Structural basis for histone H3 acetylation by Rtt109

in complex with histone chaperones Asf1 and Vps75

Von der Naturwissenschaftlichen Fakultät der Gottfried

Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

Nataliya Danilenko, Diploma (Russia)

2020

SUMMARY

In eukaryotes, genetic material is stored as chromatin, a DNA-protein complex whose structure is tightly regulated. Histones are the main protein components of chromatin and their post-translational modifications (PTMs) influence several cellular processes, from gene expression to epigenetic regulation and inheritance. Histone acetyltransferases carry out the acetylation of histones, an abundant modification, the function of which is determined by the position of the modified residue. Histone chaperones can influence the specificity of the histone acetyltransferases, however, the underlying mechanism of this process is not fully understood.

Rtt109 is a fungal histone acetyltransferase which is essential for the acetylation of newly synthesized histone H3. Two distinct histone chaperones, Asf1 and Vps75, have been reported to alter its activity and specificity: while Asf1 is necessary for H3 K56 acetylation, Vps75 promotes acetylation of H3 K9, K23 and K27, which are located in a long disordered N-terminal tail of H3. Despite the availability of structures of Rtt109 in complex with Vps75, the mechanism of regulation of Rtt109 activity by Asf1 and Vps75 remains elusive.

In order to understand how Asf1 and Vps75 stimulate the Rtt109 activity towards specific substrates, I reconstituted *in vitro* the complex containing Rtt109, histones H3:H4 and both chaperones. With multi angle light scattering and nuclear magnetic resonance (NMR), I could show that the Vps75 dimer assembles a non-symmetric complex with one copy of Rtt109 and Asf1-bound histones. Using an integrative structural biology approach combining distance restraint information from NMR and low-resolution shape information from small-angle neutron scattering (SANS) data, I could obtain a structural model of this complex. The structure revealed that the chaperones form a bagel-shaped complex with Rtt109 and the histones, bringing the enzyme and the substrate together and positioning H3 K56 next to the Rtt109 active center. A combination of NMR data with biochemical experiments and computational studies, revealed that the flexible H3 tail is chaperoned by Asf1 and is guided towards the catalytic pocket of Rtt109 by both folded and unfolded structural elements of Vps75. These results, taken together with existing literature and further mutational studies, allowed me to propose a mechanism by which the histone chaperones promote acetylation of the disordered H3 N-terminal tail.

Keywords: histone acetylation, histone chaperone, disorder, structural biology.

ZUSAMMENFASSUNG

In Eukaryonten wird das genetische Material im Chromatin gespeichert, diese DNA-Protein Komplexe sind streng reguliert. Histone sind die wichtigsten Proteinkomplexe des Chromatins. Posttranslationale Proteinmodifikationen (PTMs) beeinflussen zentrale zelluläre Prozesse, wie die Expression von Genen, epigenetische Regulationen und die Vererbung von Eigenschaften. Eine eine häufig auftretende PTM ist die Acetylierung. Diese Modifikation wird durch Histon- Acetyltransferasen katalysiert, dabei ist die Position der Modifikation von entscheidender Bedeutung für die Regulation. Histon-Chaperone haben einen Einfluss auf die Spezifität der Histon-Acetyltransferasen, dabei ist der zugrundeliegende Prozess noch nicht vollständig aufgeklärt.

Rtt109 ist eine in Pilzen vorkommende Histon-Acetyltransferase welche essenziell für die Acetylierung von neu synthetisierten Histon H3 Proteinen ist. Zwei unterschiedliche Histon-Chaperone, Asf1 und Vps75, haben einen Einfluss auf die Aktivität und Spezifität von Rtt109. Asf1 ist notwendig für die Acetylierung von H3 K56, dagegen acetyliert Vps75 H3 K9, K23 und K27, welche sich in dem unstrukturierten N-terminalen Teil des H3 Moleküls befindet. Ungeachtet der verfügbaren Strukturen von Rtt109 mit Vps75, ist der Mechanismus der Regulation von Rtt109 durch Asf1 und Vps75 nicht vollständig aufgeklärt.

Um den zugrundeliegenden Mechanismus aufzuklären, wie Asf1 und Vps75 die Aktivität von Rtt109 hinsichtlich bestimmter Substrate stimuliert, habe ich in der vorliegenden Arbeit den Komplex Rtt109 mit den Histonen H3:H4 und beiden Chaperonen in vitro rekonstruiert. Mit Hilfe von Mehrwinkellichtstreuung und Kernspinresonanzspektroskopie (NMR) konnte ich zeigen, dass das Vps75 Dimer sich als nicht-symmetrischer Komplex mit einer Kopie von Rtt109 und Asf1, in Verbindung mit Histonen assembliert. Die Verbindung von integrativ strukturbiologischen Ansätzen mit Entfernungsangaben durch NMR und gering-auflösender Strukturinformationen von Kleinwinkel - Neutronenstreuung (SANS), ermöglichte es mir ein Strukturmodell dieses Komplexes zu erstellen.

Abbreviations and symbols

bp: base pairs
CSP: chemical shift perturbation
CTAD: C-terminal acidic domain
CTD: C-terminal domain
D_{max}: maximum dimension of the particle
EM: electron microscopy
HAT: histone acetyltransferase
HDAC: histone deacetylase
HMQC: heteronuclear multiple-quantum coherence
HSQC: heteronuclear single-quantum coherence
IDR: intrinsically disordered region
I_{para}/I_{dia}: ratio of peak intensity in the paramagnetic and diamagnetic states
K(number)ac: N-ε-acetyl-lysine at position (number)
MALS: multi angle light scattering
MD: molecular dynamics
M_w: molecular weight
NMR: nuclear magnetic resonance
NOESY: nuclear Overhauser effect spectroscopy
NTD: N-terminal domain
o-RMSD: orientational root-mean-square deviation
PRE: paramagnetic relaxation enhancement
PTM: post-translational modification
R_2^{dia}: transverse relaxation rate
R_2^{diaH}: ^{1}H single-quantum coherence transverse relaxation rate
R_2^{diaHC}: ^{1}H-^{13}C multiple-quantum coherence transverse relaxation rate
R_g: radius of gyration
RVAH: Rtt109–Vps75$_2$–Asf1–H3:H4
SANS: small-angle neutron scattering
SAS: small-angle scattering
SEC: size exclusion chromatography
SLD: scattering length density
TROSY: transverse relaxation optimized spectroscopy
Γ_2: paramagnetic contribution to the transverse relaxation rate
τ_C: correlation time for the electron-nucleus interaction vector
WT: wild-type

1. Introduction

1.1. Architecture and mechanism of nucleosome assembly.

In eukaryotes, genetic information is present in the form of chromatin, a DNA-protein complex which ensures packing of long DNA molecules. The main protein components of chromatin are histones, small alkaline proteins which are able to form nucleosomes, around which DNA is wrapped. The first level of chromatin organisation is the formation of DNA-histone complexes, the nucleosomes, which are regularly spaced along the DNA strand. After this step, DNA can be further compacted into higher order structures (Fig. 1.1.1.)
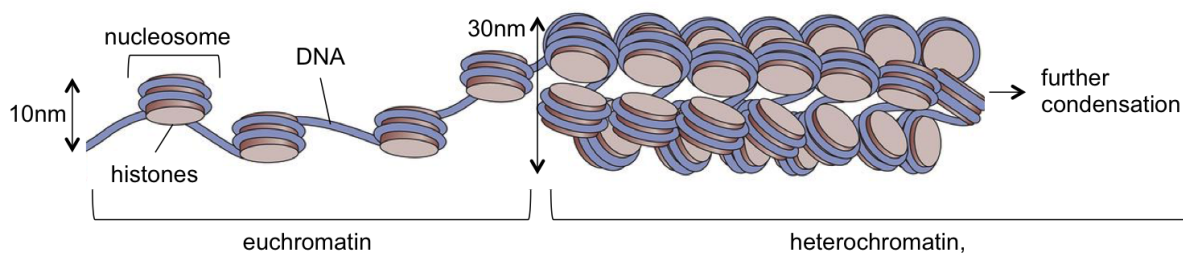
Figure 1.1.1. DNA packing. The first level of DNA packing is the assembly of nucleosomes, namely the formation of so-called "beads on a string"; DNA can be further compacted with the aid of histone H1 and scaffold proteins into 30 nm fibres that can, in turn, form condensed loops. The figure is adapted from (Kristie, 2016).

DNA compaction provides the cell with a way to regulate gene expression by regulating the accessibility of protein-coding regions. The level of packing determines the accessibility of DNA to the protein machinery and therefore regulates DNA replication, repair and gene expression. Cellular DNA is packed inhomogeneously and contains both the tightly packed, transcriptionally-inactive heterochromatin and the more "open" transcriptionally-active euchromatin. Differential gene expression is regulated at the nucleosome-DNA level by selective methylation of DNA, inclusion of specific histone variants into nucleosomes and, most importantly, by histone post-translational modifications (PTMs). Additionally, epigenetic euchromatin and heterochromatin markers may be inherited from a cell to its progeny.

Nucleosomes consist of ~146 bp of DNA wrapped around eight core histones, one H3:H4 tetramer and two H2A:H2B dimers, and linker DNA (the DNA sequences between the nucleosomes, up to ~80 bp) (Luger *et al*, 1997). The linker DNA at the nucleosome entrance and exit can be stapled with the linker histone H1 to form a more compact structure.
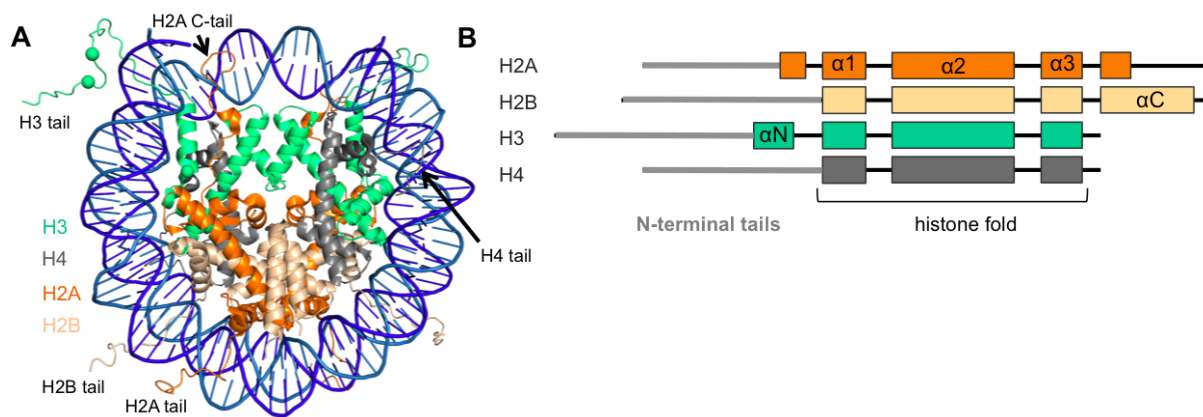
Figure 1.1.2. Structural organisation of the nucleosome and its components. A. Nucleosome core structure (PDB entry 1eqz). DNA is in blue, histones H3, H4, H2A and H2B are shown in green, grey, orange and beige, respectively. N-terminal tails of H2B and H3 exit through the minor groove formed by two DNA strands, the N-terminal tails of H2A and H4 exit without going through the nucleosomal DNA; H2A C-terminal tail is located at the entrance/exit point of the nucleosomal DNA. Accessible histone tails interact with various protein partners and also influence the higher order packing. **B**. Schematic representation of the structural domains of the histones. Colored boxes represent α-helices.

All core histones contain a histone-fold motif, which consists of three helices (α1, α2 and α3) connected by 2 short loops, and posses unfolded N- and C-terminal tails (Figure 1.1.2) (Xu *et al*, 2010; Arents & Moudrianakis, 1995). This conserved fold allows histones to adopt regular structures such as a heterodimer, tetramer and, ultimately, the octamer. The long unstructured tails, in turn, act as a platform for numerous post-translational modifications (PTMs), which play a role in nucleosome assembly, gene expression regulation and epigenetic mechanisms.

The composition of the nucleosomes and their positioning on the DNA is tightly regulated by the following functional groups of proteins:

1) Histone chaperones, which facilitate histone folding, transport and nucleosome assembly in an ATP-independent manner (Eitoku *et al*, 2008). As histones have a high positive charge at neutral pH, they are prone to aggregation and therefore require chaperones to handle them from the time they are synthesized until the incorporation into chromatin. The term "histone chaperone" was introduced by Ron Laskey for nucleoplasmin, a protein which prevents nonspecific interactions between H2A, H2B and DNA (Laskey *et al*, 1993). Nowadays histone chaperones are known to include over seven families of proteins, which are structurally diverse but share the ability to bind and shield histones from promiscuous interactions (Fig. 1.1.3).
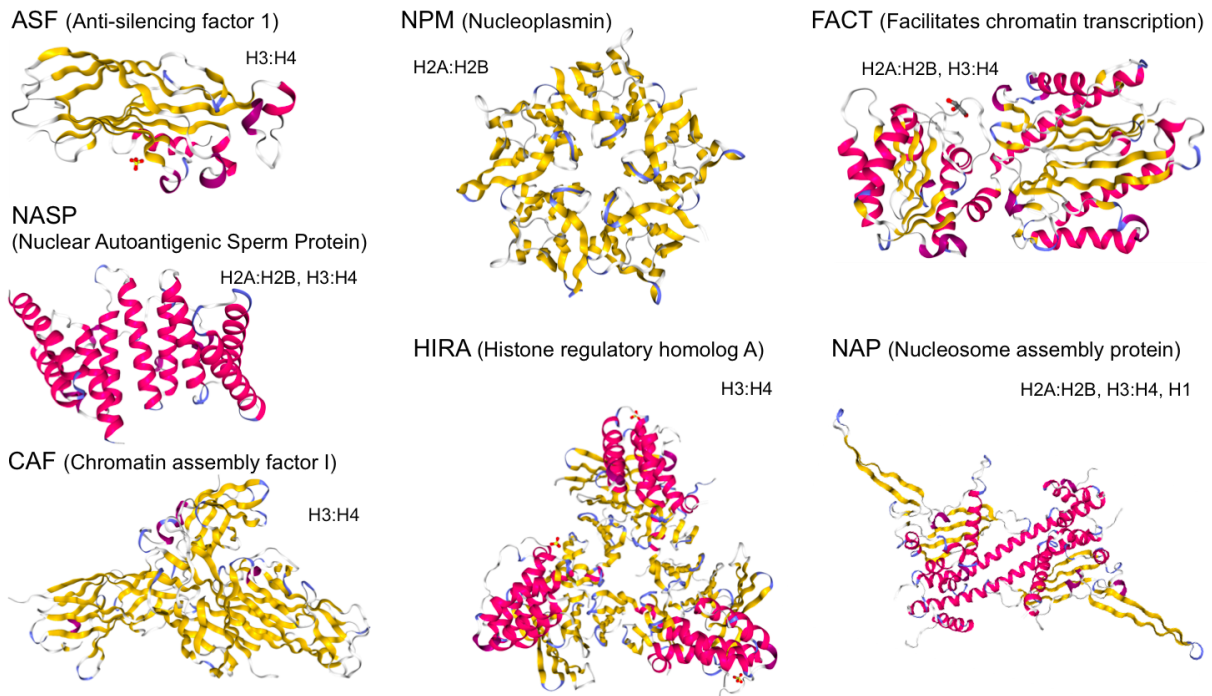
Fig. 1.1.3. Major families of histone chaperones. Histone chaperones are classified into families by sequence homology. They are structurally very diverse, ranging from a small single subunit (ASF) and homodimeric (NAP) proteins to large multisubunit complexes (HIRA, CAF-1, FACT). While some families show a general preference towards certain histones or specific histone variants, others include members that can interact with all classes of histones. The structures of the representatives of the chaperone families are shown in cartoon representation and are colored by the secondary structure elements (pink, α-helices; yellow, ß-sheets); in addition to the structured elements, many histone chaperones possess intrinsically disordered acidic regions.

2) Histone modifying enzymes, which introduce and remove PTMs. Histones are known to undergo at least 12 different kinds of PTMs: methylation, acetylation, phosphorylation, sumoylation, ubiquitination, poly-ADP-ribosylation, formylation, propionylation, butyrylation, citrullination, proline isomerization, and crotonylation (Tan *et al*, 2011). The unstructured N-terminal tails are the main, although not sole, platforms for these modifications, which can be mutually exclusive or complementary. Specific PTMs of histones act to change the interaction mode between the histones and the DNA, as well as recruit various protein partners (known as PTM readers) to their location. Histone PTMs (the "histone code") are the main players in epigenetic mechanisms, as the number of PTM combinations significantly exceeds the number of histone variants.

3) Chromatin remodelers, which rearrange the nucleosomes in an ATP-dependent manner. The currently known chromatin remodelers are classified into 5 families; all of them contain an ATPase domain and act to alter chromatin architecture. Among other functions, chromatin remodelers promote nucleosome sliding and regulate nucleosome spacing as well as eviction and exchange of histone variants.

