The chromatin remodeller ATRX: A repeat offender in human disease

David Clynes¹, Douglas R. Higgs¹ and Richard J. Gibbons¹

1. MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine,
University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK

Key words: ATRX, G4-quadruplex DNA, DNA replication, Alternative Lengthening of Telomeres
Abstract

The regulation of chromatin structure is of paramount importance for a variety of fundamental nuclear processes including; transcription, DNA repair, replication and recombination. The ATP-dependent chromatin-remodeling factor ‘ATRX’ (alpha thalassemia/mental retardation X-linked) has emerged as a key player in each of these processes. Exciting recent developments suggest that ATRX plays a repertoire of key roles at tandem repeat sequences within the genome; including the deposition of a histone variant, DNA replication, and the suppression of a homologous recombination based pathway of telomere maintenance. Here we provide a mechanistic overview of the role of ATRX in each of these processes, discuss how they may be connected and their relationship to human disease.

ATRX: From thalassemia to cancer

To understand the molecular mechanisms underlying many human diseases we need to consider DNA in the context of chromatin. Chromatin consists of 147 bp of DNA wrapped around a histone octamer which is then subsequently assembled into compacted higher order structures. The role of chromatin extends far beyond that of a simple packaging tool, its structure is dynamically regulated allowing for the modulation of a wide variety of biological processes. The importance of chromatin structure in human disease has become increasingly apparent through the identification of human diseases caused by mutations in genes that encode proteins required for the remodeling of, or epigenetic modification of chromatin. One such disorder ‘ATR-X syndrome’ is caused by mutations in the ATRX gene, of which approximately 127 unique mutations have been identified. ATRX encodes a
chromatin remodeler (ATRX) which is a 280 kDa protein that includes an unusual N-terminal plant homeodomain (PHD) designated the ATRX-DNMT3-DNMT3L (ADD) [1] domain owing to its similarity to a protein region found in this group of methyl transferases [2-4]. Located at the C-terminus are seven helicase subdomains, conferring ATPase activity and identifying ATRX as a snf2 family member of chromatin associated proteins, many of which characteristically slide, remodel or remove histones in *in vitro* assays [5]. It is now known that ATRX, in collaboration with its interaction partner DAXX, functions as a histone chaperone complex for the deposition of the histone variant H3.3 [6-9].

ATR-X syndrome is characterised by a variety of clinical features that include mental retardation, facial, skeletal and urogenital abnormalities as well as mild α-thalassemia [10, 11]. The latter is attributable due to reduced expression of the alpha globin genes located on chromosome 16. ATRX was hence considered to be an X encoded *trans*-acting factor that facilitates the transcription of a diverse repertoire of disparate genes. Subsequent studies in several model organisms have since uncovered defects in multiple important cellular processes upon perturbation of ATRX function including, defective sister chromatid cohesion and congression [12, 13], telomere dysfunction [14] and aberrant patterns of DNA methylation (at ribosomal DNA, subtelomeric and heterochromatic repeats [15]). Exciting recent developments have also implicated loss of ATRX function in a specific subset of malignancies that depend on a telomerase independent-based pathway of telomere maintenance, ‘the Alternative Lengthening of Telomeres (ALT)’ pathway [16-18].

Such a broad spectrum of pathologies raises interesting questions as to whether the underlying molecular mechanisms behind these pathologies are interrelated. Here
we address this question by outlining the recent advances made in our understanding of ATRX cellular function and what this tells us about the role of chromatin in human disease.

**ATRX: a repeat offender on heterochromatin**

A prominent clue to the physiological role of ATRX has come from studying the distribution of ATRX in the nucleus. Indirect immunofluorescence studies have revealed that ATRX has a strong preference for binding to repetitive heterochromatic regions; at rDNA repeats, telomeric repeats, pericentric DNA repeats and within promyelocytic leukaemia (PML) bodies [15, 19, 20]. The advent of deep sequencing technologies expanded on these observations by showing that ATRX has an overwhelming preference to bind tandem repetitive (TR) DNA, especially G-rich repeats as well as CpG islands (regions of the genome with a high frequency of unmethylated cytosine guanine dinucleotides) [21].

