNA damage sensors. *Nature Structural and Molecular Biology*, 22(9):736–743, 2015.

- [180] Aude Dupré, Louise Boyer-Chatenet, and Jean Gautier. Two-step activation of ATM by DNA and the Mre11-Rad50-Nbs1 complex. *Nature Structural and Molecular Biology*, 13(5):451–457, 2006.
- [181] Zhongsheng You, Julie M. Bailis, Sam A. Johnson, Stephen M. Dilworth, and Tony Hunter. Rapid activation of ATM on DNA flanking double-strand breaks. *Nature Cell Biology*, 9(11):1311–1318, 2007.
- [182] Graeme C Smith, Robert B Cary, Nicholas D Lakin, Byron C Hann, Soo-Hwang Teo, David J Chen, and Stephen P Jackson. Purification and DNA binding properties of the ataxia-telangiectasia gene product ATM. *Proceedings of the National Academy of Sciences of the United States of America*, 96: 11134–11139, 1999.
- [183] O. Llorca, A. Rivera-Calzada, J. Grantham, and K. R. Willison. Electron microscopy and 3D reconstructions reveal that human ATM kinase uses an arm-like domain to clamp around double-stranded DNA. *Oncogene*, 22:3867–3874, 2003.
- [184] Marina K Ayrappetov, Ozge Gursoy-Yuzugullu, Chang Xu, Ye Xu, and Brendan D Price. DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin. *Proceedings of the National Academy of Sciences*, 111(25):9169–9174, 2014.
- [185] Yong-chul Kim, Gabi Gerlitz, Takashi Furusawa, Frédéric Catez, Andre Nussenzweig, Kyu-seon Oh, Kenneth H Kraemer, Yosef Shiloh, and Michael Bustin. Activation of ATM depends on chromatin interactions occurring before induction of DNA damage. *Nature Cell Biology*, 11(1):92–96, 2009.
- [186] Manuela Pellegrini, Arkady Celeste, Simone Difilippantonio, Rong Guo, Weidong Wang, Lionel Feigenbaum, and André Nussenzweig. Autophosphorylation at serine 1987 is dispensable for murine Atm activation in vivo. *Nature*, 443:222–225, 2006.
- [187] Jeremy A. Daniel, Manuela Pellegrini, Ji Hoon Lee, Tanya T Paull, Lionel Feigenbaum, and André Nussenzweig. Multiple autophosphorylation sites are dispensable for murine ATM activation in vivo. *Journal of Cell Biology*, 183(5):777–783, 2008.
- [188] Luke A Yates, Rhys M Williams, Sarem Hailemariam, Rafael Ayala, Peter Burgers, and Xiaodong Zhang. Cryo-EM Structure of Nucleotide-Bound Tel1 ATM Unravels the Molecular Basis of Inhibition and Structural Rationale for Disease-Associated Mutations. *Structure*, 28:1–9, 2020.
- [189] Jianxiong Xiao, Mengjie Liu, Yilun Qi, Yuriy Chaban, Chao Gao, Beiqing Pan, Yuan Tian, Zishuo Yu, Jie Li, Peijun Zhang, and Yanhui Xu. Structural insights into the activation of ATM kinase. *Cell Research*, 29:683–685, 2019.
- [190] Yi Zhou, Ji Hoon Lee, Wenxia Jiang, Jennie L. Crowe, Shan Zha, and Tanya T. Paull. Regulation of the DNA Damage Response by DNA-PKcs Inhibitory Phosphorylation of ATM. *Molecular Cell*, 65: 91–104, 2017.
- [191] Di Michela Virgilio, Carol Y Ying, and Jean Gautier. PIKK-dependent phosphorylation of Mre11 induces MRN complex inactivation by disassembly from chromatin. *DNA Repair*, 8(11):1311–1320, 2009.
- [192] Aaron A Goodarzi, Jyoti C Jonnalagadda, Pauline Douglas, David Young, Ruiqiong Ye, Greg Bg, and Kum Kum Khanna. Autophosphorylation of ataxia-telangiectasia mutated is regulated by protein phosphatase 2A. *The EMBO Journal*, 23(22):4451–4461, 2004.

- [193] Aimin Peng, Andrea L Lewellyn, William P Schiemann, and James L Maller. Repo-Man Controls a Protein Phosphatase 1-Dependent Threshold for DNA Damage Checkpoint Activation. *Current Biology*, 20:387–396, 2010.
- [194] Michele Fiscella, HongLiang Zhang, Saijun Fan, Kazuyasu Sakaguchi, Songfa Shen, W. Edward Mercer, George F. Vande Woude, Patrick O’Connor, and Ettore Appella. Wip1 , a novel human protein phosphatase that is induced in response to ionizing radiation in a p53-dependent manner. *Proceedings of the National Academy of Sciences*, 94:6048–6053, 1997.
- [195] Nnennaya Kanu and Axel Behrens. ATMIN defines an NBS1-independent pathway of ATM signalling. *The EMBO Journal*, 26:2933–2941, 2007.
- [196] Ji Hoon Lee, Aaron A. Goodarzi, Penny A. Jeggo, and Tanya T. Paull. 53BP1 promotes ATM activity through direct interactions with the MRN complex. *The EMBO Journal*, 29(3):574–585, 2010.
- [197] Cai Bowen, Jeong-Ho Ju, Ji-Hoon Lee, Tanya T Paull, and Edward P Gelmann. Functional Activation of ATM by the Prostate Cancer Suppressor NKX3 . 1. *Cell Reports*, 4(3):516–529, 2013.
- [198] Naokazu Chiba, Valentine Comaills, Bunsyo Shiotani, Fumiyuki Takahashi, Toshiyuki Shimada, and Ken Tajima. Homeobox B9 induces epithelial-to-mesenchymal transition-associated radioresistance by accelerating DNA damage responses. *Proceedings of the National Academy of Sciences*, 109(8):1–6, 2012. doi: 10.1073/pnas.1018867108.
- [199] Jessie Yanxiang Guo, Ayumi Yamada, Taisuke Kajino, Judy Qiju Wu, Wanli Tang, Christopher D Freel, Junjie Feng, B Nelson Chau, Michael Zhuo Wang, Seth S Margolis, Hae Yong Yoo, Xiao-fan Wang, William G Dunphy, Pablo M Irusta, J Marie Hardwick, and Sally Kornbluth. Aven-Dependent Activation of ATM Following DNA Damage. *Current Biology*, 18:933–942, 2008.
- [200] Han Chi Lim, Li Xie, Wei Zhang, Rong Li, Zhong-can Chen, Guang-zhi Wu, Shu-sen Cui, Eng King Tan, and Li Zeng. Ribosomal S6 Kinase 2 (RSK2) Maintains Genomic Stability by Activating the Atm / p53-Dependent DNA Damage Pathway. *PLoS ONE*, 8(9):1–10, 2013.
- [201] Nikola Kellner, Johannes Schwarz, Miriam Sturm, Javier Fernandez-Martinez, Sabine Griesel, Wenzhu Zhang, Brian T. Chait, Michael P. Rout, Ulrich Kück, and Ed Hurt. Developing genetic tools to exploit *Chaetomium thermophilum* for biochemical analyses of eukaryotic macromolecular assemblies. *Scientific Reports*, 6:1–10, 2016.
- [202] Kilian R Knoll, Sebastian Eustermann, Vanessa Niebauer, Elisa Oberbeckmann, Gabriele Stoehr, Kevin Schall, Alessandro Tosi, Marianne Schwarz, Andrea Buchfellner, Philipp Korber, and Karl-peter Hopfner. The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. *Nature Structural & Molecular Biology*, 25:823–832, 2018.
- [203] Anna Gaciarz, Johanna Veijola, Yuko Uchida, Mirva J Saaranen, Chunguang Wang, Sohvi Hörkkö, and Lloyd W Ruddock. Systematic screening of soluble expression of antibody fragments in the cytoplasm of *E. coli*. *Microbial Cell Factories*, 15(22):1–10, 2016.
- [204] Van Dat Nguyen, Feras Hatahet, Kirsi E H Salo, Eveliina Enlund, Chi Zhang, and Lloyd W Ruddock. Pre-expression of a sulfhydryl oxidase significantly increases the yields of eukaryotic disulfide bond containing proteins expressed in the cytoplasm of *E. coli*. *Microbial Cell Factories*, 10(1):1–13, 2011.