The major biological processes that depend on nucleosome biogenesis involve DNA replication, transcription, and DNA repair. The general pathway of nucleosome assembly in eukaryotes is shown in Fig. 1.1.4:
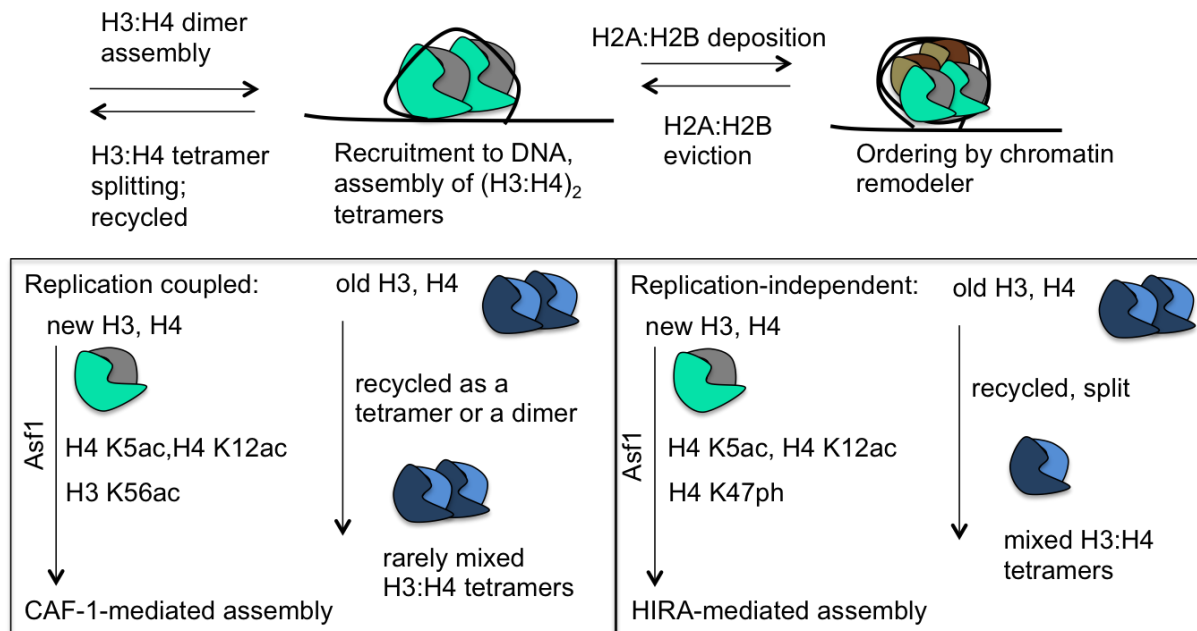
Figure 1.1.4. Key steps of nucleosome assembly. Histones H3 and H4 are shown in mint and grey (newly synthesized) or in dark blue and light blue (recycled), respectively; H2A and H2B are shown in brown. First, the (H3:H4)$_2$ tetramer is assembled onto the DNA, next, two H2A:H2B dimers are added. In both replication-coupled and replication-independent nucleosome assembly pathways old H3:H4 histones are recycled for the assembly. In replication-coupled nucleosome assembly, old H3:H4 are rarely split into dimers and are recycled as the tetramer; in replication-independent pathway, H3:H4 tetramers (if evicted) can be split and reassembled with both old and newly synthesized histones. "Mixed" tetramers consist of the H3:H4 dimers from previous nucleosomes and of the newly synthesized H3:H4 dimers. Newly synthesized H3 and H4 carry specific PTMs (H3 K56ac or H4 S47ph) required for their incorporation into nucleosomes by chromatin assembly complexes, CAF-1 and HIRA (conserved from yeast to humans). The bottom panels are partially adapted from (Liu & Churchill, 2012).

The assembly of the histone octamer follows a strict order: first, two H3:H4 dimers form a tetramer, the most stable part of the nucleosome, which is loaded onto DNA. The (H3:H4)$_2$ tetramer then acts as a scaffold to bind two H2A:H2B dimers. Nucleosome disassembly follows the reverse sequence, whereby, H2A:H2B dimers are removed first; then H3:H4 tetramers are evicted and recycled.

A proper regulation of histone recycling and nucleosome assembly is essential for the inheritance of epigenetic information. The DNA methylation pattern can be copied from the original strand after replication by the DNA-methyltransferase DNMT1, which exhibits preference for the hemi-methylated DNA, and is subsequently "proofread" by DNMT3 (Holliday & Pugh, 1975; Jones & Liang, 2009). The preservation and dynamics of the histone

code, however, require a more genuine regulation, which is fine-tuned for either replication-dependent or -independent nucleosome assembly; some of the basic mechanisms of such regulation are described below.

Histone recycling in transcription. During transcription-coupled nucleosome assembly (right panel, Fig. 1.1.4), chromatin remodeling depends on the degree of the transcription activity. In the genes that are transcribed at low to medium rate, H2A:H2B are found to be exchanged at a significantly higher rate than H3:H4. *In vitro* RNA polymerase II can progress through the DNA without complete displacement of the histone octamer, but requires the temporary dissociation of only one H2A:H2B dimer (Kulaeva *et al*, 2013). This mechanism ensures that H3:H4 tetramers are kept at their original positions and that the PTM pattern is maintained. However, at a high rate of transcription, nucleosomes are completely disassembled; recycled H3:H4 tetramers are split into dimers prior to nucleosome reassembly, resulting in mixed H3:H4 tetramers, consisting of both old and new H3:H4 dimers (Katan-Khaykovich & Struhl, 2011). In this case the PTMs are thought to be propagated by trans-modification, based on the PTM pattern of the old H3:H4 dimers located in the vicinity of the new ones.

Histone recycling in replication. During replication, the nucleosomes are disassembled completely. Pulse-chase experiments have revealed that, unlike in transcription, old H3:H4 tetramers do not split (Jackson, 1988); this observation has been supported by several other studies (Das & Tyler, 2013). Current data shows that the splitting of the H3:H4 tetramer depends on the H3 variant present in the tetramer: histone H3.1 (the major variant of mammalian H3) preserved the H3:H4 tetrameric form, while H3.3 promotes spitting of the tetramer into dimers during replication-coupled disassembly (Xu *et al*, 2010). This histone variant, which differs from H3.1 by only 5 amino acids, is associated with transcriptionally-active chromatin and is deposited onto the DNA during both replication and transcription (Ahmad & Henikoff, 2002). The difference in the behaviour of these histone variants suggests that the mechanisms of inheritance of histone PTMs might depend on the state of chromatin: the propagation of euchromatin marks requires H3:H4 tetramer splitting, while the heterochromatic modifications are preserved through retention of the original H3:H4 tetramers (Das & Tyler, 2013). While it is unclear whether these two modes result in a different precision of PTM conservation, it is clear that careful re-deposition of old H3:H4 (as dimers or tetramers) to their original position on the DNA is essential (Reverón-Gómez *et al*, 2018).

Regulation of inclusion of new H3:H4 into nucleosomes. The majority of newly synthesized histones is incorporated into chromatin during DNA replication (~85% of novel

histones included) and transcription of active genes (~10%) (Dion *et al*, 2007; Das & Tyler, 2013). Newly synthesized histones are targeted to the different nucleosome assembly pathways by a complex network of specific PTMs.

Nascent H3 and H4 are assembled into dimers and handled by the ASF1 chaperone, which is conserved throughout eukaryotes (Daganzo *et al*, 2003). Prior to nucleosome assembly, novel histones are acetylated in the cytoplasm at H4 K5 and K12 (Fig. 1.1.3). This diacetylation is introduced by homologous HATs in budding yeast and humans, and appears to be an abundant and conserved mark of a new histone H4 (Sobel *et al*, 1995). After the dimer assembly, acetylation of histone H4 at lysines 5 (H4 K5ac) and 12 (H4 K12ac) stimulates the import of H3:H4 into the nucleus (Alvarez *et al*, 2011). In addition, this diacetylation is associated with active chromatin (Turner, 1991) and was recently shown to play a role in the deposition of CENP-A onto centromeres (Shang *et al*, 2016). These acetylations do not appear to be specific for either the replication-dependent or -independent nucleosome assembly pathways (Liu & Churchill, 2012). H4 K5ac and H4 K12ac were also detected in histones co-purified with the chromatin assembly complex CAF-1, which incorporates the new histones into nucleosomes during replication. However, the double mutation of H3 K5, K12 is not lethal in *S. cerevisiae* and the nucleosome assembly still occurs, suggesting that nuclear import and other functions of these acetylations may be promoted by an alternative redundant mechanism (Ma *et al*, 1998).

After these general steps, histones are further modified depending on their specific function. Prominent H3/H4 modifications include phosphorylation of H4 S47 (H3 S47ph) and acetylation of H3 K56 (H3 K56ac), which are associated with HIRA/transcription-coupled and CAF-1/replication-coupled nucleosome assemblies, respectively.

In human cells, H4 S47ph promotes H3.3:H4 deposition onto DNA by increasing the affinity of H3.3:H4 to HIRA and simultaneously reducing its affinity to CAF-1 (Kang *et al*, 2011). While it is currently unclear whether the function of this modification is conserved in other eukaryotes, H4 S47ph is speculated to be inhibitory to transcription silencing and mutations of this residue were linked to deficiency in transcription in *S.cerevisiae* (Yu *et al*, 2011; Hainer & Martens, 2011).

Acetylation of H3 K56 is essential for the replication-coupled nucleosome assembly pathway in fungi (Li *et al*, 2008) and is crucial for the DNA damage response and genomic stability in both mammalians and yeast (Yuan *et al*, 2009; Tropberger & Schneider, 2013). The detected levels of H3 K56ac are lower in mammalian cells. Indeed, the exact pathway involving this modification in mammals is less understood than in yeast (Drogaris *et al*, 2012; Garcia *et al*, 2007). In *S.cerevisiae*, H3 K56ac has been directly linked to the incorporation of newly

synthesized H3 into chromatin. This function is carried out through the recognition of H3:H4 dimers by the histone chaperone Rtt106 (fungi-specific) and the increase of the histone binding affinity to CAF-1 (Fig. 1.1.5).
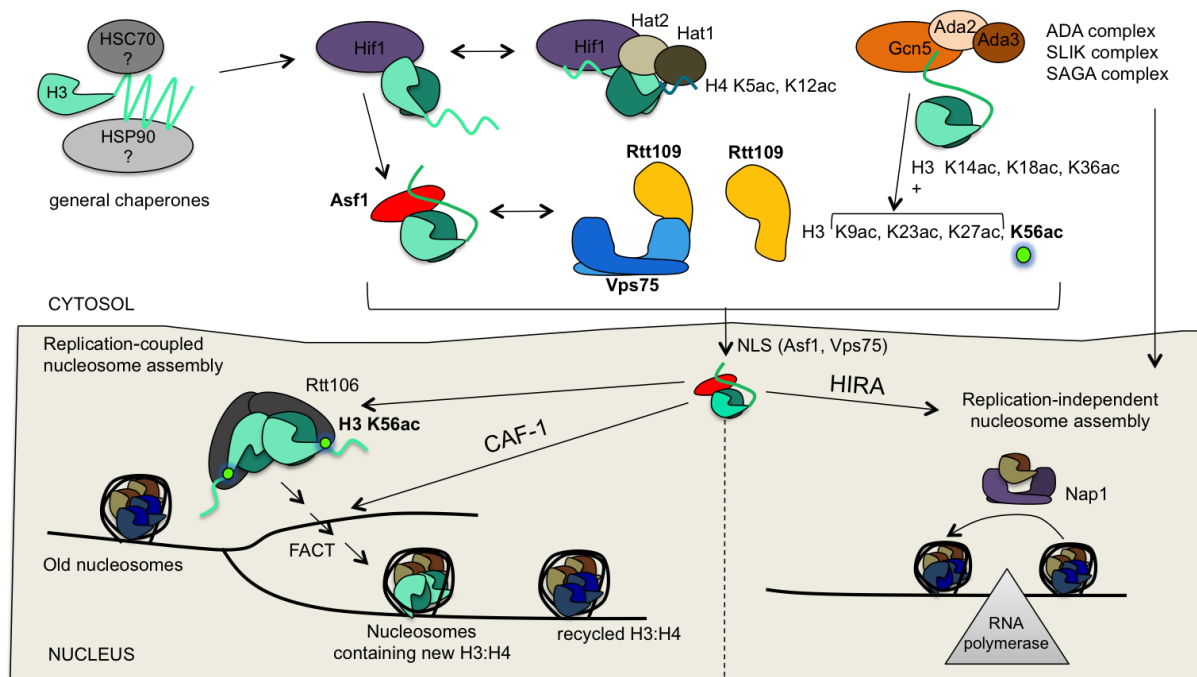
Figure 1.1.5. Pathway of incorporation of newly synthesized histone H3 in *S. cerevisiae*. Nascent histone H3 is first handled by general chaperones that assist its folding. After that it is handed over to the histone chaperone Hif1 (homolog of human sNASP), which facilitates the assembly of the H3:H4 dimer, and associates with Hat1, the histone acetyltransferase (HAT) that introduces the H4 K5ac and H4 K12ac (Campos *et al*, 2010). Hif1 transfers the H3:H4 dimer to Asf1 (homolog of human Asf1a and Asf1b) (Parthun *et al*, 1996). Asf1 aids H3 K56 acetylation by HAT Rtt109, which also acetylates other lysines in the H3 N-terminal tail. Asf1 supplies the H3:H4 dimers for both replication- and transcription-coupled nucleosome assemblies. However, H3 K56ac appears to be crucial only for the former. In this case, H3 K56ac is recognized by the downstream chaperone Rtt106 and H3:H4 dimers are assembled into tetramers and incorporated into the nucleosomes (Liu & Churchill, 2012; Burgess & Zhang, 2013).

In fungi, acetylation of H3 K56 is performed solely by HAT Rtt109 (Schneider *et al*, 2006; Tsubota *et al*, 2007). Over the past decade a number of studies focused on this enzyme due to its potential therapeutic importance (Rosa *et al*, 2010) and its intriguing acetylation mechanism. The functional and structural features of Rtt109 and its histone chaperone regulators, Asf1 and Vps75, are discussed in the sections below.

1.2. Rtt109: enzyme which requires two histone chaperones for the acetylation of H3.

Histone acetyltransferases (HATs) are the enzymes that transfer an acetyl group from acetyl-CoA to the ε-amino group of histone lysines. HATs are divided into more than four families based on the sequence conservation: GNAT (Gcn5-related N-acetyltransferases), MYST (MOZ, Ybf2/Sas3, Sas2, and Tip60), p300/CBP (p300, CREB-binding protein), Rtt109 (Regulator of Ty1 Transposition gene product 109) and others. While the p300/CPB and Rtt109 are metazoan and fungi-specific, respectively, the two biggest families, GNAT and MYST, are conserved from yeast to humans. Despite the fact that all HATs carry out the same reaction, the sequence homology between the different families is very low; HATs from different families contain multiple additional domains and display different specificity, selectivity and acetylation mechanisms.

Rtt109 is one of the latest discovered HAT families and remains one of the least studied overall. Rtt109 was identified in a genome-wide screen for genes required for *S. cerevisiae* resistance to ionizing damage and has been well-characterised since then for its role in resistance to DNA damage (Bennett *et al*, 2001; Wang *et al*, 2008). Despite a very limited sequence similarity to p300, the fold of the Rtt109 active site is similar to that of the HAT domain of p300 and the two proteins are considered to be distant orthologs. Interestingly, both enzymes use autoacetylation to regulate their activity (Thompson *et al*, 2004; Tang *et al*, 2008a). However, the autoacetylation and substrate acetylation mechanisms between the two HATs significantly differ: p300 autoacetylation is intermolecular and the acetylation of the substrate K is done with the catalytic tyrosine residue via a "hit-and-run" mechanism; Rtt109 employs the intramolecular autoacetylation and although the details of the acetylation mechanism of Rtt109 are unclear, Rtt109 active center does not contain any typical catalytic residues (Bazan, 2008).

Rtt109 consists of a globular folded domain and two disordered regions (the α2β5 loop and the C-terminal tail); the folded domain contains a conserved acetyl-CoA binding site and the active center (Fig. 1.2.1).
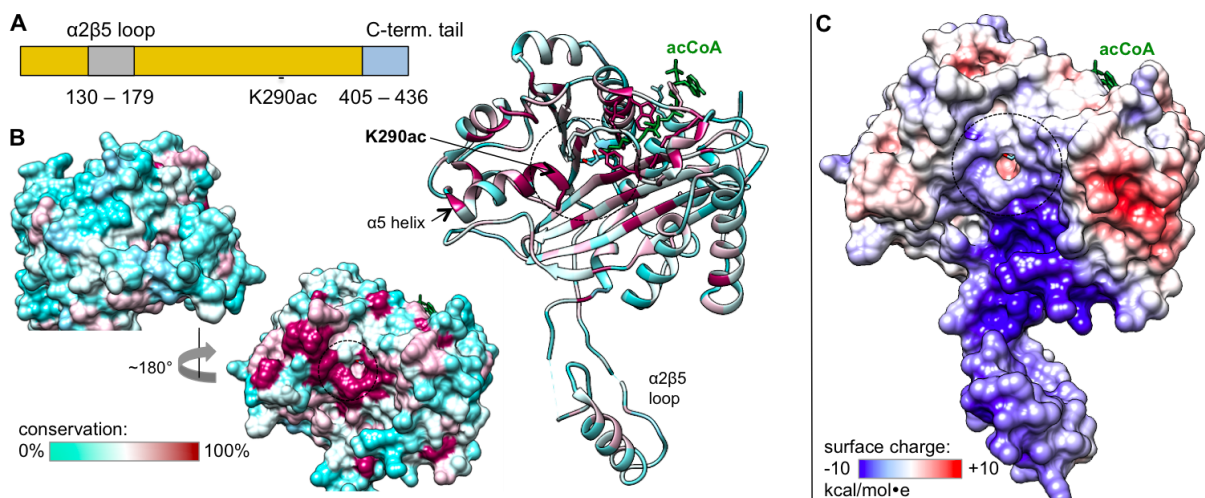
Figure 1.2.1. Structural organisation of Rtt109. A. Rtt109 contains two unstructured regions, the α2β5 loop required for Vps75 binding and a positively charged C-terminal tail. **B**. Conservation of Rtt109 in fungi. The sequence alignment was done for Rtt109 homologs from *S. cerevisiae, Candida albicans, Schizosaccharomyces pombe, Zygosaccharomyces bailii, Candida parapsilosis, Candida orthopsilosis*. The conservation is mapped on Rtt109 from PDB entry 3q33 (in complex with Vps75); on the left in surface representation only the core of Rtt109 is shown; on the right, in cartoon representation, both the core and the Vps75-bound α2β5 loop are shown. The dashed circle indicates the active center of Rtt109; K290ac is the auto-acetylated lysine residue. **C**. Surface charge representation of Rtt109. Surface electrostatic potential was generated with APBS (pH 7.5, 298 K). The active center of Rtt109 is apolar with a highly positively charged patch below.

Two X-ray structures of acetyl-CoA-bound Rtt109 were reported in 2008 (PDB entry 3qm0, Fig. 1.2.2 (Tang & Marmorstein, 2008); PDB entry 3cz7 (Stavropoulos *et al*, 2008)). The two structures are nearly identical, with a Cα-atoms RMSD below 0.5 Å. The small differences between the two structures map on the end of the α5 helix and the α6β8 loop, suggesting that these conserved regions are mobile and their flexibility might be required for the interaction with the histones and the histone chaperones or for the rearrangement of the active center (Fig. 1.2.2).