Further work has subsequently given mechanistic insight as to how ATRX is recruited to heterochromatin. Yeast two hybrid screens have shown that ATRX interacts with a murine homologue (mHP1α) of the Drosophila Heterochromatic protein HP1 [22]. Furthermore, targeting of ATRX to pericentric heterochromatin was found to be dependent on trimethylation of histone H3 at Lys9 on its N-terminal tail. Knockout of the methyltransferases Suv39H1 and Suv39H2, responsible for ‘writing’ this modification completely abrogate its recruitment [23]. An appealing hypothesis to emerge from these two observations was that HP1 might recruit ATRX to pericentric heterochromatin indirectly by serving as a protein scaffold, facilitating the recruitment of ATRX through binding of H3 K9me3 via its N-terminal chromodomains [24]. A
major caveat to this hypothesis, however, is that it would provide limited specificity
given that HP1 is a widely distributed, constitutive component of heterochromatin
with a large repertoire of binding partners [25]. Recent work by Eustermann et al.,
and lwase et al., elegantly addresses this question by showing that in addition to
binding HP1, ATRX binds directly to the histone H3 tail. Importantly, specificity for
this interaction is governed through two distinct binding pockets located in the N-
terminal ADD domain, one accommodating unmodified Lys4 and the other restricted
to di- or trimethylated Lys9. This allows for a combinatorial readout of K4me0 and
K9me3 (or K9me2)* on the histone H3 N-terminal tail, with recruitment to K9me3
further enhanced by the previously characterised interaction with HP1 [26, 27]
(Figure 1). ATRX recruitment to heterochromatin thereby adds credence to the
existence of a multi-faceted ‘histone code’ that is ‘read’ by a combination of
multivalent effector-chromatin interactions. The importance of the ATRX ADD
domain in the underlying etiology of ATR-X syndrome is highlighted by the finding
that it serves as a ‘hot spot’ for syndrome- associated mutations [28].

Recruitment of ATRX to its target sites may not, however, be restricted to its
interaction with the histone H3 tail. The favoured target of ATRX, G-rich tandem
repetitive DNA, has the potential to adopt a range of non-canonical DNA
conformations. Of particular note, approximately 50% of ATRX target sites, including
telomeres, are predicted to adopt the G-quadruplex conformation [21] and there is
mounting evidence that these structures form in vivo [29-31]. Electrophoretic mobility
shift assays have demonstrated that recombinant ATRX protein preferentially binds
to G4- structured DNA over an unfolded telomere-derived sequence [21], strongly
suggesting ATRX may directly bind these DNA structures in vivo.
ATRX: a histone chaperone

Once recruited to its target sites how does ATRX ensure appropriate gene expression and/or maintain genomic stability? It is now known that ATRX, in combination with its interaction partner DAXX, facilitates incorporation of the histone variant H3.3 into pericentric, telomeric and ribosomal repeat sequences [6, 8, 9, 32]. DAXX functions as a highly specific histone chaperone, able to discriminate H3.3 from the other major histone H3 variants H3.1 and H3.2, whereas ATRX is involved in targeting DAXX to repetitive sequences and enhancing this deposition [6, 8, 32]. Understanding the function of H3.3 at ATRX target sequences is therefore likely to prove pivotal in understanding ATRX function in relation to human disease. In contrast to the canonical histones, H3.3 is synthesised throughout the cell cycle and predominantly undergoes replication-independent deposition [33, 34]. H3.3 was traditionally considered as a marker for highly dynamic, active chromatin and its deposition in heterochromatic regions might therefore seem somewhat counter-intuitive. Repetitive DNA, however, is predicted to be an inherently “unfriendly” substrate for nucleosomes, leading to rapid turnover. ATRX/DAXX-dependent H3.3 deposition may be necessary to stabilise a B-DNA conformation in tandem repetitive ATRX target sites. This raises the question of whether the pathologies observed upon loss of ATRX function could be at least partially attributable to an increase in structured DNA at these sites?
**ATRX and alpha thalassemia**

ATR-X syndrome is characteristically associated with alpha-thalassemia, attributable to the down-regulation in expression of the alpha-globin genes located close to the telomere of chromosome 16. ChIP-seq data suggest that tandem repetitive DNA does indeed have a crucial role to play in this phenomenon by showing that instead of binding directly to the alpha globin genes, ATRX localises several kb upstream within a TR sequence, denoted Ψζ. Strikingly, in ATR-X syndrome patients it was observed that the genes immediately proximal to this repeat were the most down-regulated, and those furthest away (~10Kb) the least perturbed. It was further noted that the size of the Ψζ TR varies between patients and the degree of thalassemia was proportional to the length of the Ψζ TR, explaining the variable degrees of thalassemia observed between patients with identical ATRX mutations [21]. The Ψζ TR adopts a G-quadruplex conformation *in vitro* and a plausible explanation for this observation is that the longer the TR sequence, the higher the probability that it will adopt a non-canonical secondary structure such as a G-quadruplex. Consistent with such a hypothesis previous studies have shown that the length of a G-quadruplex forming minisatellite, CEB1, is related to its instability upon deletion of the known G-quadruplex resolving helicase, Pif1, in yeast cells [35].