- [205] Jasenko Zivanov, Takanori Nakane, Bjorn Forsberg, Dari Kimanius, Wim J.H. Hagen, Erik Lindahl, and Sjors H.W. Scheres. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife*, 7(e42166):1–16, 2018.
- [206] Shawn Q. Zheng, Eugene Palovcak, Jean Paul Armache, Kliment A. Verba, Yifan Cheng, and David A. Agard. MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*, 14(4):331–332, 2017.
- [207] Alexis Rohou and Nikolaus Grigorieff. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, 192:216–221, 2015.
- [208] Domagoj Baretic, Hannah K. Pollard, David I. Fisher, Christopher M. Johnson, Balaji Santhanam, Caroline M. Truman, Tomas Kouba, Alan R. Fersht, Christopher Phillips, and Roger L. Williams. Structures of closed and open conformations of dimeric human ATM. *Science Advances*, 3(5):1–10, 2017.
- [209] Marco Biasini, Stefan Bienert, Andrew Waterhouse, Konstantin Arnold, Gabriel Studer, Tobias Schmidt, Florian Kiefer, Tiziano Gallo Cassarino, Martino Bertoni, Lorenza Bordoli, and Torsten Schwede. SWISS-MODEL : modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42:252–258, 2014.
- [210] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. UCSF Chimera — A Visualization System for Exploratory Research and Analysis. *Journal of Computational Chemistry*, 25(13):1605–1612, 2004.
- [211] Paul D Adams, V Pavel, Vincent B Chen, W Ian, Nathaniel Echols, Nigel W Moriarty, Randy J Read, David C Richardson, S Jane, and C Thomas. PHENIX: a comprehensive Python-based system for macromolecular structure solution research papers. *Acta Crystallographica Section D: Biological Crystallography*, D66:213–221, 2010.
- [212] Paul Emsley and Kevin Cowtan. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, D60:2126–2132, 2004.
- [213] Daniel W A Buchan and David T Jones. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Research*, 47(April):402–407, 2019.
- [214] Thomas D Goddard, Conrad C Huang, Elaine C Meng, Eric F Pettersen, Gregory S Couch, John H Morris, and Thomas E Ferrin. UCSF ChimeraX : Meeting modern challenges in visualization and analysis. *Protein Science*, 27:14–25, 2018.
- [215] Kazutaka Katoh, Kei-ichi Kuma, Hiroyuki Toh, and Takashi Miyata. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, 33(2):511–518, 2005.
- [216] Andrew M Waterhouse, James B Procter, David M A Martin, Michèle Clamp, and Geoffrey J Barton. Jalview Version 2 — a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9):1189–1191, 2009.
- [217] P. Stothard. The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. *BioTechniques*, 28(6):1102–1104, 2000.
- [218] Alessandra Blasina, Brendan D Price, Gaetan A Turenne, and Clare H McGowan. Caffeine inhibits the checkpoint kinase ATM. *Current Biology*, 9(19):1135–1138, 1999.