Interestingly, the catalytic site of Rtt109 is mostly apolar but features a positively charged patch next to it (in Fig. 1.2.1 C the positively charged patch is below the catalytic site). The substrate lysine is accommodated in the hydrophobic tunnel (the opening is visible in Fig. 1.2.1; the depths of the tunnel (~7 Å) corresponds well to the length of the extended lysine side chain (~6.8 Å)); the walls of the tunnel are built by the hydrophobic Rtt109 residues and by K or R aliphatic side-chains. Mutations of the Rtt109 residues in the tunnel or next to it result in a medium-to-complete loss of acetylation activity (Stavropoulos *et al*, 2008, Tang *et al*, 2008a).

Mutation of K290, the auto-acetylated lysine, also impairs acetylation catalysis: the K290R mutation decreases the k_{cat} ~100-fold *in vitro* and cannot rescue the *RTT109Δ* phenotype *in vivo* (Albaugh *et al*, 2011). Notably, the K290Q mutation has the same effect, showing that glutamine does not properly mimic the acetylated K290. Nearly 100% of Rtt109 purified from *E. coli* carries K290ac and the presence of this acetylation does not depend on the presence of Asf1 or Vps75. The K290ac is introduced via intramolecular acetylation and presents a self-activation mechanism, as non-acetylated Rtt109 is nearly inactive. In early work, it was hypothesized that K290ac acts as the catalytic residue, accepting the acetyl group from acetyl-CoA and transferring it to the H3 lysine residue; however, subsequent studies proved that binding of H3 and acetyl-CoA by Rtt109 does not follow any order and that K290ac is required for high affinity to acetyl-CoA (Albaugh *et al*, 2011). The search for the Rtt109 catalytic residue carrying out the acetyl-group transfer did not yield any conclusive answer: Rtt109 has neither a glutamate (typical for GNAT family) nor a tyrosine (p300) in its active center. D287 and D288 were considered as the potential catalytic bases; however their mutation did not increase the pH dependence of k_{cat} as expected if they were catalytic bases, but even slightly decreased it. Instead, D288 has been proposed to form a hydrogen bond with K290ac, which is thought to stabilize the position of the α5 helix and the architecture of the active center. The D288N mutation decreases the k_{cat}/K_M for acetyl-CoA ~10^3-fold; this effect is similar to that of the K290ac mutation, suggesting that both residues and possibly their mutual interaction aid acetyl-CoA binding (Albaugh *et al*, 2010). Overall, the data support the direct transfer of the acetyl group from acetyl-CoA to the H3 lysine, with the reaction intermediate being stabilised by the enzymatic pocket (Albaugh *et al*, 2011).
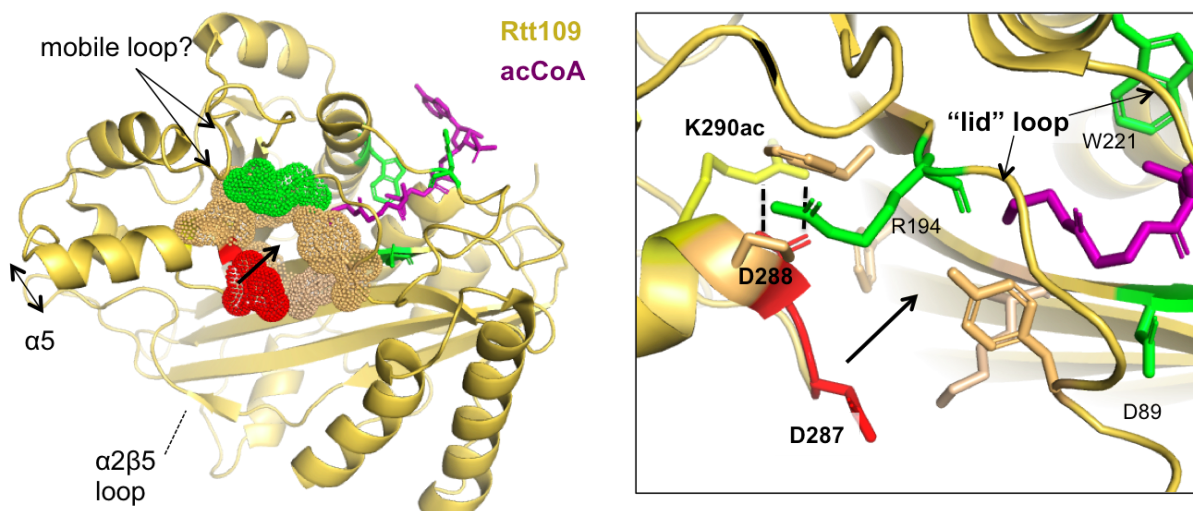
Figure 1.2.2. Rtt109 active center. The residues involved in the formation of the hydrophobic tunnel and the Rtt109 catalytic site are shown in sticks. The residues forming the walls of the tunnel are shown in the mesh representation. Conserved residues whose mutations interfere with Rtt109 acetylation activity are shown in green (D89, R194, W221 and H211); the negatively charged residues

16

are shown in red; K290ac is shown in yellow. D288 makes hydrogen bonds with K290ac, while D287 forms the base of the hydrophobic channel.

An intriguing feature of Rtt109 is the requirement of two different histone chaperones for its histone acetylation activity (Fig. 1.2.1). *In vitro*, activity of Rtt109 increases ~100-fold when Asf1 and/or Vps75 are present (Albaugh *et al*, 2010). These histone chaperones are thought to promote Rtt109-dependent H3 acetylation via different mechanisms. While Asf1 interacts with the enzyme transiently, Vps75 forms a tight complex with Rtt109. The chaperones also alter the specificity of the enzyme. Asf1 is known to promote H3 K56 acetylation, but Vps75 increases the acetylation of H3 K9, K23, K27 (discussed in detail in sections 1.3, 1.4). The mechanism by which the Asf1 and Vps75 chaperones regulate Rtt109 specificity is one of the focuses of this thesis work.
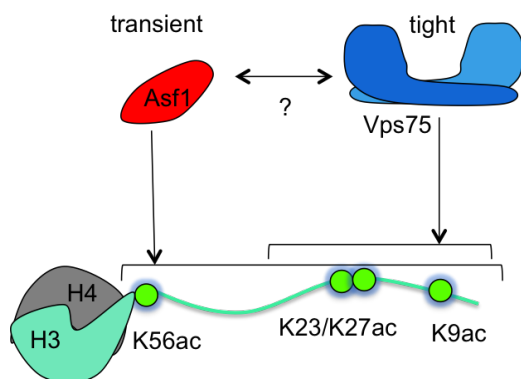
Figure 1.2.1. H3 acetylation by Rtt109 depends on histone chaperones. Asf1 is associated with all Rtt109-dependent acetylation marks of H3 and is essential for the H3 K56ac. Vps75, in turn, is not necessary for H3 K56ac *in vivo*, but aids Rtt109-dependent acetylation of the H3 N-terminal lysines. The notable difference between these acetylation sites is that H3 K56 is located in the αN helix in the nucleosome and near the folded core of H3 in the Asf1–H3:H4 complex. H3 K9, K23 and K27, on the other hand, are located in the long disordered N-terminal tail.

Rtt109-dependent acetylation marks of H3 are associated either with newly synthesized histones and/or with transcriptionally active chromatin. Apart from its function in replication-coupled nucleosome assembly (section 1.1), H3 K56ac promotes the so-called "nucleosome breathing", slightly weakening the contact between the DNA and the histone octamer in the nucleosome, and allowing the DNA to slide past the octamer. This function of H3 K56ac is required for the nucleosome dis- and re-assembly by the FACT complex during transcription (McCullough *et al*, 2019). Acetylation of H3 K56 is reversible and the mark is removed in yeast by Hst3 and Hst4 HDATs in the G_2/M-phase of the cell cycle (Maas *et al*, 2006). The cyclic nature of H3 K56 acetylation and deacetylation contributes to the fidelity of replication and protection from mutations.

The H3 N-terminal tail acetylation is in general linked to the activation of transcription and is incompatible with the methylation of the lysines, which in turn, is associated with silenced chromatin. In particular, H3 K9ac levels are elevated at the 5′ ends of actively transcribed genes and are thought to promote transcription by facilitating the interaction between TFIID (transcription factor of RNA polymerase II) and chromatin (Fillingham *et al*, 2008). Similarly to H3 K9ac, H3 K27 acetylation is abundant in the promoter regions, where it enhances the transcription of the genes (Chen *et al*, 2019a).

A number of H3 acetylation marks introduced by Rtt109 can be also introduced by Gcn5 in *S. cerevisiae*. The functions of the two HATs partially overlap, as both of them acetylate H3 K9, K23 and K27 *in vivo*. Unlike Gcn5, Rtt109 does not contribute to H3 K14 acetylation (Berndsen *et al*, 2008) and does not show any detectable activity towards nucleosomal H3 (Han *et al*, 2007). Gcn5, on the other hand, can acetylate the H3 N-terminal tail in intact nucleosomes and aids the propagation of PTMs in chromatin (Li & Shogren-Knaak, 2009). Nuclear localization is not required for Rtt109-dependent acetylation of H3, which suggests that Rtt109 acts on newly synthesized H3 in cytosol (Keck & Pemberton, 2011), further highlighting the role of Rtt109 in nucleosome biosynthesis.

Due to its unique capacity to acetylate H3 K56, Rtt109 was identified as a potential drug target for the treatment of opportunistic fungal infections caused by *Candida albicans*, *Pneumocystis jirovecii* and, potentially, other fungi (Rosa *et al*, 2010; Dahlin *et al*, 2014). Deletion of the Rtt109 homolog in *C. albicans* leads to hypersensitivity to genotoxic agents and hence to the susceptibility to the host immune macrophage-mediated response. Mice infected with a *RTT109Δ C.albicans* strain showed a close to 100% survival rate, due to the inability of the pathogen to proliferate efficiently (Rosa *et al*, 2010). This data together with the low sequence conservation between Rtt109 and its human homolog, p300, initiated a search for the Rtt109 inhibitors. Despite the structural similarity of the active centers of p300 and Rtt109, known p300 inhibitors had a very weak effect on Rtt109 activity (Tang *et al*, 2008a). Several small-molecule compounds were identified as potential Rtt109 inhibitors from high-throughput screening (Rosa *et al*, 2013; Dahlin *et al*, 2013). However, a large portion of them failed to proceed through clinical trials either because of their promiscuous activity (due to their ability to nonspecifically covalently bind to cysteine residues) or to their inefficiency *in vivo* (Dahlin *et al*, 2015). Recently, inhibitors for *P. jirovecii* Rtt109 were designed via virtual screening in combination with homology modelling and computational docking; however, the activity of these compounds *in vivo* has not been reported yet

(Sugumar *et al*, 2016). Further research in this direction could clearly benefit from the more extensive structural data and understanding of the mechanism of Rtt109 activity.

1.3. Vps75 and its role in Rtt109-dependent acetylation of H3.

The *VPS75* gene (vacuolar protein sorting) was first described during a genome-wide screen for genes involved in vacuolar protein sorting in *S. cerevisiae* (Bonangelino *et al*, 2002). Only later the protein product of this gene, Vps75, was identified as a histone chaperone belonging to the nucleosome assembly protein (NAP) family (Selth & Svejstrup, 2007; Bowman *et al*, 2014). NAP family chaperones are characterised by a conserved "headphone" fold and constitutive homodimerization (Bowman *et al*, 2014; Park & Luger, 2006). Members of this family are present in all eukaryotes, where they take part in chromatin-related processes, influencing nucleosome assembly, single-stranded DNA nicking during apoptosis, transcription and cell-cycle regulation (Park & Luger, 2006; Fan *et al*, 2003). Members of the NAP family of chaperones have been reported to bind a wide range of histone forms, ranging from the core histone dimer and octamer to the linker histone H1 (Shintomi *et al*, 2005). While Vps75 shares the characteristic fold of proteins in this family, it displays several atypical features connected to its role in promoting H3 acetylation by Rtt109 (Tang *et al*, 2008b; Selth & Svejstrup, 2007).

The Vps75 monomer consists of three regions: the dimerization helix, the earmuff domain (named so after the "headphone" homodimer shape) and an unstructured C-terminal acidic domain (CTAD) (Fig. 1.3.1). The most notable difference between Vps75 and other chaperones of this family is the conserved positively charged area around the α5 helix, which together with the neighbouring acidic patch creates an interaction site for both Rtt109 binding and for self-tetramerization (Tang *et al*, 2008b; Selth & Svejstrup, 2007; Bowman *et al*, 2014; Su *et al*, 2011).
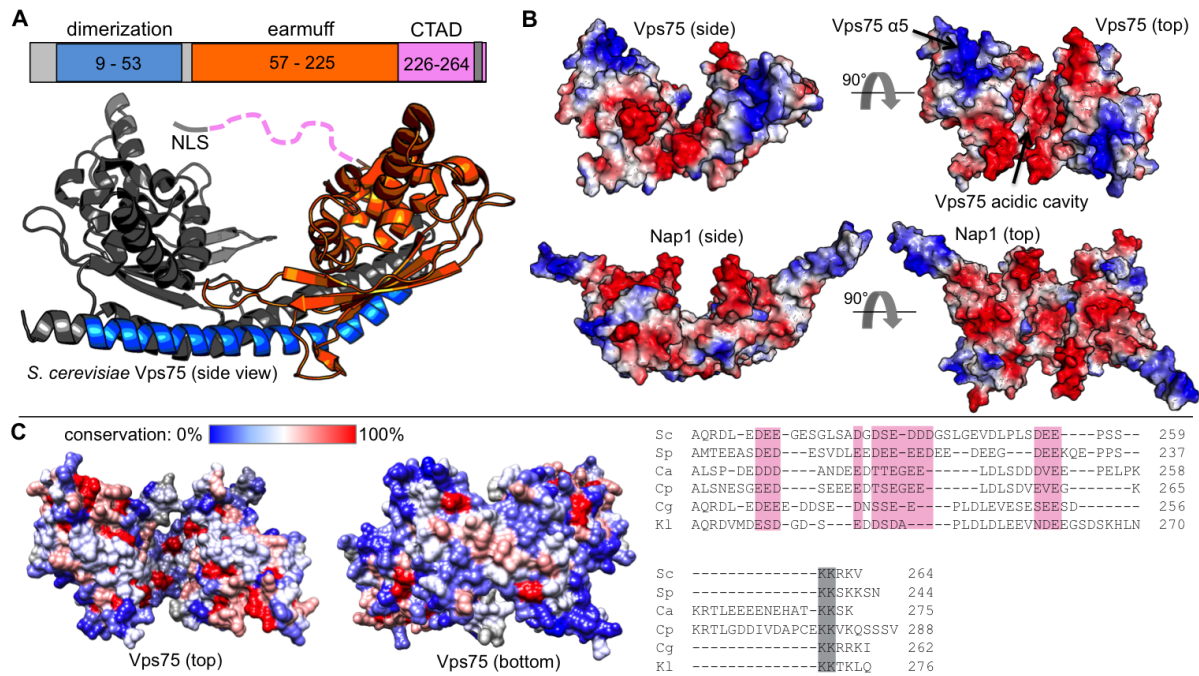
Figure 1.3.1. Vps75 structural features. A. Vps75 structural domains: dimerization helix (blue), earmuff domain (orange) and the unstructured C-terminal acidic domain (CTAD, pink) which includes nuclear localization signal (NLS, grey). Structural features are displayed on one Vps75 monomer, Vps75(A), the other one is shown in dark grey. **B.** Surface charge representation of Vps75 and its closest homolog in *S. cerevisiae*, Nap1, which is also able to self-tetramerize and bind Rtt109, but does not stimulate its activity (Park *et al*, 2008; Bowman *et al*, 2014). While both chaperones share the overall fold and possess an internal acidic cavity, the Vps75 earmuff domain lacks the long protruding β sheets and has a positively charged α5 helix instead. **C**. Conservation of Vps75 in fungal species in which histone acetylation is dependent on Rtt109 (*Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp), *Candida albicans* (Ca), *Candida parapsilosis* (Cp), *Candida glabrata* (Cg), *Kluyveromyces lactis* (Kl)). Right panel, manual sequence alignment of the unstructured CTAD sequences of Vps75 orthologs. Despite the low sequence conservation all CTADs share negatively charged patches (pink) and contain an NLS (grey). The panel is adapted from Danilenko *et al*, 2019, and was originally produced by myself.

Vps75 is well-known for its role in Rtt109-mediated acetylation, but was also demonstrated to have an Rtt109-independent role in transcription and nucleosome dynamics. Notably, *VPS75* expression levels are stable throughout the cell cycle, while *RTT109* expression rises only before the S phase (Selth *et al*, 2009). In accordance with this, a portion of Vps75 in the cell is not Rtt109 bound, but Rtt109 exists in complex with Vps75 *in vivo* and typically co-purifies in complex with this chaperone in pull-down experiments (Selth & Svejstrup, 2007; Keck & Pemberton, 2011; Krogan *et al*, 2006).

Although deletion of *VPS75* in *S. cerevisiae* is not lethal, the cells lacking *VPS75* display an increased sensitivity to DNA damage; specifically, the *VPS75* deletion strains (*VPS75Δ*) displayed a lower efficiency of DNA double strand break repair and the loss of Rtt109-dependent acetylation of the H3 N-terminal tail (Park *et al*, 2008; Jessulat *et al*, 2008;

Keck & Pemberton, 2011)[1]. *In vivo* data clearly shows the importance of Vps75 for the Rtt109-dependent H3 K9, K23 and K27 acetylation (Fig 1.3.2). On the contrary, the impact of *VPS75* deletion on H3 K56 acetylation is much lower and is incomparable with the deletion of *RTT109* or *ASF1* (Keck & Pemberton, 2011).