**ATRX and DNA replication**

Assuming such a hypothesis is correct, it raises the intriguing question as to how the presence of an aberrant DNA secondary structure may lead to the transcriptional silencing of genes located several Kb downstream of the TR DNA. Structured DNA, such as G-quadruplexes, are known to form barriers to several nuclear processes,
including DNA replication and transcription [31, 36, 37]. Consistent with this, recent studies have shown that loss of ATRX function is associated with replication defects, as identified by an increase in replication fork stalling, a prolongation in S-phase and concomitant accumulation of p53 and the γ-H2AX DNA damage marker, particularly at telomeres, a known site for G-quadruplex formation and ATRX binding [14, 38, 39]. Accumulating evidence is emerging that problems in replicating structured DNA can also cause perturbations in local gene expression. The accurate propagation of histone marks during chromosomal replication is thought to rely on the tight coupling of replication with the recycling of parental histones. Work by Sarkies et al., has shown that interruption of processive DNA replication at G4 sequences in cells lacking various G4 helicases results in this tight coupling being lost, leading to the biased incorporation of newly synthesised histones (associated with H4 N-terminal acetylation but not other marks of activation or repression). Genes that were previously active became repressed and conversely genes that were previously repressed became active [40, 41]. One could also envisage an alternative explanation attributable to epigenetic silencing arising from the creation of a double strand break, which is known to be generated upon collapse of a stalled replication fork (for comprehensive reviews see [42, 43]). Work by the Greenberg laboratory has recently demonstrated such a phenomenon by showing that the induction of a double strand break can lead to the silencing of a reporter gene some 4kb downstream from the break site through H2A ubiquitylation dependent pausing of RNAPII [44] (Figure 2). Further work is needed to dissect whether either of these events are indeed occurring at the alpha-globin locus in ATR-X syndrome, through
careful profiling of DNA replication through the associated Ψζ TR and the detection
of any associated epigenetic changes arising from this site.

Given that ATRX appears to have general roles in the deposition of a histone variant
and DNA replication, one might expect that the cellular consequences of losing
ATRX function would include a loss of genome stability and chromosomal
congregation defects. In the appropriate cellular context this does indeed appear to
be the case, with loss of ATRX associated with extensive genomic rearrangements,
DNA damage and telomere fusions [14, 18, 38]. Furthermore, loss of ATRX function
has been shown to result in both mitotic and meiotic defects, including defective
spindle formation, chromosome condensation and alignment [12, 13, 38]. A link
between defective replication and chromosome condensation has long been
established (for review see [45]) and it is therefore possible that these phenotypes
may, at least partially, be attributable to the DNA replication defects observed upon
loss of ATRX function.

**ATRX and Cancer**

Of particular recent interest is the finding that mutations in ATRX are associated with
a specific subgroup of tumours that employ a telomerase-independent mechanism of
telomere maintenance, termed the Alternative Lengthening of Telomeres (ALT)
pathway. To grow indefinitely, cancer cells must circumvent the progressive attrition
of telomeric DNA that inevitably occurs with each cell division, ultimately leading to
cellular senescence or crisis. In the majority of cancer cells this is potentiated
through re-activation of telomerase, a specialised reverse transcriptase that
synthesises the tandem TTAGGG hexanucleotide repeat by using its own RNA
subunit as a template [46, 47]. However, 10 – 15% of cancers lack detectable
telomerase and are instead reliant on a telomere maintenance pathway that depends
on homologous recombination (HR) between telomeric sequences (for extensive
reviews see [48, 49]). Strikingly, a recent study has documented that loss of ATRX is
a trait of 90% of in vitro immortalised ALT cell lines [18]. Moreover, mutations in
ATRX and/or DAXX have been identified in many tumours exhibiting ALT, including
pancreatic neuroendocrine tumours, pediatric glioblastomas, oligodendrioglomas
and medulloblastomas [16]. More recently additional mutations in H3.3 have been
identified in ALT-positive paediatric glioblastomas [17], strongly suggesting that the
ATRX/DAXX/H3.3 pathway has a role in normally suppressing the ALT pathway.