- [219] Matthias P Wymann, Ginette Bulgarelli-Leva, Marketa J Zvelebil, Luciano Pirola, Bart Vanhaesebroeck, Michael D Waterfield, and George Panayotou. Wortmannin Inactivates Phosphoinositide 3-Kinase by Covalent Modification of Lys-802 , a Residue Involved in the Phosphate Transfer Reaction. *Molecular and Cellular Biology*, 16(4):1722–1733, 1996.
- [220] S. E. Golding, E. Rosenberg, N. Valerie, I. Hussaini, M. Frigerio, X. F. Cockcroft, W. Y. Chong, M. Hummersone, L. Rigoreau, K. A. Menear, M. J. O’Connor, L. F. Povirk, T. van Meter, and K. Valerie. Improved ATM kinase inhibitor KU-60019 radiosensitizes glioma cells, compromises insulin, AKT and ERK prosurvival signaling, and inhibits migration and invasion. *Molecular Cancer Therapeutics*, 8(10):2894–2902, 2009.
- [221] Ian Hickson, Yan Zhao, Caroline J Richardson, Sharon J Green, Niall M B Martin, Alisdair I Orr, Philip M Reaper, Stephen P Jackson, Nicola J Curtin, and Graeme C M Smith. Identification and Characterization of a Novel and Specific Inhibitor of the Ataxia-Telangiectasia Mutated Kinase ATM. *Cancer Research*, 64:9152–9159, 2004.
- [222] Qingsong Liu, Chunxiao Xu, Sivapriya Kirubakaran, Xin Zhang, Wooyoung Hur, Yan Liu, Nicholas Kwiatkowski, Jinhua Wang, Kenneth D Westover, Peng Gao, Dalia Ercan, Mario Niepel, Carson C Thoreen, Seong A Kang, Matthew P Patricelli, Yuchuan Wang, Tanya Tupper, Abigail Altabef, Hidemasa Kawamura, Kathryn D Held, Danny M Chou, Stephen J Elledge, Pasi A Janne, Kwok-Kin Wong, David M Sabatini, and Nathanael S Gray. Characterization of Torin2 , an ATP-Competitive Inhibitor of mTOR, ATM and ATR. *Cancer Research*, 73(8):2574–2587, 2013.
- [223] Werner Kühlbrandt. The Resolution Revolution. *Science*, 343:1443–1445, 2014.
- [224] Wilson CY Lau, Yinyin Li, Zhe Liu, Yuanzhu Gao, Qinfen Zhang, and Michael SY Huen. Structure of the human dimeric ATM kinase. *Cell Cycle*, 15(8):1117–1124, 2016.
- [225] Xuejuan Wang, Huanyu Chu, Mengjuan Lv, Zhihui Zhang, Shuwan Qiu, Haiyan Liu, Xuotong Shen, Weiwu Wang, and Gang Cai. Structure of the intact ATM/Tel1 kinase. *Nature Communications*, 7: 1–8, 2016.
- [226] Marta Sawicka, Paulina H. Wanrooij, Vidya C. Darbari, Elias Tannous, Sarem Hailemariam, Daniel Bose, Alena V. Makarova, Peter M. Burgers, and Xiaodong Zhang. The Dimeric Architecture of Checkpoint Kinases Mec1 ATR and Tel1 ATM Reveal a Common Structural Organization. *Journal of Biological Chemistry*, 291:13436–13447, 2016.
- [227] Jiyu Xin, Zhu Xu, Xuejuan Wang, Yanhua Tian, Zhihui Zhang, and Gang Cai. Structural basis of allosteric regulation of Tel1/ATM kinase. *Cell Research*, 29:655–665, 2019.
- [228] Evgeny Krissinel and Kim Henrick. Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, 372(3):774–797, 2007.
- [229] Kum Kum Khanna, Katherine E Keating, Sergei Kozlov, Shaun Scott, Magtouf Gatei, Karen Hobson, Yoichi Taya, Brian Gabrielli, Doug Chan, Susan P Lees-miller, and Martin F Lavin. ATM associates with and phosphorylates p53: mapping the region of interaction. *Nature*, 20:398–400, 1998.
- [230] Jeffrey J Seidel, Carol M Anderson, and Elizabeth H Blackburn. A Novel Tel1/ATM N-Terminal Motif, TAN, Is Essential for Telomere Length Maintenance and a DNA Damage Response. *Molecular and Cellular Biology*, 28(18):5736–5746, 2008.

- [231] Brad Nolen, Susan Taylor, and Gourisankar Ghosh. Regulation of Protein Kinases : Controlling Activity through Activation Segment Conformation. *Molecular Cell*, 15:661–675, 2004.
- [232] John G Tate, Sally Bamford, Harry C Jubb, Zbyslaw Sondka, David M Beare, Nidhi Bindal, Harry Boutselakis, Charlotte G Cole, Celestino Creatore, Elisabeth Dawson, Peter Fish, Bhavana Harsha, Charlie Hathaway, Steve C Jupe, Yin Kok, Kate Noble, Laura Ponting, Christopher C Ramshaw, Claire E Rye, Helen E Speedy, Ray Stefancsik, Sam L Thompson, Shicai Wang, Sari Ward, J Campbell, and Simon A Forbes. COSMIC : the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research*, 47:941–947, 2019.
- [233] UniProt. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*, 47:506–515, 2019.
- [234] Katsunori Sugimoto. Branching the Tel2 pathway for exact fit on phosphatidylinositol 3-kinase-related kinases. *Current Genetics*, 64(5):965–970, 2018.
- [235] Hiroyuki Takai, Richard C. Wang, Kaori K. Takai, Haijuan Yang, and Titia de Lange. Tel2 Regulates the Stability of PI3K-Related Protein Kinases. *Cell*, 131:1248–1259, 2007.
- [236] Zuzana Horejsi, Hiroyuki Takai, Carrie A Adelman, Spencer J Collis, Helen Flynn, Sarah Maslen, J Mark Skehel, Titia De Lange, and Simon J Boulton. CK2 Phospho-Dependent Binding of R2TP Complex to TEL2 Is Essential for mTOR and SMG1 Stability. *Molecular Cell*, 39:839–850, 2010.
- [237] Séverine Boulon, Bérengère Pradet-Balade, Céline Verheggen, Stéphanie Boireau, Marya Georgieva, Karim Azzag, Yasmeen Ahmad, Henry Neel, Angus I Lamond, and Edouard Bertrand. HSP90 and Its R2TP/Prefoldin-like Cochaperone Are Involved in the Cytoplasmic Assembly of RNA Polymerase II. *Molecular Cell*, 39(6):912–924, 2010.
- [238] Angel Rivera-Calzada, Mohinder Pal, Hugo Muñoz-Hernández, Juan R. Luque-Ortega, David Gil-Carton, Gianluca Degliesposti, J. Mark Skehel, Chrisostomos Prodromou, Laurence H. Pearl, and Oscar Llorca. The Structure of the R2TP Complex Defines a Platform for Recruiting Diverse Client Proteins to the HSP90 Molecular Chaperone System. *Structure*, 25:1145–1152, 2017.
- [239] Walid A Houry, Edouard Bertrand, and Benoit Coulombe. The PAQosome, an R2TP-Based Chaperone for Quaternary Structure. *Trends in Biochemical Sciences*, 43(1):4–9, 2018.
- [240] Hiroyuki Takai, Yihu Xie, Titia De Lange, and Nikola P Pavletich. Tel2 structure and function in the Hsp90-dependent maturation of mTOR and ATR complexes. *Genes and Development*, 24:2019–2030, 2010.
- [241] Xiaotong Yin, Mengjie Liu, Yuan Tian, Jiawei Wang, and Yanhui Xu. Cryo-EM structure of human DNA-PK holoenzyme. *Cell Research*, 27:1341–1350, 2017.
- [242] Yair Andegeko, Lilach Moyal, Leonid Mittelman, Ilan Tsarfaty, and Galit Rotman. Nuclear retention of ATM at sites of DNA double strand breaks. *Journal of Biological Chemistry*, 276(41):38224–38230, 2001.
- [243] Julyun Oh, Amr Al-Zain, Elda Cannavo, Petr Cejka, and Lorraine S. Symington. Xrs2 Dependent and Independent Functions of the Mre11-Rad50 Complex. *Molecular Cell*, 64:405–415, 2016.
- [244] Corinne Cassani, Jacopo Vertemara, Matteo Bassani, Antonio Marsella, Renata Tisi, Giuseppe Zampella, and Maria Pia Longhese. The ATP-bound conformation of the Mre11 – Rad50 complex is essential for Tel1/ATM activation. *Nucleic Acids Research*, 1:1–18, 2019.