In vivo data	Deletion of *RTT109*	Deletion of *VPS75*
Rtt109-dependent acetylation of H3	Berndsen *et al*, 2008*; Seth *et al*, 2007; Han *et al*, 2007; Tsubota *et al*, 2007; Fillingham *et al*, 2008;	Keck & Pemberton, 2011; Tang *et al*, 2011; Berndsen *et al*, 2008*; Seth *et al*, 2007; Han *et al*, 2007; Tsubota *et al*, 2007; Fillingham *et al*, 2008;
K9	+* / n.d./ n.d. / n.d. / -	- / - / +* / n.d. / n.d. / n.d. / -
K23	+* / n.d. / n.d. / n.d. / n.d.	n.d. / n.d. / +* / n.d. / n.d. / n.d. / n.d. /
K27	n.d.	- / - / n.d. / n.d. / n.d. / n.d. / n.d.
K56	n.d. / - / - / - / -	++/ +++ / n.d. / ++ / +++ / ++ /+++

Figure 1.3.2. Vps75 is important for the Rtt109-dependent acetylation of the H3 N-terminal tail (K9, K23, K27) *in vivo*. Summary table of the previously published data, the sign order corresponds to the order of the citations: the acetylation varies from "+++" (identical to the wild-type strain) to "-" (acetylation is completely abolished/ reduced to below detection level); "n.d." stands for "no data"; the asterisk marks the results from experiments in which a *S. cerevisiae* strain containing wild type Gcn5 was used. Gcn5 is able to acetylate H3 K9 and K23, so the effect of *RTT109*/*VPS75* deletion is partially masked. Deletion of *VPS75* abolishes the H3 N-terminal tail acetylation similarly to the deletion of *RTT109*, but has little effect on the H3 K56 acetylation.

Vps75 has a nanomolar affinity for Rtt109 (K_d =10 ± 2 nM), with which it forms a tight and stable complex (Albaugh *et al*, 2010). In addition, Vps75 also has nanomolar affinity for the H3 and H4 histones, but the interaction between Vps75/Rtt109 and the histones was poorly described at the time this thesis work began, preventing investigation of Vps75-promoted acetylation (Berndsen *et al*, 2008). Thus, the data known at the time suggested three possible mechanisms by which Vps75 might stimulate Rtt109 activity:

1. Conformational change. It was speculated that Rtt109 in isolation adopts a catalytically inactive conformation, while association with Vps75 causes a conformational change and stabilization of an active state of the enzyme (Kolonko *et al*, 2010). However, neither of the structures of Rtt109–Vps75 complexes, discussed below, revealed significant changes in the Rtt109 active center, challenging this hypothesis (Fig. 1.3.3).

---

[1] In two other studies, deletion of *VPS75* was reported to have no influence on the resistance to the DNA-damaging agents: in Seth *et al*, 2007 and Tsubota *el al*, 2007 *VPS75Δ* was not sensitive to 100 mM hydroxyurea and camptothecin, respectively. However, the assays were optimised for comparison to the strain with *RTT109* deletion, which exhibits extreme sensitivity to the DNA damage. In the studies cited above, *VPS75Δ* was sensitive to such DNA damaging agents as 200 mM hydroxyurea, methyl methanesulfonate and UV.

2. Stabilization of Rtt109. The long unstructured Rtt109 α2β5 loop contributes to the enzyme instability *in vivo*. This loop is necessary for the formation of the complex with Vps75, where it becomes structured and protected from degradation (Tang *et al*, 2008a). Experiments with the Rtt109 lacking the α2β5 loop demonstrated that *in vivo* this mutant is functional, stable and its stability no longer depends on the presence of Vps75. Acetylation of H3 K56 was unaffected by the α2β5 loop deletion, but H3 K9 and K27 acetylation was significantly decreased. This suggested that Vps75 has more functions in H3 N-terminal tail acetylation than just the maintenance of Rtt109 levels in the cell (Keck & Pemberton, 2011).

3. Substrate presentation. While Rtt109 is able to bind the H3:H4 by itself, its affinity to the histones is ten times lower than the one for Vps75 and H3:H4 together; thus Vps75 might facilitate a tighter interaction between the enzyme and the histones (Kolonko *et al*, 2010). Binding to Vps75 could also lead to a correct positioning of H3:H4 in relation to Rtt109 and a better accessibility of the H3 N-terminal tail for the enzyme (Tang *et al*, 2011; Keck & Pemberton, 2011).

In total, three X-ray structures of Rtt109-Vps75 complexes were reported (Tang *et al*, 2011; Kolonko *et al*, 2010; Su *et al*, 2011). Unexpectedly, the most notable difference between the structures is the stoichiometry between the Rtt109 and Vps75. In the structure from Tang *et al*, each Vps75 monomer is bound to one copy of Rtt109, resulting in the formation of a 2:2 complex (Rtt109$_2$–Vps75$_2$, PDB entry 3q33, 3q35). In the structures from Su *et al* and Kolonko *et al*, one Rtt109 is bound to the Vps75 homodimer (1:2 complex, Rtt109–Vps75$_2$, PDB entry 3q66, 3q68 (Su *et al*); the low resolution of the data from Kolonko *et al* prevented the authors from refining the structure to a high level of detail) (Fig. 1.3.3). Notably, all structures show the same interaction between the Rtt109 α2β5 loop and the Vps75 dimerisation helix, as well as between the charged patches on the Vps75 earmuff domain and the complementary surface of Rtt109. These contacts were proven to be essential for the formation of the Rtt109-Vps75 complex (Tang *et al*, 2011). In addition to that, the 1:2 structure has a small interaction interface between Rtt109 α5 and Vps75(A) α5. In this case, mutations reversing charge distribution or deletion of Vps75 α5 weakened the interaction between the two proteins (Tang *et al*, 2011).
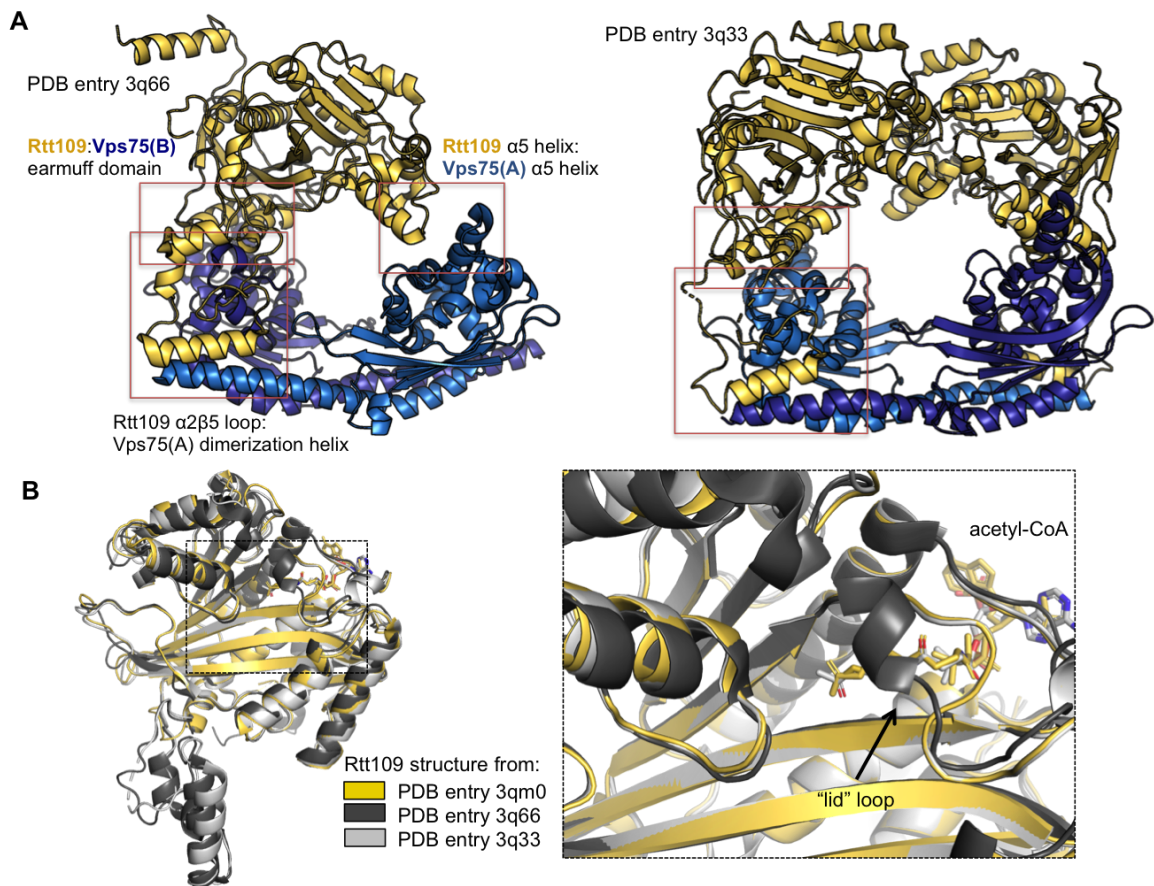
Figure 1.3.3. Interaction of Rtt109 with Vps75. A. The structures of 1:2 and 2:2 Rtt109–Vps75 complexes (left and right, respectively). The interaction interfaces between Rtt109 and Vps75 are highlighted by red boxes. The 1:2 complex has an additional small interaction interface between the α5 helix of Rtt109 and the α5 helix of Vps75. **B**. The active center of Rtt109 remains unchanged upon Vps75 binding. The site in the dashed box is shown in an expanded view on the right. The differences in the lid loop of Rtt109 from the PDB entry 3q66 are due to the absence of bound acetyl-CoA.

The possibility of the formation of both 1:2 and 2:2 complexes arises from the flexibility of the Vps75 homodimer. The dimerisation helix of Vps75 can adopt various conformations, thereby affecting the degree of the "opening" of the Vps75 homodimer, allowing to accommodate one or two copies of Rtt109 (Fig 1.3.4). The free Vps75 homodimer displays the same type of flexibility, and it was speculated that these conformational dynamics may be required for its function in H3 acetylation as well as for the self-tetramerization (Berndsen *et al*, 2008; Bowman *et al*, 2014).
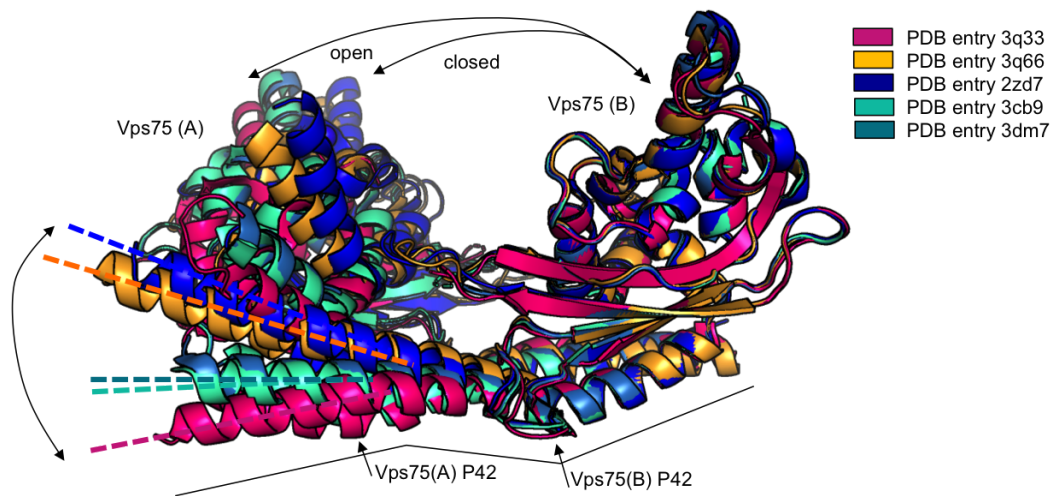
Fig. 1.3.4. Conformational flexibility of the Vps75 homodimer. The structures of Vps75 dimers in isolation (blue, mint, teal) or from 1:2 and 2:2 Rtt109–Vps75 complexes (orange and pink, respectively). The kink in the dimerization helix determines the degree of the opening of the homodimer, which enables binding of two copies of Rtt109 and may play a role in the accommodation of the histones. The Vps75 dimers are aligned on the Vps75(B) chain.

The different stoichiometry does not stem from *in vitro* reconstitution protocols, as all crystals were produced from co-expressed in *E. coli* and co-purified Rtt109 and Vps75. The placement of the affinity tag does not explain the enrichment in Rtt109 or Vps75 either, as the Rtt109$_2$–Vps75$_2$ structure was obtained with the GST-tagged Vps75 and untagged Rtt109. The major differences between the protocols lie in the co-expression strategies and in the Vps75 constructs. Tang *et al* used a truncated construct of Vps75 (lacking the CTAD) and co-expressed both proteins from the same vector, which possibly favoured the 2:2 stoichiometry (Tang *et al*, 2011). In the two other cases, the full-length Vps75 and Rtt109 were co-expressed from separate vectors (Kolonko *et al*, 2010; Su *et al*, 2011).

Interestingly, biochemical data did not discriminate clearly between the two stoichiometries. Analytical ultracentrifugation experiments on the 2:2 mixture of Rtt109 and Vps75 or Vps75 lacking the CTAD showed the formation of complexes with identical sedimentation coefficients, together with the presence of unbound material. This suggested that the stoichiometry of the Rtt109–Vps75 complex does not depend on the Vps75 CTAD; the presence of unbound material is likely due to the formation of a 1:2 complex with the excess of Rtt109 appearing in the free form (Park *et al*, 2008). On the other hand, native gel shift assays demonstrated a shift at a Rtt109:Vps75 ratio of 1:2 and a super shift corresponding to the 2:2 complex: this supershift was also independent of the presence of the Vps75 CTAD (Park *et al*, 2008). In any case, Rtt109 activity was found to be higher in the 1:2 Rtt109–Vps75 complex (Kolonko *et al*, 2010).

The mutations of Vps75 which disrupted or weakened its interaction with Rtt109 also reduced the H3 N-terminal tail acetylation *in vivo* similarly to the *VPS75* deletion, confirming that the acetylation of these sites is performed by the Rtt109–Vps75 complex in the cell (Tang *et al*, 2011; Kolonko *et al*, 2010; Su *et al*, 2011). Apart from the Vps75 structural elements directly involved in the interaction with Rtt109, mutagenesis analysis identified two other Vps75 sites important for its function: Vps75-E206 and E207, located on the inner surface of the Vps75 dimer, and the Vps75 CTAD. As Vps75 E206/E207 mutations decreased H3 K9 acetylation *in vivo*, but did not affect binding to the histones or the affinity to Rtt109, it was speculated that they contribute to the dynamics of the Vps75 dimer that may be required for the H3 acetylation (Tang *et al*, 2011; Berndsen *et al*, 2008).

The presence of the CTAD is conserved among Vps75 homologs, but the sequence conservation of this acidic region is low (Fig. 1.3.1). While Vps75 lacking the CTAD was not able to rescue the *VPS75Δ* phenotype, its effect on the H3 acetylation *in vivo* was not characterised (Park *et al*, 2008). The last eight residues of the Vps75 CTAD contain a nuclear localization signal (NLS), which promotes Rtt109 import into the nucleus. However, Vps75 NLS is not crucial for Rtt109 stability or Rtt109-dependent acetylation of H3 K9, which is in agreement with the model where Rtt109 acetylates the newly synthesized H3 in cytoplasm (Keck & Pemberton, 2011).

In vitro acetylation data on Rtt109 activity partially contradict the *in vivo* observations. For instance, Vps75 strongly stimulated Rtt109 activity towards H3 K56 *in vitro* even in the absence of Asf1, and in some cases, H3 K56 was reported to be the preferred acetylation site for the Rtt109-Vps75 complex (Fillingham *et al*, 2008; Tsubota *et al*, 2007; Radovani *et al*, 2013; Berndsen *et al*, 2008). However, the results of these *in vitro* assays were affected by the use of the histones from chicken erythrocytes, in which H3 is pre-acetylated at K14 and K23. These pre-existing modifications were found to influence the specificity of Rtt109 in favour of H3 K56 acetylation (Cote *et al*, 2019; Kuo *et al*, 2015). Another complication with the interpretation of the *in vitro* activity assays is the differences in stoichiometry between the components: Vps75 was added to Rtt109 either in 1:2 or 2:2 ratio, while the substrate histones were presented either as H3:H4 dimer or tetramer or as an isolated H3. Nevertheless, the *in vitro* activity assays mostly confirmed the effect of the mutations observed *in vivo*. In particular, deletion of the Vps75 CTAD and the double mutation of Vps75 E206K, E207K reduced the Rtt109 activity rate ~4- and 2-fold, respectively (Tang et al, 2011). Deletion of the CTAD did not affect the Vps75 affinity to Rtt109 or to the histones;

however, its role in acetylation was prominent and could not be replaced by the addition of a polyglutamate stretch to the reaction mixture (Park et al, 2008; Tang et al, 2011).

Overall, *in vivo* data suggest that Vps75 is not necessary for the H3 K56 acetylation, but instead promotes the acetylation of the N-terminal tail of H3 by Rtt109, acting as a specificity switch of the enzyme. The likely substrate for Rtt109 in complex with Vps75 *in vivo* is the Asf1-bound histones (discussed further in the chapter 1.4; Bowman *et al*, 2016).