Increased stalling of replication forks, a known trigger of HR, at telomeric sites may
contribute to the initiation of aberrant telomeric recombination and activation of the
ALT pathway in the absence of functional ATRX. Telomeres are known to be
inherently difficult sites to replicate and resemble fragile sites [50] and the
stabilisation of G-quadruplex structures using a stabilising ligand has been shown to
perturb DNA replication at telomeres and trigger a DNA damage response [37]. This
implies that an increased presence of DNA secondary structure at telomeres might
lead to increased replication fork stalling. Furthermore, the unidirectional nature of
telomere replication [50] means that a converging fork cannot salvage a stalled
replication fork, presumably increasing the reliance on HR mediated fork restart
(Figure 2).

Several unanswered questions remain. Knockdown of ATRX in a telomerase positive
cell line is not sufficient to trigger the ALT pathway [18]. The cellular context,
including the telomerase-negative status of ALT cells, must therefore be an important additional factor in the development of the ALT pathway and more research will be required to determine what additional factors are required to trigger ALT in an ATRX null genetic landscape.

**Summary**

The study of ATR-X syndrome illustrates the benefit of studying rare genetic disorders and highlights the importance of chromatin remodellers in human disease. From gaining a detailed molecular insight as to how and where ATRX is recruited in the genome and identifying the ATRX/DAXX complex as a novel histone chaperone for H3.3 we have now begun to understand how ATRX regulates gene expression and gained insight into how ATRX functions as an important tumour suppressor. Surprisingly, insights gained from the study of one disorder can be applied to understand the role of ATRX in seemingly disparate pathologies, highlighting the importance of combining basic research with clinical genetics.

**Box 1.**

Recent research has also identified an interaction between ATRX and another histone variant, macroH2A1 [51]. MacroH2A1 is generally considered to associate with transcriptionally inert chromatin and is frequently found associated with heterochromatin [52]. Ratnamukar et al. show that in contrast to the role of ATRX/DAXX in histone H3.3 deposition, ATRX acts as a negative regulator of macroH2A chromatin association. Knockdown of ATRX resulted in an increase in the accumulation of macroH2A1 at telomeres and at the alpha globin cluster in a human erythroleukemia cell line, inferring that ATRX may additionally promote alpha-globin
expression by maintaining chromatin in an active configuration through sequestration of macroH2A1.

Acknowledgements

Thank you to Ricarda Gaentzsch and Barbara Xella for critical reading of the manuscript. This work was supported by the Medical Research Council and the University of Oxford.

Reference list


Figure 1. Recruitment of ATRX to chromatin. Schematic depicting binding of ATRX to histone H3 at heterochromatic sites through an interaction of the ADD domain with the histone H3 N-terminal tail, trimethylated at Lys9 and unmodified at Lys4. Recruitment is enhanced by a third interaction through HP1 that also recognises trimethylated Lys9 [23, 27]. ATRX also binds directly to G-quadruplex structured DNA (far right) [21]. Once recruited to its target sites ATRX, in combination with its interaction partner DAXX facilitate the deposition of the histone variant H3.3 [6, 8, 9, 32], which may maintain DNA in the B-form.
Figure 2. Stalling of DNA replication may lead to transcriptional silencing and aberrant telomeric recombination upon loss of ATRX function. The persistence of DNA secondary structures, particularly on the lagging strand, in the absence of functional ATRX may lead to replication fork stalling. Upon collapse of the fork a double strand break is generated which may result in the silencing of genes distal to the break site through RNF8 and RNF168 mediated ubiquitylation of H2A and the subsequent stalling of RNAPII [44]. Double strand breaks are a known trigger for homologous recombination mediated repair, which at telomeric sites may in turn trigger the ALT pathway in malignant cells lacking ATRX.
G-quadruplex formed on replication

Stalling of replication fork at G-quadruplex

ATRX

ATP

ADP

Collapse of replication fork and double strand break formation

RNF8

RNF168

Double strand break induced transcriptional silencing through H2A ubiquitylation dependent pausing of RNAPII

Aberrant homologous recombination - At telomeric sites, triggering of ALT