- [245] Marcel Hohl, Youngho Kwon, Sandra Muñoz Galván, Xiaoyu Xue, Cristina Tous, Andrés Aguilera, Patrick Sung, and John H J Petrini. The Rad50 coiled-coil domain is indispensable for Mre11 complex functions. *Nature Structural & Molecular Biology*, 18(10):1124–1132, 2011.
- [246] Rajashree A Deshpande, Gareth J Williams, Oliver Limbo, R Scott Williams, Jeff Kuhnlein, Ji-hoon Lee, Scott Classen, Grant Guenther, Paul Russell, John A Tainer, and Tanya T Paull. ATP-driven Rad 50 conformations regulate DNA tethering , end resection , and ATM checkpoint signaling. *The EMBO Journal*, 33(5):482–500, 2014.
- [247] Venugopal Bhaskara, Aude Dupré, Bettina Lengsfeld, Ben B Hopkins, Angela Chan, Ji-hoon Lee, Xiaoming Zhang, Jean Gautier, Virginia Zakian, and Tanya T Paull. Rad50 Adenylate Kinase Activity Regulates DNA Tethering by Mre11/Rad50 Complexes. *Molecular Cell*, 25(5):647–661, 2007.
- [248] Feras E Machour and Nabieh Ayoub. Transcriptional Regulation at DSBs: Mechanisms and Consequences. *Trends in Genetics*, pages 1–17, 2020.
- [249] Yu Chen, Haiping Zhang, Zhu Xu, Huanyin Tang, Anke Geng, Bailian Cai, Tao Su, Jiejun Shi, Cizhong Jiang, Xiao Tian, Andrei Seluanov, Jun Huang, Xiaoping Wan, Ying Jiang, Vera Gorbunova, and Zhiyong Mao. A PARP1 – BRG1 – SIRT1 axis promotes HR repair by reducing nucleosome density at DNA damage sites. *Nucleic Acids Research*, 47(16):8563–8580, 2019.
- [250] Zhao Qin Bao, Douglas M Jacobsen, and Matthew A Young. Briefly Bound to Activate: Transient Binding of a Second Catalytic Magnesium Activates the Structure and Dynamics of CDK2 Kinase for Catalysis. *Structure*, 19(5):675–690, 2011.
- [251] Rhys M Williams, Luke A Yates, and Xiaodong Zhang. Structures and regulations of ATM and ATR , master kinases in genome integrity. *Current Opinion in Structural Biology*, 61:98–105, 2020.
- [252] Qinhui Rao, Mengjie Liu, Yuan Tian, Zihan Wu, Yuhan Hao, Lei Song, Zhaoyu Qin, Chen Ding, Hong-Wei Wang, Jiawei Wang, and Yanhui Xu. Cryo-EM structure of human ATR-ATRIP complex. *Cell Research*, 28:143–156, 2017.
- [253] Xuejuan Wang, Tingting Ran, Xuan Zhang, Jiyu Xin, Zhihui Zhang, Tengwei Wu, Weiwu Wang, and Gang Cai. 3.9 Å structure of the yeast Mec1-Ddc2 complex, a homolog of human ATR-ATRIP. *Science*, 358:1206–1209, 2017.
- [254] Bancinyane L. Sibanda, Dimitri Y. Chirgadze, David B. Ascher, and Tom L. Blundell. DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand break repair. *Science*, 355: 520–524, 2017.
- [255] Humayun Sharif, Yang Li, Yuanchen Dong, Liyi Dong, Wei Li Wang, Youdong Mao, and Hao Wu. Cryo-EM structure of the DNA-PK holoenzyme. *Proceedings of the National Academy of Sciences*, 114(28):7367–7372, 2017.
- [256] Xizi Chen, Mengjie Liu, Yuan Tian, Jiabei Li, Yilun Qi, Dan Zhao, Zihan Wu, Min Huang, Catherine C L Wong, Hong-Wei Wang, Jiawei Wang, Huirong Yang, and Yanhui Xu. Cryo-EM structure of human mTOR complex 2. *Cell Research*, 28:518–528, 2018.
- [257] Huirong Yang, Jia Wang, Mengjie Liu, Xizi Chen, Min Huang, Dan Tan, Meng Qiu Dong, Catherine C.L. Wong, Jiawei Wang, Yanhui Xu, and Hong Wei Wang. 4.4 Å Resolution Cryo-EM structure of human mTOR Complex 1. *Protein and Cell*, 7(12):878–887, 2016.

- [258] Domagoj Baretić, Alex Berndt, Yohei Ohashi, Christopher M Johnson, and Roger L Williams. Tor forms a dimer through an N-terminal helical solenoid with a complex topology. *Nature Communications*, 7:1–9, 2016.
- [259] Edward Stutfeld, Christopher H S Aylett, Stefan Imseng, Daniel Boehringer, Alain Scaiola, Evelyn Sauer, Michael N. Hall, Timm Maier, and Nenad Ban. Architecture of the human mTORC2 core complex. *eLife*, 7:1–12, 2018.
- [260] Christopher H.S. Aylett, Evelyn Sauer, Stefan Imseng, Daniel Boehringer, Michael N. Hall, Nenad Ban, and Timm Maier. Architecture of human mTOR complex 1. *Science*, 351(6268):48–52, 2016.
- [261] Haijuan Yang, Xiaolu Jiang, Buren Li, Hyo J. Yang, Meredith Miller, Angela Yang, Ankita Dhar, and Nikola P. Pavletich. Mechanisms of mTORC1 activation by RHEB and inhibition by PRAS40. *Nature*, 552:368–373, 2017.
- [262] Yair Gat, Jan Michael Schuller, Mahesh Lingaraju, Elisabeth Weyher, Fabien Bonneau, Mike Strauss, Peter J Murray, and Elena Conti. InsP6 binding to PIKK kinases revealed by the cryo-EM structure of an SMG1 – SMG8 – SMG9 complex. *Nature Structural & Molecular Biology*, 26:1089–1093, 2019.
- [263] Li Zhu, Liang Li, Yilun Qi, Zishuo Yu, and Yanhui Xu. Cryo-EM structure of SMG1–SMG8–SMG9 complex. *Cell Research*, 0:1–8, 2019.
- [264] Yingli Sun, Xiaofeng Jiang, Shujuan Chen, Norvin Fernandes, and Brendan D Price. A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *Proceedings of the National Academy of Sciences*, 102(37):13182–13187, 2005.
- [265] Yingli Sun, Ye Xu, Kanaklata Roy, and Brendan D Price. DNA Damage-Induced Acetylation of Lysine 3016 of ATM Activates ATM Kinase Activity. *Molecular and Cellular Biology*, 27(24):8502–8509, 2007.