1.4. Asf1 and its role in Rtt109-dependent acetylation of H3.

Asf1 (anti-silencing function protein 1) belongs to the Asf1-like histone chaperone family, which includes eukaryotic Asf1 homologs (Daganzo *et al*, 2003). *S.cerevisiae ASF1* was identified in the screening for genes which take part in the anti-silencing of the transcriptionally repressed mating type loci (Le *et al*, 1997). All Asf1-like proteins possess a conserved N-terminal domain (residues 1-155 of the *S.cerevisiae* Asf1), while the presence, the length and the charge of the disordered C-terminal acidic region varies among different species (Fig 1.4.1.). The chaperones of this family are known to interact with the heterodimer H3:H4 and participate in histone processing and modification as well as nucleosome assembly and disassembly (Hamiche & Shuaib, 2013).
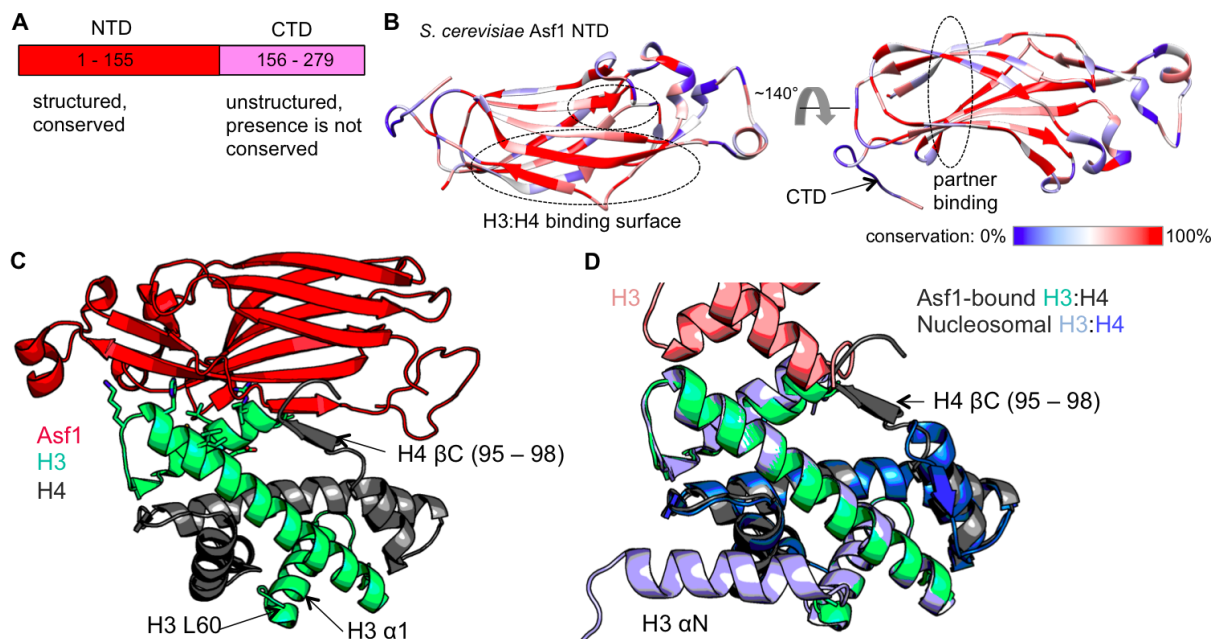
Figure 1.4.1. Asf1 binds H3:H4 via conserved surface on its N-terminal domain and prevents H3:H4 tetramerization. A. Scheme of structural organisation of *S. cerevisiae* Asf1. **B.** Conservation of Asf1 in eukaryotes (the sequence alignment was done for Asf1 homologs from *S. cerevisiae, Candida albicans, Schizosaccharomyces pombe, Gibberella zeae, Debaryomyces hansenii, Cryptococcus neoformans, Puccinia striiformis, Drosophila melanogaster, Xenopus laevis, Xenopus tropicalis, Mus musculus, Homo sapiens*). The H3:H4 interaction interface of Asf1 is highly conserved.

Several interaction partners of Asf1 (Cac2 (CAF-1 subunit, *S. pombe*), Hip1 (HIRA subunit, *S.pombe*), HIRA (*H. sapiens*), the C-terminal peptide of Rad53 (*S. cerevisiae*)) bind to the opposite side of the Asf1 NTD. **C.** Asf1 (red) engages the H3:H4 dimer (green and grey, respectively; PDB entry 2hue). The H3 residues involved in direct interaction with Asf1 are shown in sticks. The H3 N-terminal tail is unstructured; the H4 C-terminal tail folds into a short β-sheet (bC) when in complex with Asf1. **D.** Alignment of Asf1-bound H3:H4 from (**C**) to the H3:H4 in the context of nucleosome (the H3:H4 dimer aligned to the Asf1-bound histones is shown in light blue and blue; the second copy of H3 in the H3:H4 tetramer is shown in pink; PDB entry 1eqz). In the nucleosome the same surface of H3 that binds Asf1 is required for the formation of the H3:H4 tetramer. Also, the N-terminal tail of H3 forms an additional helix (aN), which comprises H3 K56.

The N-terminal domain of Asf1 interacts with the H3:H4 dimer, engaging the same surface of H3 that is responsible for the H3:H4 tetramerization (Fig. 1.4.1). Consequently, it has been shown that Asf1 can disrupt $(H3:H4)_2$ *in vitro*, but is unable to bind cross-linked tetramers (Bowman *et al*, 2011). The C-terminal domain (CTD) of *S. cerevisiae* Asf1 is dispensable for the histone binding as well as for the anti-silencing function of Asf1 and H3 K56 acetylation (Minard *et al*, 2011; Recht *et al*, 2006). The CTD was demonstrated to contribute to nonspecific association of Asf1 with chromatin and to increase the affinity of Asf1 for the histones (Minard *et al*, 2011; Dennehey *et al*, 2013). Residue 210 of the Asf1 CTD substituted to an unnatural amino acid cross-links to H3 *in vivo*; Asf1 E210 is surrounded by poly-E and poly-D stretches, suggesting possible interactions between the Asf1 CTD and H3:H4 positively charged DNA-binding surface, which is not engaged by the Asf1 NTD (Dennehey *et al*, 2013).

Apart from the H3:H4 dimer, Asf1 interacts with CAF-1 and HIRA complexes, involved in replication-dependent and -independent nucleosome assembly (Sharp *et al*, 2001; Green *et al*, 2005; Mousson *et al*, 2007); with Rad53 to function in the DNA-damage response (Jiao *et al*, 2012); with the SAS complex for chromatin remodelling and silencing (Reiter *et al*, 2015); with the replication factor C (Franco *et al*, 2005), with a subunit of the MCM helicase to facilitate eviction, splitting and redeposition of $(H3:H4)_2$ (Huang *et al*, 2015); and other protein partners. Some of these interactions are compatible with the Asf1–H3:H4 complex, as they involve the surface of the conserved N-terminal domain of Asf1 which is not occupied by the H3:H4 dimer (Fig. 1.4.1B). This binding mode facilitates Asf1's functions in the histone handover and its deep integration in multiple histone-related processes.

It is well established that the Rtt109-dependent H3 K56 acetylation necessitates Asf1 (Fig. 1.4.2). Alike the deletion of *RTT109*, *ASF1* deletion is not lethal but *ASF1Δ* cells display high sensitivity to DNA damaging stress and proliferate slower than the WT cells (Mousson *et al*,

2007). The *ASF1Δ* phenotype is similar to that of the H3 K56R mutation (Recht *et al*, 2006); while the H3 K56Q mutation, which mimics acetylated H3 K56, is able to partially rescue the *ASF1Δ* phenotype (Masumoto *et al*, 2005). Both *ASF1* and *RTT109* are expressed most actively during the S-phase and deletion of either one completely abolishes H3 K56 acetylation (Le *et al*, 1997). In addition to H3 K56, Asf1 is required for Rtt109-dependent H3 K9 acetylation. However, the role of Asf1 for the H3 N-terminal tail acetylation by Rtt109 *in vivo* is less well described.

In vivo	Deletion of *RTT109*	Deletion of *ASF1*
Rtt109-dependent acetylation of	Berndsen *et al*, 2008*; Seth *et al*, 2007; Han *et al*, 2007; Tsubota *et al*, 2007; Fillingham *et al*, 2008; Kuo *et al*, 2015*	Recht *et al*, 2006; Seth *et al*, 2007; Tsubota *et al*, 2007; Adkins *et al*, 2007; Fillingham *et al*, 2008; Kuo *et al*, 2015*
K9	+* / n.d./ n.d. / n.d. / - / +*	n.d. / n.d. / n.d. / +* / - / +*
K56	n.d. / - / - / - / - / n.d.	- / - / - / - / - / n.d.

Figure 1.4.2. Asf1 is important for Rtt109-dependent acetylation of H3 *in vivo*. Summary table of the previously published data; the sequence of the symbols in each column corresponds to the order of the citations: the acetylation varies from "+++" (identical to the wild-type strain) to "-" (acetylation is completely abolished/ reduced to below detection level); "n.d." stands for "no data"; the asterisk marks the results from experiments that used a *S. cerevisiae* strain containing wild type Gcn5. Gcn5 is able to acetylate H3 K9, so the effect of *RTT109/ASF1* deletion is partially masked. Deletion of *ASF1* abolishes the Rtt109-dependent acetylation of H3 K9 and K56.

Deletion of either *ASF1* or *VPS75* abolishes Rtt109-dependent H3 K9 acetylation (Fillingham *et al*, 2008), suggesting that both of these chaperones are required for the H3 N-terminal tail acetylation by Rtt109 (Fig. 1.4.2, Fig. 1.3.2) As Asf1 co-immunoprecipitated with Vps75-bound Rtt109 and this interaction was dependent on the ability of Asf1 to bind H3:H4, it was speculated that Asf1 can act to present H3:H4 as a substrate to the Rtt109−Vps75 *in vivo* (Han *et al*, 2007).

In vitro, Asf1 is able to stimulate Rtt109 activity only towards the H3:H4 dimers, but not H3 in isolation, suggesting that Asf1 promotes H3 acetylation via a substrate presentation mechanism. H3 K56, K9 and K23 acetylation is enhanced by the addition of Asf1 to the Rtt109–Vps75 complex (Han *et al*, 2007; Kuo *et al*, 2015). Mass-spectrometry also suggested that Asf1 does not alter the range of the residues acetylated by Rtt109–Vps75, but increases its selectivity towards H3 K9 and K23 (Kuo *et al*, 2015).

Co-purification of Asf1 with Rtt109 could only be achieved by cross-linking (Han *et al*, 2007), suggesting that unlike Vps75, Asf1 interacts with Rtt109 either transiently or solely through the H3:H4. In another study it was demonstrated that truncation of the last 11 C-terminal residues of Rtt109 (Rtt109ΔC) reduces acetylation of H3 K9 and K56 *in vivo* (Radovani *et al*,

2013). Notably, the effect of the Rtt109ΔC mutation on H3 K56ac is partially compensated by Vps75, suggesting that (i) the Rtt109ΔC mutation affects the Asf1-dependent acetylation, (ii) acetylation of H3 K56 *in vivo* is done by Rtt109 in complex with Vps75.

In parallel with the work in this thesis, an X-ray structure of *A. fumigatus* (*Af*) Rtt109 in complex with *Af* Asf1−H3:H4 was reported (Fig. 1.4.3) (Zhang *et al*, 2018). While *Af* Rtt109 does not bind to the *Af* homolog of Vps75, the sequence identity between the *S. cerevisiae* and *Af* Rtt109 is ~20% and the two proteins adopt a very similar fold in the active center.
In the *Af* structure, Asf1 promotes H3 K56 acetylation by substrate presentation: Asf1-bound H3:H4 interacts with Rtt109 in a way that allows positioning H3 K56 near the Rtt109 active center. H3 K56 remains in the unfolded state, similarly to the Asf1−H3:H4 complex. Asf1 does not contact Rtt109 directly, but mediates the interaction of the H4 βC with Rtt109.

Mutation of *S. cerevisiae* proteins	Corresponding interaction in the *Af* Rtt109–Asf1–H3:H4	Effect on H3 K56ac in *S. cerevisiae*
Rtt109 R292E, R318E	*Af* Rtt109 R265E, R306E; interact with H3 E94.	-
H3 E94R	Interacts with *Af* Rtt109 R265E, R306E.	+
H4 R95D	Part of H4 β-sheet, interacts with Rtt109.	+/++
Asf1 V146P, T147P	Part of *Af* Asf1 β-sheet interacting with the H4 β-sheet.	-
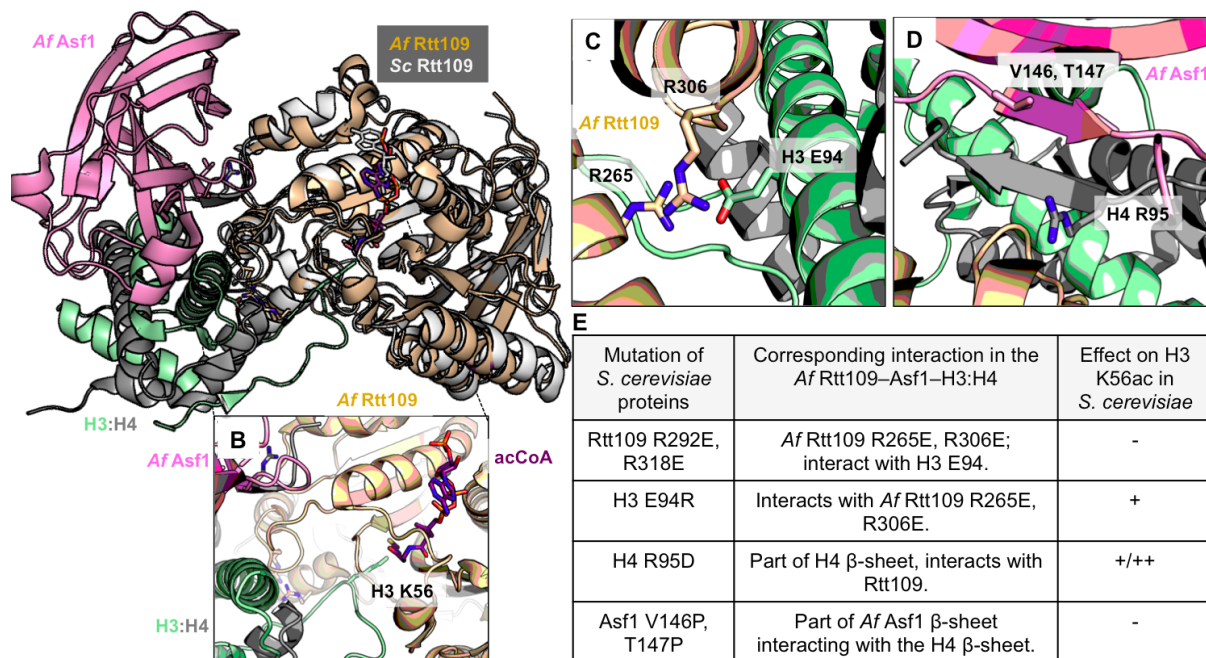
Figure 1.4.3. Structure of *Af* Rtt109 in *Af* Rtt109–Asf1–H3:H4 complex and comparison to *S. cerevisiae* (*Sc*) Rtt109. A. Structure of *Sc* Rtt109 (PDB entry 3qm0; white) is aligned to the *Af* Rtt109–Asf1–H3:H4 (PDB entry 5zba; beige, Rtt109; pink, Asf1; green, H3; grey, H4). **B**. Zoom on the region of the Rtt109 active center and H3 K56 in the *Af* Rtt109–Asf1–H3:H4 complex. **C**. Interaction of *Af* Rtt109 with H3 α2 helix. **D**. Interaction between *Af* Asf1 β11 with H4 C-terminal β-sheet, which makes contacts with *Af* Rtt109. **E**. Western blotting results from Zhang *et al* report that mutations of the residues in *Sc* proteins corresponding to the *Af* protein residues shown in (**C**, **D**) reduce H3 K56 acetylation in *S.cerevisiae in vivo* (the acetylation degree is denoted as in Fig. 1.4.2).

The structure of the *Af* Rtt109 complex was used by the authors to design mutants for *in vivo* acetylation studies in *S. cerevisiae* (Fig. 1.4.3). The mutations of conserved residues that participate in the *Af* Rtt109−Asf1−H3:H4 complex significantly reduce H3 K56 acetylation in

S. cerevisiae, suggesting that the Rtt109−Asf1−H3:H4 interaction interfaces are conserved between *Af* and budding yeast.

Although this data provides valuable insight into the interaction of *S. cerevisiae* Rtt109 with Asf1−H3:H4, it remains unclear how the acetylation of the H3 N-terminal tail (H3 K9, K23 and K27) is promoted by the chaperones, and which role Vps75 plays. Given the fact that *Af* Vps75 is unable to interact with *Af* Rtt109, it remains unclear how similar the *Af* and *S. cerevisiae* acetylation mechanisms are and whether Asf1−H3:H4 is also a substrate of Vps75-bound Rtt109.

1.5. Methodological background.

Methyl-TROSY NMR for high weight biomolecules.

NMR allows to study biomolecules in solution in a wide range of conditions, including intact cells. NMR experiments provide high-resolution information on the structure and dynamics of molecules. However, some of the drawbacks of the method include its low sensitivity and an upper limit in the size of molecules that can be studied.

The signal-to-noise ratio in a mono-dimensional NMR depends on several parameters:

$$S/N \sim N \; \gamma_{exc} \; \gamma_{det}^{(3/2)} \; B_0^{(3/2)} \; T_2^{(1/2)} \; NS^{(1/2)} \qquad (1)$$

where *N* is the number of atoms (for a fixed volume, this is equivalent to the sample concentration), γ_{exc} and γ_{det} are the gyromagnetic ratios of excited and detected spins, respectively, B_0 is the strength of the external magnetic field, T_2 is the transverse relaxation time-constant of the detected nucleus, and NS is the number of scans in the NMR experiment.

The T_2 relaxation time-constant becomes shorter as the molecular weight increases, and therefore the NMR signal typically decays rapidly for high-molecular-weight systems. Over the past two decades, the size limit of the systems that can be studied with NMR has increased from ~35 kDa to ~1 MDa due to the increase in the available magnetic field strengths (B_0) and, most importantly, the development of novel labelling schemes and pulse sequences.

TROSY (transverse-relaxation-optimized spectroscopy) NMR experiments are specifically designed to detect the slowest-relaxing transitions in a particular spin-system, therefore allowing the observation of NMR signals from high-molecular-weight systems. In an NMR experiment, the nuclear spins are initially aligned with the static magnetic field B_0, resulting in

a non-zero bulk magnetisation in the direction of the +z-axis. A short radio-frequency pulse transfers the bulk magnetisation into the transverse plane (x-y plane); the magnetization then precesses around the z-axis, generating a detectable NMR signal. In addition to the precession, the magnetisation begins to return towards its equilibrium state through a process called relaxation. There are two distinct relaxation processes: transverse relaxation, which is characterized by the T_2 time-constant and is the decay of the transverse component of the transverse, observable magnetisation to zero, and longitudinal relaxation, which is characterized by the T_1 time-constant and is the recovery of the bulk magnetisation along the +z-axis. In biomolecular NMR, T_2 is always shorter than T_1, and thus shortening of T_2 with increasing molecular weight determines the limits of sensitivity and resolution of NMR spectra for high-molecular-weight biomolecules.