## List of abbreviations

°C	degrees Celcius
53BP1	p53-binding protein 1
6-FAM	6-carboxyfluorescein
6-his	6-histidine-tag
Å	Ångstrom
aa	amino acid
ABC	ATP Binding Cassette
AFM	atomic force microscopy
alt-NHEJ	alternative non-homologous end joining
AMPPNP	adenosine 5'-( $\beta$ - $\gamma$ -imido)triphosphate
ASTRA	ASsembly of Tel, Rvb and Atm-like kinase
A-T	Ataxia-Telangiectasia
ATLD	Ataxia-Telangiectasia like disease
ATM	Ataxia-Telangiectasia mutated
ATP	adenosine triphosphate
ATP $\gamma$ S	adenosine 5'-( $\gamma$ -thio)triphosphate
ATR	Ataxia telangiectasia and Rad3 related
ATRIP	ATR Interacting Protein
BER	Base Excision Repair
BLM	Bloom syndrom protein
bp	base pair
BRCA1/2	breast cancer 1/2
BRCT	breast cancer carboxy terminus
CCM	<i>Chaetomium thermophilum</i> culturing medium
CHK	cyclin dependent kinase
cNHEJ	classical non-homologous end joining
COSMIC	Catalogue Of Somatic Mutations In Cancer
cryo-EM	cryo-electron microscopy
CST	Cdc13-Stn1-Ten1
C-terminus	carboxy terminus
CTF	contrast transfer function
CtIP	CtBP-interacting protein
DDR	DNA damage response
D-loop	displacement loop
DNA	deoxyribonucleic acid
Dna2	DNA replication ATP-dependent helicase/nuclease 2
DNA-PKcs	catalytic subunit of DNA-dependent protein kinase
DSB	DNA double strand break
dsDNA	double stranded DNA
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
ETAA1	Ewings tumor-associated antigen 1
Exo1	exonuclease 1
FANC	Fanconi Anaemia
FAT	FRAP-ATM-TRRAP domain

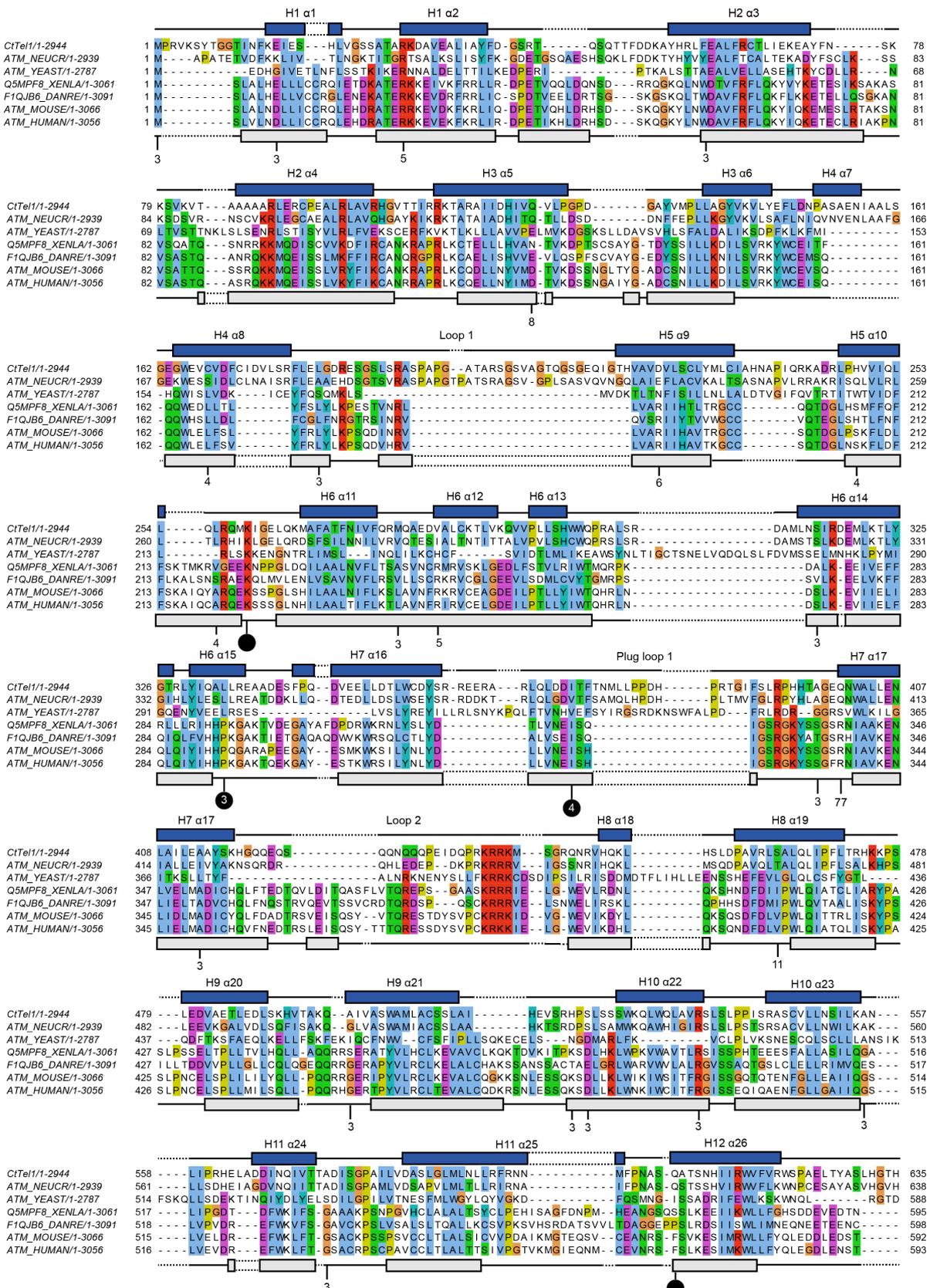
FATC	C-terminus of FAT domain
FATKIN	FAT and kinase domain
FHA	forkhead associated
FRB	FKBP12-Rapamycin binding
FSC	Fourier shell correlation
fw	forward
G phase	gap phase
GST	Glutathione S-transferases (GSTs)
h	hour
HEAT	Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR1
HPLC	high performance liquid chromatography
HR	homologous recombination
Ig	immunoglobulin
INO80	INOsitol requiring 80
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
IR	ionizing radiation
$K_D$	dissociation constant
kDA	kilo dalton
keV	kilo electron volt
LB	lysogeny broth
LBE	LST8 binding element
LST8	Lethal with Sec Thirteen
mAB	mouse antibody
MDC1	Mediator Of DNA Damage Checkpoint 1
Mec1	mitosis entry checkpoint 1
Mg	magnesium
min	minute
MMEJ	micro-homology mediated end joining
MMR	mismatch repair
MMS	methyl methanesulfate
Mn	manganese
M phase	mitotic phase
Mre11	meiotic recombination 11
MRN	Mre11-Rad50-Nbs1
MRX	Mre11-Rad50-Xrs2
mTOR	mammalian target of rapamycin
mTORC1	mTOR complex 1
mTORC2	mTOR complex 2
NBD	nucleotide binding domain
Nbs1	Nijmegen breakage syndrome 1
NER	nucleotide excision repair
Ni <sup>2+</sup> -NTA	nickel nitrilotriacetic acid
NMD	nonsense mediated decay
N-terminus	amino terminus

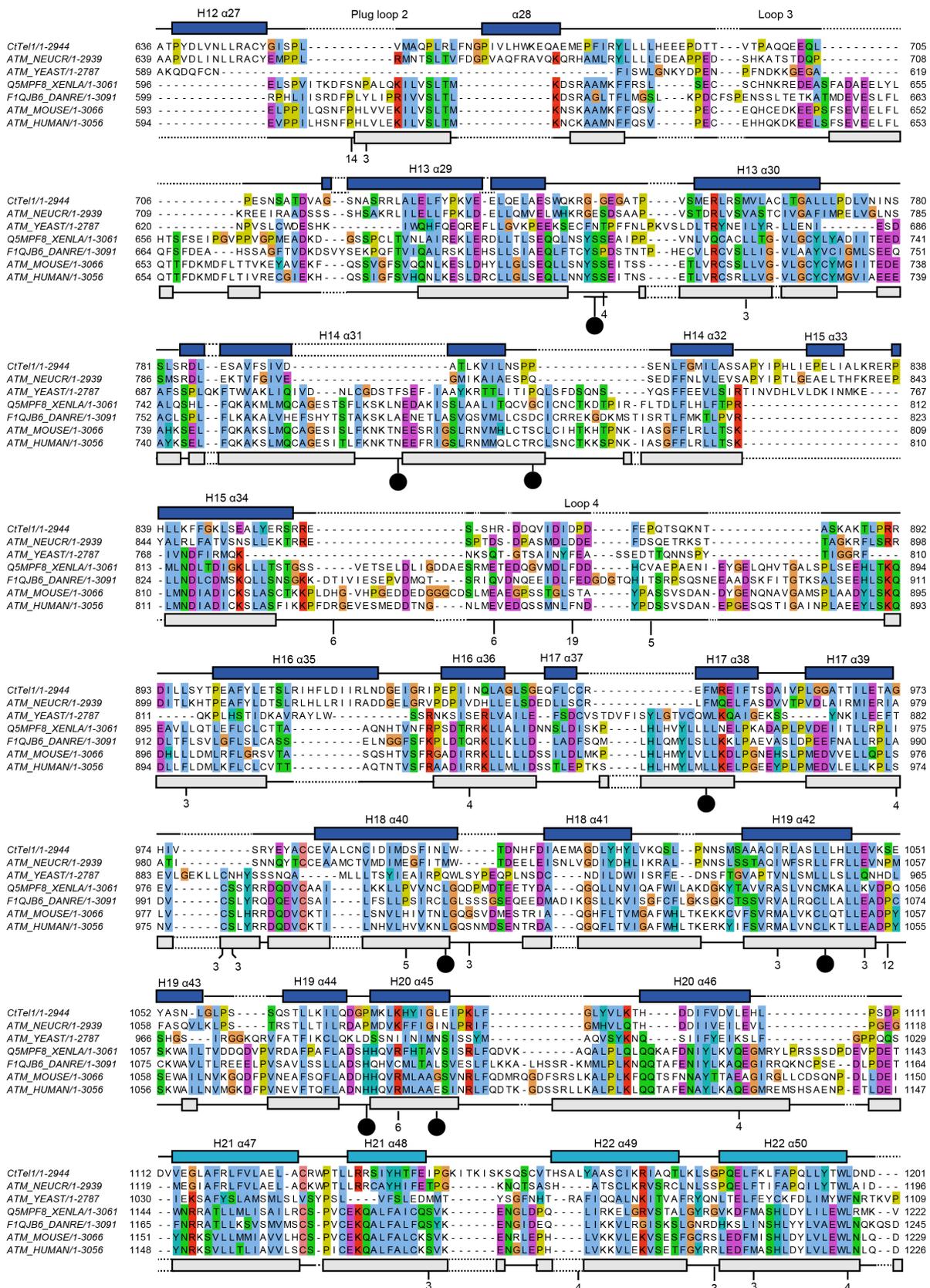
NuA4	nucleosome-acetyltransferase of histone H4
OD <sub>600</sub>	optical density at 600 nm
ORF	open reading frame
PAGE	polyacrylamide gelelectrophoresis
PBAF	Polybromo-associated BAF complex
PCR	polymerase chain reaction
PDB	Protein Databank
pH	potential of hydrogen
PI-3 kinase	phosphatidylinositol-3
PIKK	PI-3 kinase-like kinase
P-loop	phosphate-binding loop
PNK	polynucleotide kinase
PP1 $\gamma$	protein phosphatase $\gamma$ 1
PP2A	protein phosphatase 2
PRD	PIKK regulatory domain
PTM	post-translation modification
PVDF	polyvinylidene difluoride
Rad50/51	radiation sensitive 50
Rheb	Ras homolog enriched in brain
Rif1	Replication Timing Regulatory Factor 1
R-loop	RNA displacement loop
RNF8	Ring Finger Protein 8
ROS	reactive oxygen species
RPA	replication protein A
rpm	rotations per minute
RT	room temperatute
rv	reverse
SAGA	Spt-Ada-Gcn5-acetyltransferase
SDS	sodium dodecyl sulfate
sec	second
SMC	structural maintenance of chromosomes
SMG1	Suppressor with morphogenetic effect on genitalia 1
SSB	single strand break
S phase	synthesis phase
Spo11	SPOrulation 11
ssDNA	single stranded DNA
SWI/SNF	SWItch/Sucrose Non-Fermentable
TCR	transcription-coupled repair
Tdt	terminal deoxynucleotidyl transferase
Tel1	Telomere length regulation protein 1
TEV	tobacco etch virus protease
T-loop	telomere loops
TLS	translesion DNA repair
TopB1	DNA topoisomerase 2-binding protein 1
TOR	target of rapamycin
Tra1	Transcription-associated protein 1
TRRAP	Transformation/transcription domain-associated protein
TRIS	tris(hydroxymethyl)aminomethane
UPF1	Up-Frameshift Suppressor 1
UV	ultraviolet