The relaxation is caused by time-varying fluctuations of the local magnetic field at each nuclear spin; the principal relaxation sources are the fluctuating fields generated by dipolar (DD) interactions and chemical shift anisotropy (CSA) tensors. DD relaxation is caused by the rapid changes in the orientation of nucleus-nucleus dipoles with respect to the magnetic field due to molecular tumbling in solution. Similarly, CSA relaxation arises from the reorientation of the chemical shift anisotropy tensor with respect to the magnetic field due to molecular tumbling in solution; the change of orientation of the CSA tensor gives rise to a time-dependent change in the resonance frequencies, which causes loss of coherence. The interference between different relaxation mechanisms is called cross-correlated relaxation and can give rise to variations in the net relaxation rates for different transitions between spin-states in multi-spin systems.

A TROSY experiment is designed to selectively detect only the slowest-relaxing component of the magnetisation transfer pathway occurring during an heteronuclear 2D correlation experiment(Pervushin *et al*, 1997): here the relaxation interference of 1H-15N dipolar relaxation and 15N CSA relaxation subtracts or adds from the transversal relaxation of different spin-spin transitions, respectively lengthening and shortening the T_2. If the slow relaxing component is not mixed with the fast relaxing component during the pulse sequence and is selected for detection during the final acquisition time, one can obtain a considerable increase in sensitivity and resolution of the 2D correlation of high-molecular weight molecules, despite the fact that only half of the magnetization (the slow relaxing component) is observed while the other half (the fast-relaxing component) is suppressed.

While the first TROSY experiment was developed for [1]H-^{15}N correlation spectra of backbone amide groups, the TROSY effect was subsequently exploited for [1]H-^{13}C methyl groups

(Ollerenshaw *et al*, 2003). Irrespective of the additional gains provided by TROSY, the advantages of methyl groups detection are the presence of three pseudo-equivalent protons, which increases the magnitude of the corresponding bulk magnetization vector, and the fast rotation of the methyl group around its three-fold symmetry axis, which partially decouples it from the slow tumbling rates of large molecules and results in a slower decay of transverse magnetization. Further reduction in the transverse relaxation rates is provided by methyl-TROSY, which exploits the relaxation interference between the ^{1}H–^{13}C and ^{1}H–^{1}H DD relaxation mechanisms within the methyl group.

The methyl-TROSY NMR experiment requires perdeuterated samples containing selectively protonated methyl-groups. This is necessary to exploit the methyl-group-specific TROSY effect the most, without leakage of magnetization to neighbouring spins. Commercially available precursors allow the selective labelling of methyl groups of I, L, V, M, A & T amino-acids; additionally it is also possible to make S-methylthiocysteine, which can be detected also within a non-deuterated protein. The labelled proteins are typically expressed in *E. coli* as recombinant proteins; it is also possible to achieve a sufficient level of deuteration and selective methyl-group labelling via expression in *Pichia pastoris* and insect cells, as well as in cell-free expression systems (Schütz & Sprangers, 2019).

The spectra of the large systems can be simplified by the use of a ProS or ProR LV precursors, which result in $^{1}H,^{13}C$ labelling of only one pro-chiral methyl group of the LV residues. This labelling scheme halves the number of peaks in the LV region and hence reduces the spectral overlap significantly, without loss of information on individual LV residues. In addition, ProS/ProR labelling can be extremely useful for the verification and transfer of assignments. The spectra of large biomolecules and complexes can be further simplified by the selective labelling of individual domains of one molecule of the complex. This approach is known as segmental labelling and LEGO NMR (Mund *et al*, 2013).

Due to the large size of the proteins studied in this thesis work (up to ~60 kDa for individual subunits; ~160 kDa protein complex), the majority of the NMR experiments presented here are methyl-TROSY experiments, that is ^{1}H-^{13}C HMQC (heteronuclear multiple-quantum correlation) spectra, recorded on perdeuterated ^{1}H-^{13}C -ILV-methyl labelled samples.

Chemical shift perturbations (CSPs).

Comparison of the positions of the chemical shifts between the spectra of the same biomolecule/protein in different experimental conditions or in the presence of potential

binding partners provides information on changes in the chemical environment of specific sites. These changes are caused either by a conformational change of the protein or by a direct interaction. Therefore, monitoring chemical shift perturbations (CSPs) is a way to gain insights into the binding affinity of protein-protein and protein-ligand complexes. Moreover, due to the nature of the chemical shift evolution in an NMR experiment, CSPs can inform on the exchange regime between two protein states (Fig. 1.5.1).

In an experiment with a ligand (L) and a protein (P), where ligand reversibly binds to the single site on the protein, the observed states in the reaction can be described as:

$$L + P \leftrightarrow LP \quad (2)$$
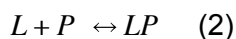
where LP is a ligand-protein complex. The protein in this case is observed in the free and the ligand-bound states, which interconvert as in:

$$d[P]/dt = -K_1[P] + K_{-1}[LP] \quad (3.1)$$

$$d[LP]/dt = K_1[P] - K_{-1}[LP] \quad (3.2)$$

where [P] and [LP] are the concentrations of the protein in the free form and in the ligand-bound state, respectively; K_1 and K_{-1} are the forward (P to LP) and back (LP to P) rate constants. In the equilibrium the concentration change of L and LP over time equals zero, hence from the equations (3.1) and (3.2), the ratio between the equilibrium concentrations equals:

$$[LP]/[P] = K_1/K_{-1} \quad (4)$$

The exchange rate K_{ex} describes the return of the system to the equilibrium state:

$$K_{ex} = K_1 + K_{-1} \quad (5)$$

The reaction in (2) can also be described in terms of the dissociation constant k_d, which describes the dissociation rate of the LP complex. Dissociation constant is the ratio of the equilibrium concentrations of the free ligand and protein to the concentration of their complex:

$$K_d = [L][P]/[LP] = K_{off}/K_{on} \quad (6)$$

where k_{on} and k_{off} are the rate constants for the forward (L + P to LP) and back (LP to L + P) reactions of the reaction shown in (2). From the equations (4), (5) and (6), the exchange rate K_{ex} can also be defined as:

$$K_{ex} = K_{on}[L] + K_{off} \quad (7)$$
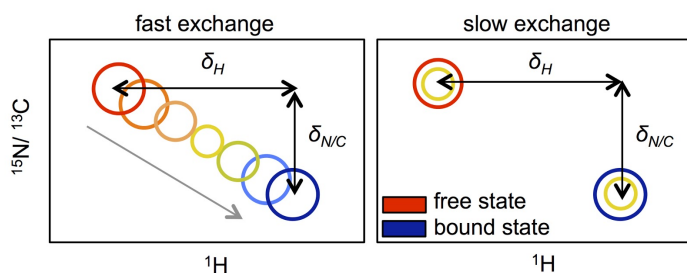
Fig. 1.5.1. CSPs. The figure shows the behaviour of protein peaks over several 2D NMR spectra acquired in the presence on increasing concentrations of a binding partner. The term "fast exchange" is used when the exchange rate K_{ex} is faster than the difference in the chemical shifts of the bound and free states in Hz ($\Delta\Omega$). Typically, fast exchange is observed for weak interactions with k_d in the μM range or above; the observed chemical shift is a population weighted average of the chemical shifts of the two states, and therefore chemical shifts change gradually as the concentration of the interaction partner increases. When $K_{ex} <<(\Delta\Omega)$, which is typically the case for interactions with sub-μM affinity, the chemical shifts of both states are observed as separate signals; the intensity of the original peak decreases gradually as the new peak appears and gains its intensity, until the protein is saturated with ligand and the signal corresponding to the free state disappears.

As CSPs are a function of changes in the chemical environment around individual nuclei, they are not a reliable source of information on the interaction interfaces, since they may arise both from direct interactions and conformational changes upon binding, which can occur also far from the binding site. However, measurement of CSPs and their visualization on the structure are useful to detect binding and obtain a candidate binding surface (to be confirmed with other methods). CSPs for an NMR experiment where N nuclei are recorded are calculated as:

$$CSP = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\alpha_i\delta_i)^2} \qquad (8)$$

where δ is a chemical shift difference and α is a scaling factor to adjust the δ of the different nuclei (Williamson, 2013).

During the titration of a ligand to a protein, their total concentrations ($[L]_t$ and $[P]_t$, respectively) can be presented as a sum of the ligand (or protein) concentration in a free state and in the ligand-protein complex (bound state). In fast exchange, observed CSP is the population weighted average of the chemical shifts of the free and the bound states and can be expressed as:

$$CSP_{obs} = CSP_{max}(([L]_t + [P]_t + K_d) - (([L]_t + [P]_t + K_d)^2 - 4[L]_t[P]_t)^{1/2})/2[P]_t \qquad (9)$$

where CSP_{max} is the maximal CSP for the ligand-saturated state; it can be either measured directly or derived from a fitting procedure. Hence, k_d can be extracted if the total concentrations of the ligand and the protein as well as the CSP_{max} are known.

Paramagnetic relaxation enhancement (PRE).

PRE experiments provide distance restraints in a range of ~10–35 Å (Battiste & Wagner, 2000). The method relies on the distance dependence of the effect that a paramagnetic group has on the nuclei observed in the NMR. Initially the method was applied to proteins that are able to bind paramagnetic metals; nowadays a selection of different paramagnetic tags is available that can be coupled to a position of interest specifically engineered in the biomolecule. In this study, I use 3-(2-iodoacetoamido)-PROXYL as paramagnetic tag (also called spin-label), which can react via its iodine group with the thiol group of a solvent-accessible cysteine (Fig. 1.5.2). The protein to be spin-labelled must contain only one solvent accessible cysteine to ensure that the paramagnetic tag is located at one specific position: this can be achieved by mutagenesis, whereby all undesired solvent accessible cysteine residues are substituted by serine and one cysteine is engineered at the desired position.

The design of the PRE experiment used in this work is shown in Fig. 1.5.2.
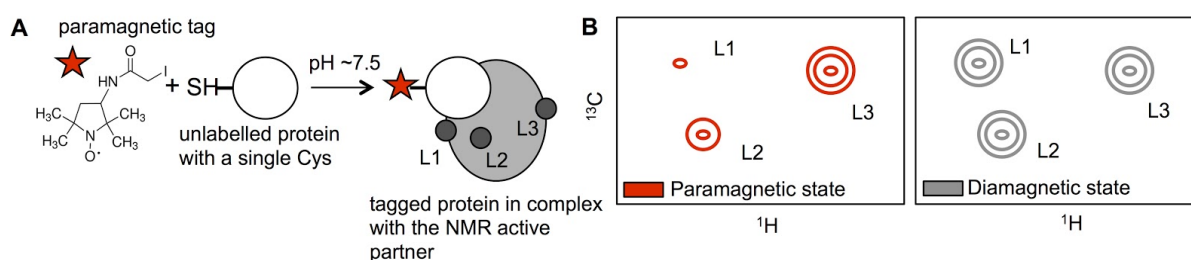
Figure 1.5.2. Scheme of a PRE experiment. A. A paramagnetic tag is covalently attached to one component of the complex at a precise position on one of the molecules of the complex. The tagged (or spin-labelled) protein is then added to the NMR active partner and spectra are recorded with the tag in paramagnetic and diamagnetic states. **B.** In the paramagnetic state, the tag increases the transverse relaxation rate of the nuclei closer than approximately 35 Å. The corresponding NMR peaks broaden and decrease in intensity in a distance-dependent manner. This effect is quantified by recording a reference spectrum where the tag is in the diamagnetic state and thus no longer changes the relaxation rates of the nuclei close to it. The diamagnetic state of the tag is obtained upon reduction with ascorbic acid.

The paramagnetic tag increases the transverse relaxation rate (R_2) of all nuclei closer than 35 Å in a distance-dependent manner. The mechanisms through which the paramagnetic center causes the increase in the relaxation rate of the nuclei are diverse and depend on the nature of the paramagnetic center. For nitroxide paramagnetic groups, the main mechanism is the electron-nucleus dipolar interaction. The difference between the transverse relaxation rates in the diamagnetic and paramagnetic states is the PRE rate (Γ_2), which is defined by the Solomon-Bloembergen equation (Solomon & Bloembergen, 1956) as:

$$\Gamma_2 = \tfrac{1}{15}\gamma^2 g^2 \mu_B^2 S(S+1)\{4J_{SB}(0) + 3J_{SB}(\omega_I)\} \qquad (10)$$

where γ is the gyromagnetic ratios of the nuclear spins, g is the electron g-factor, μ_B is the Bohr magneton, S is the spin quantum number of the electron, and J_{SB} is the generalised spectral density function. J_{SB} incorporates the distance dependence of the PRE rate:

$$J_{SB}(\omega_I) = r^{-6}\frac{\tau_c}{1+\omega^2\tau_c^2} \qquad (11)$$

where r is the distance between the free electron of the paramagnetic tag and the nucleus observed in the NMR, and τ_c is the electron-nucleus correlation time (Clore & Iwahara, 2009).

As it can be derived from equations (10) and (11), the distance can be expressed as:

$$r = \left(\frac{K}{\Gamma_2}\left(4\tau_c + \frac{3\tau_c}{1+\omega^2\tau_c^2}\right)\right)^{1/6} \qquad (12)$$

where K is a constant that contains the physical constants of equation (10) (γ, g, μ_B and S). Therefore, distances between the tag and specific nuclei can be determined by measuring τ_c and Γ_2. As the PRE effect decays with the inverse sixth power of the electron-nucleus distance, it can be estimated that for most of the available tags the influence of the paramagnetic center on the nuclei relaxation rates can be quantified up to a maximum of 35 Å.

Γ_2 can be quantified by measuring the ratio between the intensity of the peaks in the paramagnetic and diamagnetic states (Fig. 1.5.3). In this thesis work, a "single-time-point" approach is used, in which two HMQC spectra are recorded for each sample, one with the tag in the paramagnetic state and one with the tag in the diamagnetic state. The ratio between the peak heights in the two spectra depends on Γ_2 as:

$$\frac{I_{para}}{I_{dia}} = \frac{exp(-\Gamma_2 t_{hmqc})R_2^{diaHC}R_2^{diaH}}{(R_2^{diaHC}+\Gamma_2)(R_2^{diaH}+\Gamma_2)} \qquad (13)$$

where R_2^{diaH} and R_2^{diaHC} are the diamagnetic transverse relaxation rates for the ^1H single-quantum coherence and the ^1H-^{13}C multiple-quantum coherence, respectively and t_{hmqc} is the coherence transfer time during which the proton magnetisation is in the transverse plane. Thus, to extract Γ_2, R_2^{diaH} and R_2^{diaHC} must be known for each NMR peak. These rates are measured in separate NMR experiments.

The electron-nucleus correlation time τ_c can be estimated from Γ_2 rates that can be correlated with known electron-nucleus distances. In this work, as we know the structure of the individual proteins entering the complex, we could measure intra-molecular PREs and

correlate them with the electron-nucleus distances measured from the structure. In these experiments the paramagnetic tag is placed on an NMR active, isotope-labelled protein with a known structure. The experimental Γ_2, derived from equation (13) are then used to fit the Γ_2 predicted from the distances between the paramagnetic tag modelled on the protein structure and the corresponding protein nuclei. In the fitting procedure, both the conformations of the paramagnetic tag and τ_c are optimised to achieve convergence between the experimental and predicted Γ_2 (Iwahara *et al*, 2004) (the workflow for the determination of τ_c is shown in blue lines in Fig. 1.5.3).
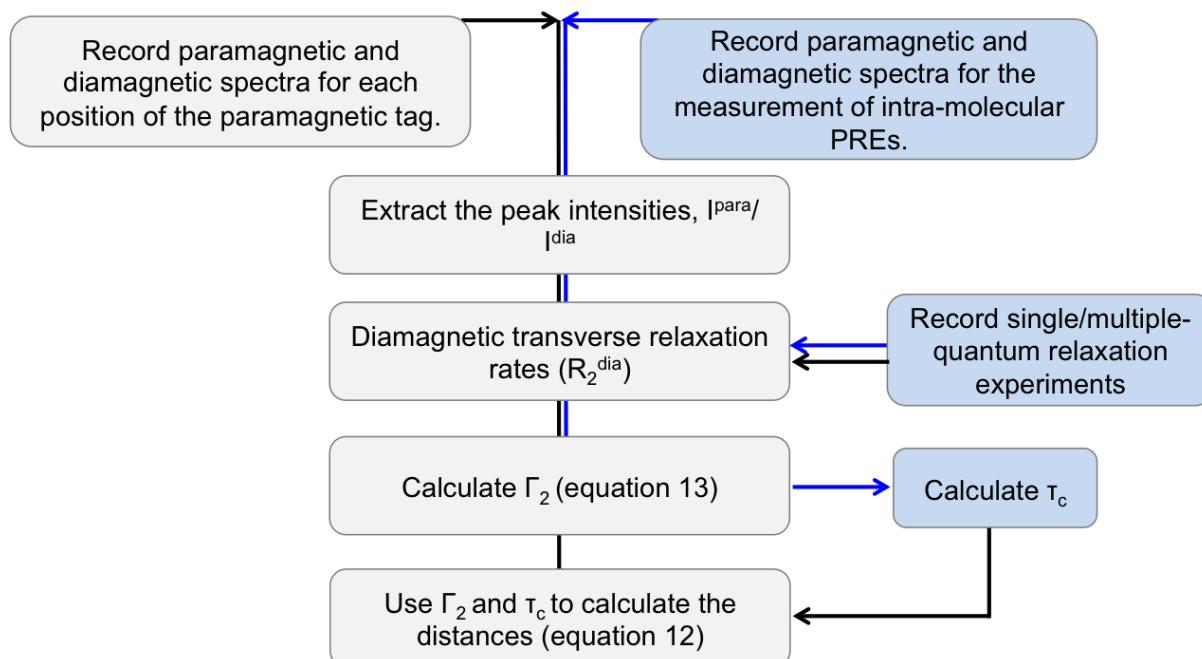
Figure 1.5.3. A simplified scheme of the distance calculation protocol used in this work. First, the diamagnetic transverse relaxation rates of each observed methyl-group are measured. After that, the τ_c is extracted from the intra-molecular PRE dataset, where the experimental Γ_2 rates are fit to the Γ_2 rates calculated from the known distances. Finally, the optimized τ_c can be used to calculate the unknown intermolecular distances.