V	volt
V(D)J	variable, diversity, joining
Wip1	Wild-Type P53-Induced Phosphatase 1
WRN	Werner syndrome ATP-dependent helicase
wt	wild type
X-gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside
XLF	XRCC4-like factor
XRCC4	X-ray repair cross-complementing protein 4
Xrs2	X-ray sensitive 2
Zn	zinc

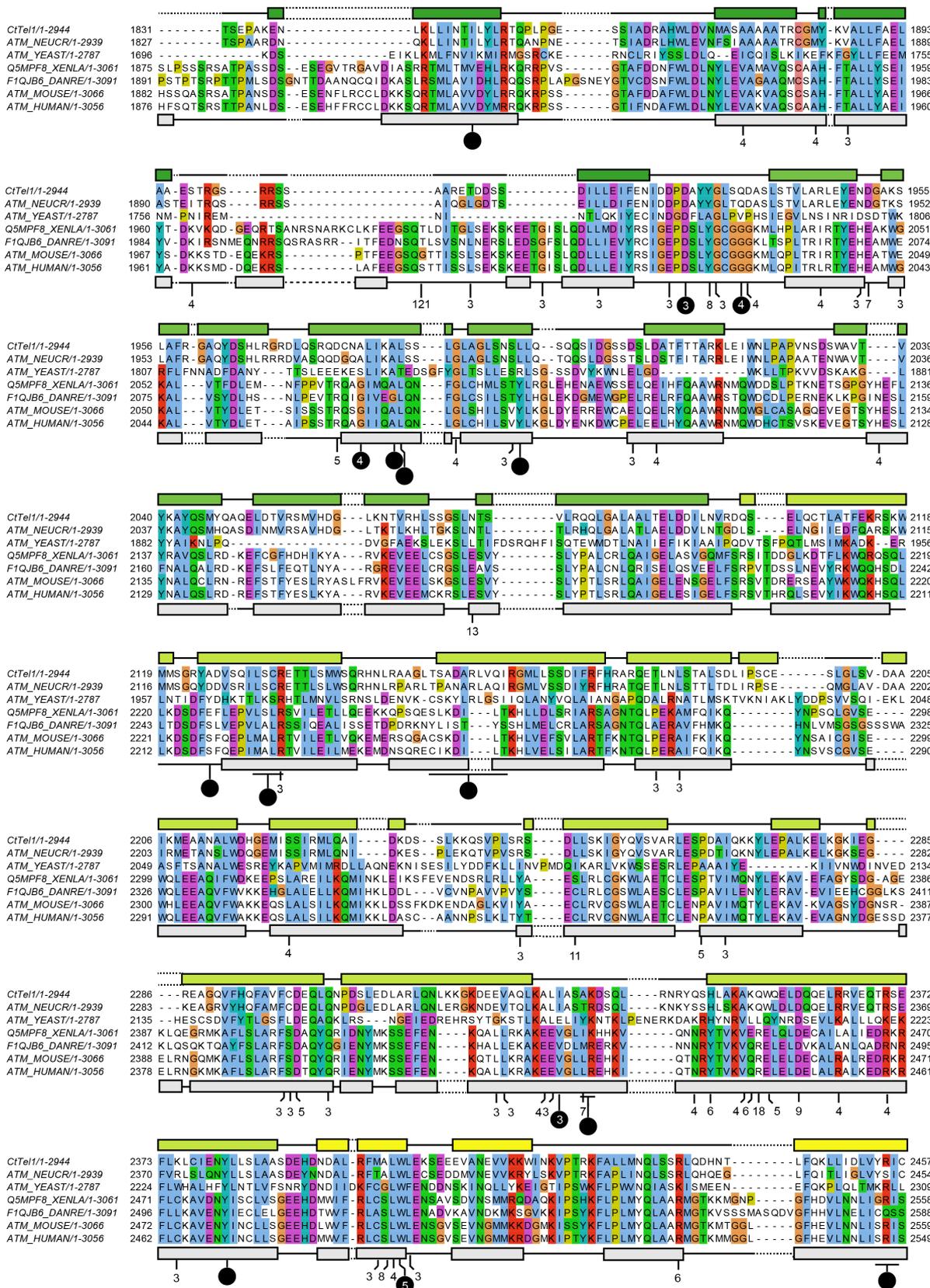
## Appendix: Multiple sequence alignment of ATM kinases

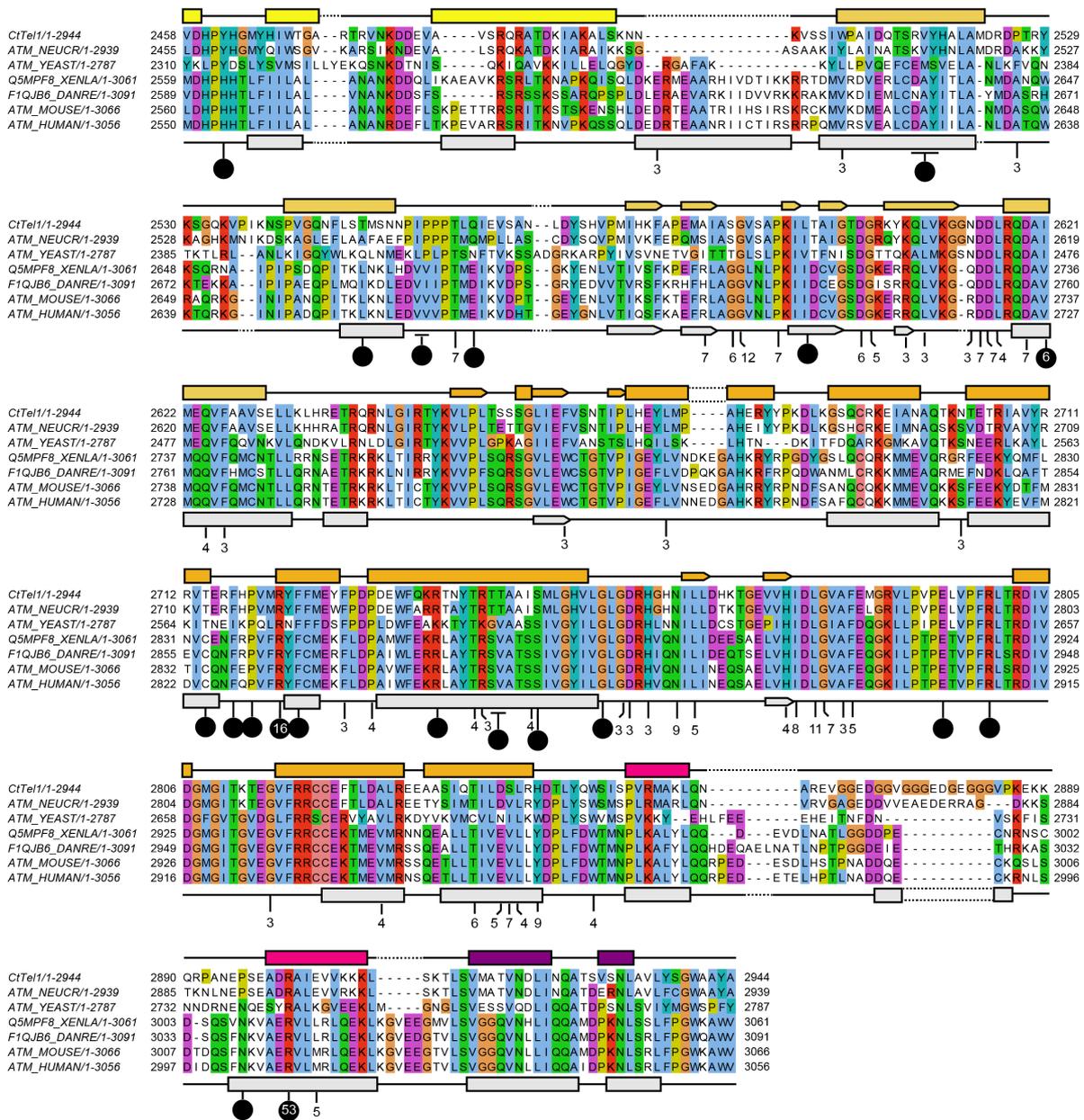
Alignment of ATMs from different species, created using MAFFT-E-INS-I and visualized using JalView.<sup>215,216</sup> Secondary structure elements displayed for ctTel1 are based on the ctTel1 structure and for hsATM on secondary structure predictions using PsiPred. For ctTel1, the domains are indicated with the same colour scheme as in Figure 3.10. On the hsATM sequence, cancer mutations from the COSMIC database are mapped by the numbers. These also indicate the frequency in which they occur in the database. A-T mutations, as listed in the UniProt database, are indicated by black circles. Black circles with a white number indicate that a cancer mutation overlaps with an A-T mutation.











## Acknowledgements

First of all, I would like to thank my PhD supervisor Prof. Dr. Karl-Peter Hopfner for giving me the opportunity to do a PhD in his lab and for allowing me to work on such a challenging and exciting project.

I would also like to thank Prof. Dr. Roland Beckmann and Dr. Boris Pfander for being part of my thesis advisory committee and sharing their valuable support and advice.

Furthermore, many thanks go to Christian Linke-Winnebeck for introducing me to the project. You did an amazing job during your postdoc and without all the groundwork you provided, there would not be a structure of ctTel1 now. Also many thanks for helping with the paper, despite having left the lab years before and having many other obligations. Although we only overlapped for a few months, it was a great pleasure to work with you.

Many thanks also go to Sebastian Eustermann. Thank you for sharing your expertise on cryo-EM data processing, your enthusiasm for science, and especially for having faith that the project would work out.

I also would like to thank Katja Lammens for all the indispensable help with the microscopy and the structure building. Many thanks as well for smuggling a ctTel1 grid into the newly-built Titan to collect one of the first (non-ribosomal) datasets.

I also thank Joe Bartho for help with microscopy (even at the most impossible times of the day) and Dirk Kostrewa for help with building and refining the structure when there was a lot of pressure on the project.

Of course, many thanks also go to all members of the DNA repair subgroup, past and present, but I would like to mention a number of people in particular. Rob Byrne: thank you for introducing me to cryo-EM data processing, and your endless patience with me when I once again had forgotten my linux password. Hinnerk Saathoff: thank you for helping me with setting up the biochemical assays. Thanks also for all the Schoki and introducing me to FM4. Claudia Litz: thanks for helping me find my way in the lab and in the project. Lisa Käshammer: thank you for all the conversations, serious and (usually) somewhat less serious, and the slow salad lunches.

Furthermore I would like to thank all the non-DNA repair members of the Hopfner lab as well for providing a pleasant working atmosphere and for bringing so many cakes.

Many thanks also go to Petra Runge-Wollmann as well for making the GRK1721 such a fantastic graduate school, consisting of a great network of PhD students all over Munich. I will fondly remember all the courses and retreats of the program, as these really added a lot to my PhD.

Special thanks also go out to the Boehringer Ingelheim Fonds, not only for the financial support but especially for the amazing network of people, and the inspiring events, after which I always had new ideas and fresh motivation.

Finally, I would like to thank my friends and my family, nearby and far away, for always supporting me and reminding me there is a world outside of the lab as well.