Small-angle neutron scattering (SANS).

SANS provides low resolution structural information on particles in solution. Compared to small-angle X-ray scattering, SANS requires longer acquisition times and higher sample concentrations but it has key advantages for studying large biomolecular complexes: the possibility of studying the system in a wider range of solvents, absence of radiation damage, and, most importantly, the ability to gain structural data on individual components within a larger particle.

The typical layout of a SANS experiment is shown in Fig. 1.5.4:
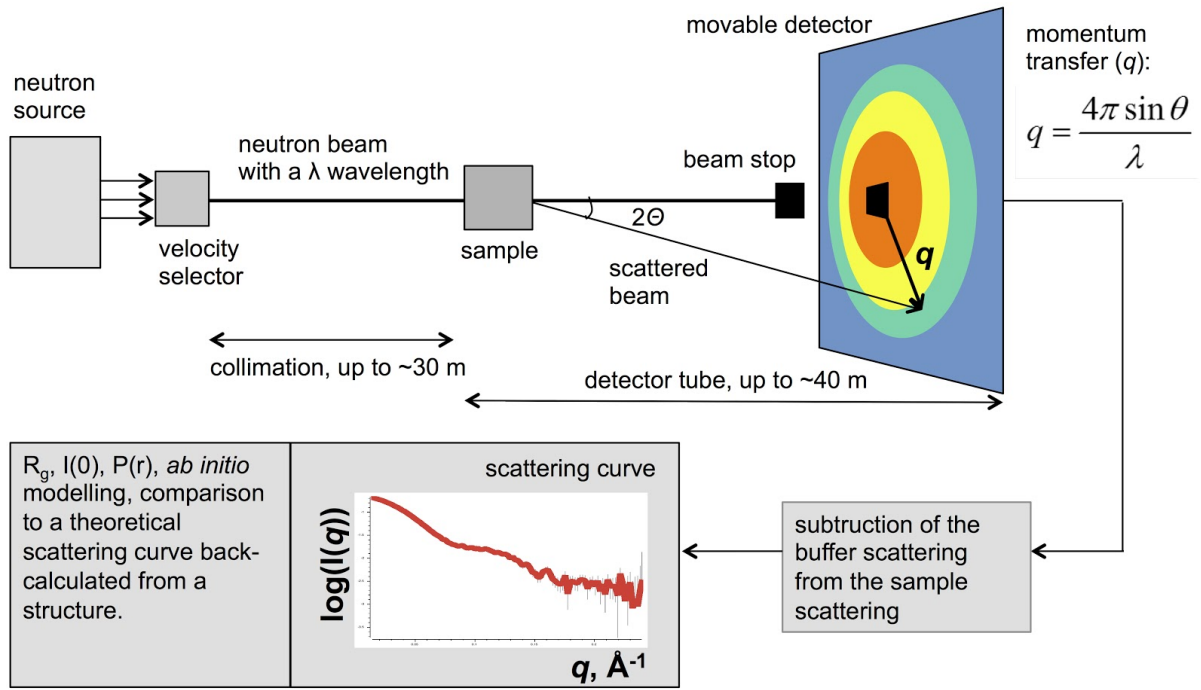
Fig. 1.5.4. SANS experimental scheme. The monochromatic neutron beam is shot at the sample and the resulting circularly symmetric scattering pattern is detected. The scattering pattern depends on the shape of the particle and the scattering intensity depends on its size and the concentration.

A neutron source is used to generate neutrons, which are moderated to the wavelength required for the particular measurement. SANS can be applied for particles with sizes ranging from ~1 nm up to ~500 nm; the length scale it probes (d) is proportional to the neutron wavelength (λ), and reversely proportional to the scattering angle (2Θ):

$$d \approx \frac{\lambda}{2\Theta} \qquad (14)$$

The neutron wavelength chosen for biological samples typically lies within 4-7 Å with a resolution ($\Delta\lambda/\lambda$) of ~5-10%. In the SANS experiment, elastic coherent neutron scattering allows to obtain structural information, when the resulting neutrons are scattered in different directions without energy transfer. In this case, the scattering intensity of a sample consisting of identical, randomly oriented particles depends on the momentum transfer, q (the distance of the direct beam to the detector, Å^{-1}) and can be expressed as:

$$I(q) = \frac{1}{V} \left| \int_V \Delta\varrho \, e^{-iqr} dr \right|^2 \qquad (15)$$

where V is the volume of the sample, r represents the coordinates of the individual scattering point, and $\Delta\varrho$ is the difference between the scattering length densities of the particle and the solvent (later referred to as the scattering contrast). The scattering length density (SLD) depends on the strength of the neutron-nucleus interaction and differs for isotopes of the

same element. The differences between the SLDs of ^{1}H and ^{2}H are exploited in a method called "contrast matching". The SLDs of ^{1}H and ^{2}H have different signs, and SLDs for various biomolecules typically have values between those of H_2O and D_2O (Fig. 1.5.5). Thus, it is possible to vary the concentration of D_2O in the solvent to match the SLD of selected biomolecules and render them "invisible". If a particle consists of different biomolecules (for example, proteins and nucleic acids) or biomolecules with different level of deuteration (for example ^{1}H and ^{2}H proteins), it is possible to use a D_2O:H_2O solvent mixture where a part of the particle is invisible, while another part contributes to the scattering signal (Jacrot, 1976).
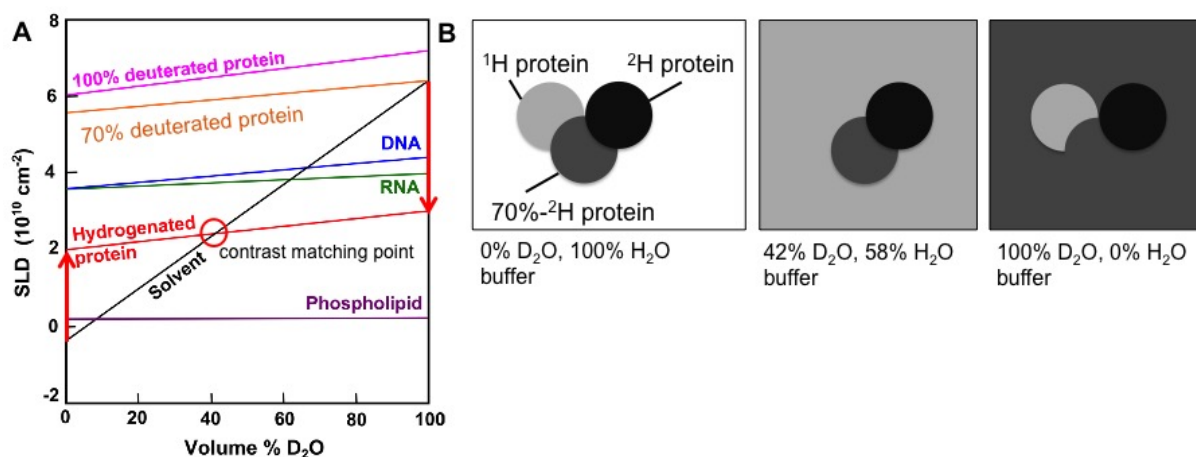
Fig. 1.5.5. Contrast matching in SANS. A. Plot of scattering length densities (SLDs) of the solvent and biomolecules depending on the percentage of D_2O in the solvent. The SLDs of DNA, RNA, proteins and phospholipids are between the SLDs of pure H_2O and D_2O. Thus, a contrast matching point between the solvent and the biomolecule (the intersection of the solvent and the sample lines on the plot) can be reached by varying the percentage of D_2O:H_2O. For example, the ^{1}H proteins have a positive scattering contrast in H_2O, are contrast-matched by the solvent in 42% D_2O and have a negative scattering contrast in 100% D_2O (the scattering contrast is indicated by the red arrows). **B.** Contrast matching labelling schemes used in this work. The panel (**A**) is adapted from (Dunne *et al*, 2017).

SANS data analysis. The structural information that can be derived from the scattering pattern includes the radius of gyration (R_g, root mean square distance of each scattering center to the center of mass), an estimated molecular weight of the particle, and a pair-distance distribution function (P(r), which represents the probability distribution of all pairwise interatomic distances within the particle).

The Guinier approximation postulates that at the small angles ($qR_g \leq 1 - 1.3$ depending on the particle shape) the scattering depends linearly on q^2 on a logarithmic scale (Feigin & Svergun, 1987; Guinier, 1939). The plot of the logarithm of the scattering intensity, $I(q)$, versus q^2 for this q range returns a straight line with a slope of ($R_g^2/3$):

$$\ln I(q) \approx \ln I(0) + \frac{q^2 R_g^2}{3} \qquad (16)$$

where $I(0)$ is the scattering intensity at the zero angle ($I(q)$ value at the intersection of the plotted line with the y-axis). Apart from deriving R_g and $I(0)$, the Guinier plot serves for the analysis of the quality of the sample, as deviation from the straight line at q values approaching zero indicates interparticle interactions such as aggregation or repulsion (Fig. 1.5.6) (Mertens & Svergun, 2010).
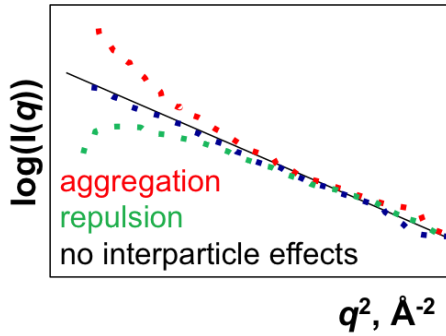
Figure 1.5.6. Guinier plot reports on the presence of inter-particle interactions. Deviation of the scattering plot from the straight line signifies attraction (red) or repulsion (green) of the particles.

The molecular weight (M_w) of the particle can be derived from $I(0)$ and depends on the concentration of the sample (c), the scattering contrast ($\Delta\varrho$), and the partial specific volume (v), depends on the type of the biomolecule, e.g. protein, DNA, RNA); N_A is the Avogadro constant:

$$I(0) = M_w \frac{c(\Delta\varrho)^2 v^2}{N_A} \qquad (17)$$

The determination of the M_w by this method is not very accurate, due to the assumptions on the sample homogeneity, the accuracy of the background subtraction, the precision of the concentration measurement and the completeness of the labelling; however, it can be used to assess the oligomeric state of the sample. An alternative method to determine M_w is based on the volume-of-correlation parameter (the ratio of $I(0)$ to the total scattering intensity) and is better suited for particles with non-spherical shape (e.g. unfolded proteins; the mass error is ~10%) (Rambo & Tainer, 2013).

P(r) is the distribution of the distances between the nuclei in the particle and can be calculated from the scattering intensity pattern:

$$P(r) = \frac{r^2}{2\pi^2} \int\limits_{q=0}^{q=\infty} q^2 I(q) \frac{sin(qr)}{qr} dq \qquad (18)$$

As the experimental scattering data do not reach from zero to infinite q, the P(r) is calculated by an indirect Fourier transformation, typically using the GNOM program (Svergun, 1992). For the calculation of the P(r) GNOM uses an estimated maximal dimension of the particle (D_{max}, ranging from $2R_g$ to $4R_g$ in the automated estimation) and a representation of the scattering curve as a sum of spline functions (smoothed curve). The quality of the D_{max}

estimation can be judged by the fitness of the smoothed scattering curve to the experimental one and by the behavior of the P(r) at the r=0 and r=D_{max} (Fig. 1.5.7).
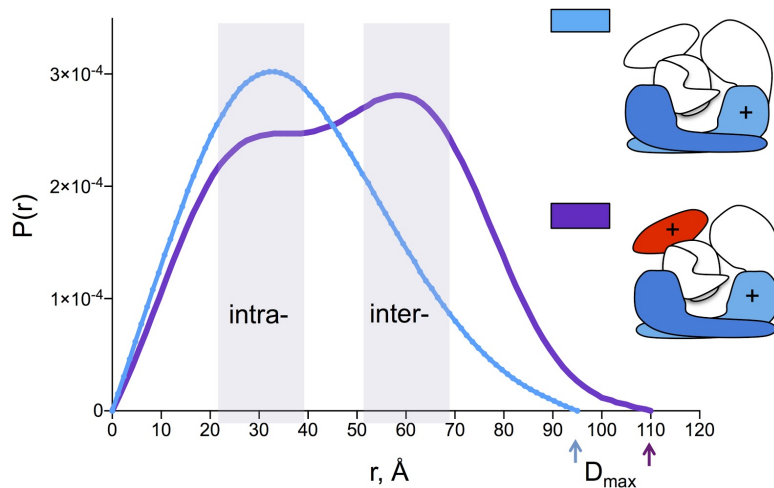
Figure 1.5.7. Pair-distance distribution. P(r) represents the probability distribution of distances between the two scattering centers within the particle. The P(r) is for a zero distance and approaches zero smoothly towards the D_{max}. Blue, P(r) for one scattering subunit (Vps75) within a contrast-matched protein complex: the D_{max} of Vps75 is 95 Å and the most probable intramolecular distance is ~31 Å. Purple, the P(r) for two scattering subunits (Vps75 and Asf1) within a contrast-matched protein complex: the P(r) function has a characteristic shape indicating a non-globular architecture of Vps75 and Asf1 together, the D_{max} of the combined subunits is 110 Å; the second maximum indicates the most probable intermolecular distance, hence the distance between the centers of mass of the two proteins (in this case ~62 Å).

As R_g can also be computed from the calculation of the P(r) by GNOM, it is important to compare the convergence of the R_g values from GNOM and from the Guinier analysis to ensure that both the data and the analysis are of high quality.

SANS in structural modelling. The scattering curve can be used either to generate an *ab initio* low-resolution structural envelope or to score independently generated structures (Fig. 1.5.8A).

A Generate low-resolution structural model:

exp. curve ──degeneracy──> *ab initio* structure

compare, χ²

back-calculated curve

Select structural model that fits the exp. data:

exp. curve   exp. structure

evaluation of the exp. structure

compare, χ²

back-calculated curve

B

exp. curve #1   structure A   χ²
                       structure B   χ²

comparison of χ² and χ² values reports on the best fitting structure

exp. curve #1   structure A   χ²
exp. curve #2   χ²

χ² and χ² values should not be be compared directly*

Fig. 1.5.8. Use of SANS data in structural modelling. A. While one structure corresponds uniquely to one scattering curve, the same scattering curve may generate from different structures. Thus, the *ab initio* structural model obtained from the scattering data has a low-confidence level (top panel). On the other hand, the unique scattering curve predicted from a given structure can be compared to the experimental curve to evaluate the fit of this structure to the SANS data (bottom panel) **B**. Due to the dependence of χ^2 on the experimental errors, the absolute χ^2 values for structures scored against different experimental curves cannot be compared directly (*the fitness of a structure to the different SAS datasets can be compared if the visual fit and the confidence in the χ^2 value are taken into account. For additional statistical methods see the review (Trewhella *et al*, 2017)).

A set of programs for *ab initio* modelling from the SAS data has been developed by the Svergun group (Franke *et al*, 2017). The generation of a dummy atom model can be done by DAMMIF, which builds the model from dummy atoms based on the agreement between the theoretical and experimental scattering curves and a set of parameters such as the model connectivity and compactness (Franke & Svergun, 2009; Svergun, 1999). However, the degeneracy of the SAS data, which arises from the loss of the orientation and phase information during the data collection, severely limits the resolution and confidence of the *ab initio* model.

In this work, the SANS data was mostly used for evaluation of the structural models generated by high-resolution methods (X-ray, NMR) and docking. The fit of the structural model to the experimental data is done with CRYSON, a program which predicts a theoretical scattering curve from the structure, taking into account the solvent shell and percentages of deuteration of the molecule and the solvent. The resulting theoretical curve is compared to the experimental one and their agreement is measured by the χ^2 parameter:

$$\chi^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left[\frac{I_{exp}(q_i) - cI_{mod}(q_i)}{\sigma(q_i)} \right]^2 \qquad (19)$$

where N is the number of data points, $I_{exp}(q)$ is the experimental scattering, $I_{mod}(q)$ is the theoretical scattering curve back-calculated from the structural model, c is a scaling factor for the theoretical curve required for the minimization of χ^2, $\sigma(q)$ is the standard error for each data point (Trewhella *et al*, 2017). As it can be seen from this equation, the χ^2 value depends on the experimental errors, as well as the number of data points along the detector axis. Error propagation, in turn, depends on the data reduction procedure, which is unique to each instrument, so the final error may be either over- or underestimated. Thus, while χ^2 values are useful to compare the fits of different models to the same scattering curve, the absolute χ^2 values do not unambiguously report on the fit quality, unless properly normalized, and should not be directly compared between different SAS curves (Fig. 1.5.8B).

Multi angle light scattering (MALS).

MALS is a technique for determination of the molecular weight (M_w) and radius of gyration (R_g) of particles in solution, which in turn can be used to assess interactions and oligomeric states of the biomolecules. Coupling of MALS with the size-exclusion chromatography (SEC) allows to define the above-mentioned parameters for the separate species present in the sample, thus enabling the analysis of heterogeneous samples.

In a MALS experiment, a laser beam is shot at the sample in solution and the scattering intensity measured at several, 2 to 20, fixed angles (Θ) by either a single movable detector or several fixed detectors. The basic principle of MALS is the proportionality of the resulting scattering, $I(\Theta)$, to the M_w, the concentration of the sample (c) and to the dn/dc ratio ($\Delta n/\Delta c$, where Δn is the change in the refractive index, Δc is the change in the concentration; reflects the polarization difference between the molecules and the solvent in a concentration dependent manner).

The reduced Zimm equation (Zimm, 1948) describes the amount of the scattered light in a M_w- and concentration-dependent way:

$$\frac{cK}{R_{(\Theta,c)}} = \frac{1}{M_w P(\Theta)} + 2A_2 c \quad (20)$$

where $R(\Theta,c)$ is excess Rayleigh ratio (the difference of the scattered light from the solution over the scattered light from solvent), $P(\Theta)$ is the angular dependence of the scattered light (includes R_g component), A_2 is the second virial coefficient describing the solubility of the biomolecule in solution, and K is a constant that contains the dn/dc:

$$K = 4\pi^2 (dn/dc)^2 n_0^2 / N_A \lambda^4 \quad (21)$$

where n_0 is the refraction index of the solvent, N_A is the Avogadro number and λ is the vacuum wave-lengths of the laser beam.

Due to the low sample concentration in each elution volume slice, the $2A_2c$ component of the equation (20) can be neglected (Johnson, 2005).

While the dn/dc value can be measured with the refractive index (RI) detectors, its values for different proteins vary very slightly and often the default value (0.185 mL/g) is used instead (recommended by Wyatt Technology). $R(\Theta,c)$, $P(\Theta)$ and the concentration are measured directly during the experiment, and the M_w determination in SEC-MALS can be done via a Zimm plot, where the scattering data from each elution volume slice is extrapolated to zero angle and zero concentration.

The R_g determination by SEC-MALS has a size limit of ~10 nm, as R_g is extracted from the anisotropic light scattering detected at different angles, and the angular dependence $P(\Theta)$ for smaller molecules is often undetectable.

Overall, SEC-MALS provides absolute mass determination, which does not depend on the elution time or the shape of the molecule. An example of the SEC-MALS profile of the molecular mass of the eluted species can be seen in the Fig. 4.1.3.

2. Aims of the thesis work

Previous studies have highlighted a unique feature of Rtt109: the dependence of its activity on the associated chaperones, Asf1 and Vps75. While the structures of Rtt109 and an Rtt109–Vps75 complex have been reported, these have given little information on the chaperone-dependent differences in the enzyme activity and specificity; instead they posed a question on the stoichiometry of the Rtt109–Vps75 complex. It remained unclear why Rtt109 requires two distinct histone chaperones to acetylate sites in the unfolded H3 N-terminal tail, while acetylation of H3 K56 requires only Asf1 *in vivo*.

Elucidating the mechanism by which Asf1 and Vps75 regulate Rtt109 activity could shed light on how enzyme binding partners can switch its specificity and how the acetylation of unfolded substrates can be promoted.

In order to answer these questions, the following objectives have been set in this thesis work:

1. To reconstitute the functional complex of Rtt109 with the chaperones and the substrate H3:H4 histones *in vitro* and determine its stoichiometry.

2. To explain the roles of Asf1 and Vps75 in Rtt109-dependent H3 acetylation by:

   - Solving the structural model of the complex by an integrative structural biology protocol combining NMR and SANS data.

   - Exploring the dynamics of the complex by molecular dynamics simulations (MD).

   - Defining the determinants of Rtt109 activity regulation by *in vitro* activity assays with proteins containing mutations designed under the guidance of the structural model.

3. Methods

The text of the sections 3.1 – 3.9 has been partially taken from the published manuscript (Danilenko *et al*, 2019), and has been originally written entirely by myself.

3.1. Protein expression and purification.

Plasmids for *Xenopus laevis* histone H3 (UniProtKb accession code P84233) and H4 (P62799), *S. cerevisiae* Vps75 (P53853) and *S. cerevisiae* Asf1 (P32447) were obtained from the Luger laboratory (University of Colorado Boulder). The *S. cerevisiae* Rtt109 sequence (Q07794), codon-optimised for expression in *E. coli*, was purchased from Geneart.

Xenopus laevis histones were used due to the better expression level in *E. coli* and high sequence identity to the *S. cerevisiae* H3 and H4 (Fig. 3.1.1).

```
H3 Xen   MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALR 50
H3 Xen#  MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALR 50
H3 Sc    MARTKQTARKSTGGKAPRKQLASKAARKSAPSTGGVKKPHRYKPGTVALR 50
         *********************:*******:*********** :*******

H3 Xen   EIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAY 100
H3 Xen#  EIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAY 100
H3 Sc    EIRRFQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAIGALQESVEAY 100
         ****:*********************************: **** : ***

H3 Xen   LVGLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA 136
H3 Xen#  LVALFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA 136
H3 Sc    LVSLFEDTNLAAIHAKRVTIQKKDIKLARRLRGERS 136
         **.*******.********* ***:****:****:

H4 Xen   MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGL 50
H4 Xen#  MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGL 50
H4 Sc    MSGRGKGGKGLGKGGAKRHRKILRDNIQGITKPAIRRLARRGGVKRISGL 50
         *********************:****************************

H4 Xen   IYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYG 100
H4 Xen#  IYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYG 100
H4 Sc    IYEEVRAVLKSFLESVIRDSVTYTEHAKRKTVTSLDVVYALKRQGRTLYG 100
         ****.*.*** ***.****.*************:::**************

H4 Xen   FGG 103
H4 Xen#  FGG 103
H4 Sc    FGG 103
         ***
```
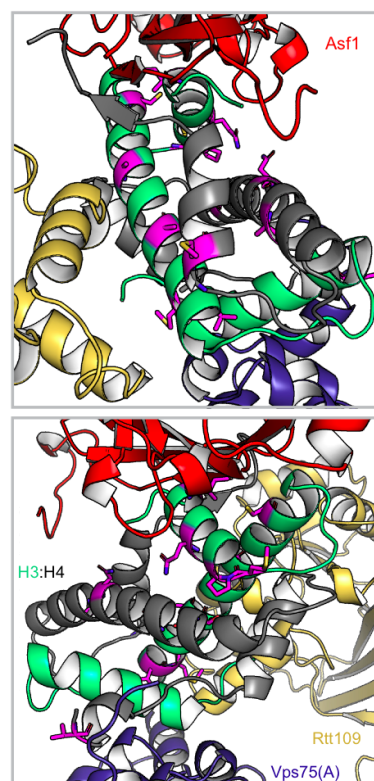

Figure 3.1.1. Comparison of the *Xenopus laevis* and *S. cerevisiae* histones H3 and H4. Left, sequence alignment of *Xenopus laevis* H3/H4 (Xen), histones used in this study (Xen#) and *S. cerevisiae* H3/H4 (Sc). Fully conserved, strongly similar and weakly similar residues are marked as asterisk, colon and dot, respectively. Right, Xen# H3/H4 residues different from the *Sc* H3/H4 sequence are shown as sticks and colored in magenta on the Rtt109–Vps75$_2$–Asf1–H3:H4 structure (PDB entry 6o22). The interfaces between Rtt109, Vps75 and the histones involve almost exclusively fully-conserved H3/H4 residues.

Asf1$^{2-169}$, Rtt109 and Vps75 were cloned into pETM11 vector (EMBL collection); the Vps75 sequence was also cloned into pACYC184 (addgene) and in araBAD (EMBL collection) vectors for co-expression with Rtt109. pET3d plasmids (Luger laboratory) were used for H3 and H4 expression.

Asf1, Vps75, and Rtt109 were expressed in *E. coli* BL21(DE3) cells. Cells were grown at 37 °C, expression was induced by addition of 0.5 mM IPTG at OD_{600} = 0.6–0.8 and continued for 14 – 16 hours at 16 °C in LB media or at 25 °C in 99% D_2O M9 minimal media. Cells were then harvested by centrifugation for 30 minutes at 4000g at 4 °C and lysed by sonication in 25 mM Tris–HCl pH 7.5, 500 mM NaCl, 5 mM β-mercaptoethanol (BME) with the addition of 1× protease inhibitors cocktail (Roche), 0.25 mg/ml lysozyme, and 1 µg/ml DNase I.

Vps75 was purified by immobilized-metal-ion affinity chromatography (IMAC) with a HisTrap HP 5 ml column (GE Healthcare). The column was equilibrated and washed with the HisTrap wash buffer (25 mM Tris–HCl pH 7.5, 500 mM NaCl, 5 mM BME, 20 mM imidazole) and Vps75 was eluted with a 0 – 100% gradient of the elution buffer (25 mM Tris–HCl pH 7.5, 500 mM NaCl, 5 mM BME, 400 mM imidazole) over 20 column volumes (CVs). The tagged protein was incubated with TEV protease (added in 1:50 protease:protein ratio) overnight at 4 °C. The cleaved protein was isolated by affinity purification with a Histrap HP column. Cleaved Vps75 was further purified by ion-exchange with a Hi-trap Q 5 ml column (GE Healthcare) in 25 mM Tris–HCl pH 7.5, 5 mM BME with a gradient from 0.15 to 1 M NaCl over 20 CV and size-exclusion chromatography (SEC) with a Superdex S200 16/60 column in 25 mM Tris–HCl pH 7.5, 500 mM NaCl, 5 mM BME. Asf1 was purified identically to Vps75, except for a different imidazole gradient (10 mM – 500 mM over 15 CV). Rtt109 was purified identically to Vps75, except for the ion-exchange step, which was performed with a Heparin HP 5 ml column (GE Healthcare) in 25 mM Tris–HCl pH 7.5, 5 mM BME, gradient from 0.3 to 1 M NaCl over 15 CV.

Expression and purification of histones H3 and H4 was customised based on the protocols from (Luger *et al*, 1999) and (Klinker *et al*, 2014), which rely on expression in inclusion bodies and refolding. The pET3d plasmid containing either H3 or H4 was transformed into *E. coli* BL21(DE3)pLysS cells. Expression was induced at OD_{600} = 0.6 with addition of 0.2 mM IPTG and continued at 37 °C for the next 2.5 hours in LB media or for the next 6 – 8 hours in 99% D_2O minimal media. Cells expressing either H3 or H4 were harvested by centrifugation

for 30 min at 4000g at 4 °C, resuspended in wash buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaC1, 1 mM Na-EDTA, 5 mM BME) and lysed by sonication on ice. Pellets containing inclusion bodies were washed twice with Triton wash buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaC1, 1 mM Na-EDTA, 5 mM BME 1% volume/volume Triton X-100) and twice with wash buffer. Pellets were then solubilized in the unfolding buffer (7 M guanidinium hydrochloride, 20 mM Tris-HCl, pH 7.5, 10 mM DTT). Solubilized histones were buffer-exchanged into SAU-200 buffer (7 M deionized urea, 20 mM sodium acetate, pH 5.2, 5 mM BME, 1 mM Na-EDTA, 200 mM NaCl) using a HiPrep 26/10 desalting column, which was followed by an ion-exchange step with HiTrap Q 5ml column in SAU-200. The ion-exchange flow-through was collected and loaded onto a HiTrap SP HP, where it was purified with a gradient from SAU-200 wash buffer to SAU-1000 elution buffer (7 M deionized urea, 20 mM sodium acetate, pH 5.2, 5 mM BME, 1 mM Na-EDTA, 1 M NaCl). H3 was eluted with the following gradient of the SAU-1000 buffer: 0 – 15% over 2 CV, followed by 15% of SAU-1000 for 3 CV, then 15 – 100% over 8 CV. The gradient steps for H4 were the same, except 15% of SAU-1000 was changed to 30%. All purification steps in the buffers containing urea were done at 4 °C in order to avoid carbamylation. Purified histones were either used for reconstitution immediately or flash-frozen and stored at –80 °C.

Perdeuterated and 70% deuterated proteins for SANS measurements were expressed in 100% D_2O M9 minimal medium supplemented with ^{2}H-glycerol or ^{1}H-glucose, respectively. ^{1}H, ^{15}N-labeled proteins were produced in H_2O M9 minimal medium supplemented with $^{15}NH_4Cl$. Isoleucine, Leucine and Valine (ILV) methyl-labeled proteins were expressed in 100% D_2O M9 minimal medium supplemented with 60 mg/L of α-ketobutyric acid sodium salt (Sigma) and 120 mg/L of α-ketoisovaleric acid sodium salt (Sigma), added 40 minutes before the induction. Pro-S LV methyl labeling was achieved with the NMR-Bio precursor (DLAM-LV^{proS} kit). Labeled proteins were purified identically to the unlabelled ones with an exception for Rtt109. Perdeuterated or 70% deuterated full-length Rtt109 was co-expressed with Vps75 in order to increase the stability of Rtt109 and reduce the amount of required minimal media and precursors. *E. coli* BL21(DE3) cells were transformed with both pETM11 and pACYC184 vectors carrying Rtt109 and Vps75, respectively. Expression and cell lysis were performed as described above. Both histidine-tagged proteins were purified with HisTrap, after which Rtt109 was separated from Vps75 with a Heparin column (25 mM Tris–HCl pH 7.5, 5 mM BME, gradient from 0.3 to 1 M NaCl over 15 CV). The C-terminally truncated Rtt109 (Rtt109$^{1-424}$) could not be separated from Vps75 by ion-exchange chromatography. Thus, Rtt109$^{1-424}$ was expressed and purified without Vps75, independent

of the labelling scheme. The expression and purification protocol for Rtt109[1-424] was identical for the full-length unlabelled Rtt109.

Final protein concentrations were determined by UV absorption using the following molar extinction coefficients (ε_{molar}, in M^{-1} cm^{-1}): 4470 for H3; 5960 for H4; 43890 for Vps75; 21430 for Asf1; 49850 for Rtt109.

3.2. Reconstitution of the Asf1–H3:H4–Rtt109–Vps75$_2$ complex.

The complex reconstitution consisted of three main steps, involving a preparation of the H3:H4 dimer and the reconstitution of the Asf1–H3:H4 sub-complex at high salt concentration, and a final step consisting of mixing the Asf1–H3:H4 and the Rtt109–Vps75$_2$ sub-complexes at lower salt concentration (150 mM NaCl) in order to allow for the formation of the full complex. The best results were obtained when the Asf1–H3:H4 sub-complex was kept in high salt (> 0.5 M NaCl) and at 4 °C until mixed with the Rtt109–Vps75$_2$ by fast pipetting in the NMR/SANS buffer (50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME). Subsequent addition of first Vps75$_2$ and then Rtt109 to the Asf1–H3:H4 sub-complex lead to the same result as addition of the Rtt109–Vps75$_2$ sub-complex.

Histones H3 and H4 were mixed in equimolar ratio in 7 M urea, 10 mM Tris–HCl pH 7.5, 1 M NaCl, 1 mM Na-EDTA, 5 mM BME and dialysed at 4 °C into the refolding buffer (10 mM Tris–HCl pH 7.5, 2 M NaCl, 1 mM Na-EDTA, 5 mM BME). The H3:H4 dimer was then separated by SEC in refolding buffer. The Asf1–H3:H4 sub-complex was prepared by mixing the H3:H4 dimer with Asf1 at a 1:1.1 molar ratio in 50 mM Tris–HCl pH 7.5, 0.5 M NaCl, 5 mM BME followed by SEC.

The Rtt109–Vps75$_2$ sub-complex was reconstituted with a 1.1:1 Rtt109:Vps75$_2$ ratio in the NMR/SANS buffer and subsequently purified by SEC.

The full complex was assembled by mixing the two sub-complexes in a 1:1 ratio followed by a SEC in the NMR/SANS buffer. All size-exclusion chromatography runs were performed with a Superdex S200 10/300 column (GE Healthcare).

3.3. Size exclusion chromatography - multi angle light scattering (SEC-MALS) experiments.

The molecular weights of the Rtt109–Vps75 and Asf1–H3:H4–Rtt109–Vps75$_2$ complexes were verified by SEC-MALS. The experiments were performed on an AKTA Pure system at room temperature (25 °C) using a Superdex S200 10/300 column (GE Healthcare) in-line with the refractive index detector Optilab T-rEX and the miniDAWN TREOS MALS detector (WYATT technology). The data was collected and analyzed using the ASTRA 7.1.4 software (WYATT technology).

All samples were run in the NMR/SANS buffer. The Asf1–H3:H4–Rtt109–Vps75$_2$ sample was obtained by mixing Rtt109–Vps75$_2$ and Asf1–H3:H4 sub-complexes in a 1:1 molar ratio, the sample concentration was 6.5 mg/mL. The Rtt109$_2$–Vps75$_2$ sample was prepared by mixing Rtt109 and Vps75 in a 2:2 molar ratio in the NMR/SANS buffer, the sample concentration was 4.3 mg/mL. Before each experiment the delay volume was calibrated with bovine serum albumin (BSA). The refractive index and viscosity of the NMR/SANS buffer were defined as 1.331 and 1.05 cP. The dn/dc (refractive index increment value) was set to 0.185 mL/g (default value for protein samples).

3.4. NMR assignment of methyl groups and backbone.

Assignment experiments for Vps75 were performed on ^{2}H,^{13}C,^{15}N-ILV-methyl-protonated Vps75$^{1–225}$ at 0.8 mM protein concentration in 20 mM Tris–HCl pH 7.5, 200 mM NaCl, 1 mM DTT at 300 K. The assignment strategy is described below and schematically shown in Fig. 3.4.1.

Backbone amide resonances were assigned using TROSY-HNCA, TROSY-HNCO, TROSY-HN(CA)CO, TROSY-HN(CO)CA and TROSY-HNCACB experiments (Salzmann *et al*, 1998). A subset of the ILV methyl resonances (I28, I54, I60, I127, I138, I155, I160, I179, V121, L16 & L223) were assigned through Ile,Leu-(HM)CM(CGCBCA)NH and Val-(HM)CM(CBCA)NH experiments (Tugarinov & Kay, 2003). Methyl resonances of I28, I54, I72, I127, V25, V32, V61, V124, V213 & L188 were assigned or verified via point mutations. The remaining methyl resonances were assigned using a 4D HCCH HMQC-NOESY-HMQC spectrum recorded with non-uniform sampling (NUS) (Vuister *et al*, 1993) in combination with the predicted NOEs from the PDB entry 2zd7 (Vuister *et al*, 1993; Park *et al*, 2008).

Stereospecific assignment of the L and V methyl resonances was achieved using pro-S LV-labeling. Identification of the prochiral methyl-groups of L/V residues was performed with a 3D experiment designed by Dr. J. Kirkpatrick, in which the ^{1}H and ^{13}C resonances of the methyl groups were correlated with the ^{13}C resonances of Cγ and Cβ of leucines and valines, respectively.

Through-bond: backbone assignment, connection of the ILV methyl group resonances to the backbone resonances, identification of L/V prochiral methyl groups.
I28, I54, I60, I127, I138, I155, I160, I179, V121, L16, L223

Mutational analysis: point mutations of ILV residues.
I28, I54, I72, I127, V25, V32, V61, V124, V213, L188

Through-space: NOESY experiments in combination with the known structure and assignments. All remaining assignments and confirmation of previous assignments.

Stereospecific assignment: ProS labeling.
V88 and confirmation of previous assignments.

Figure 3.4.1. Workflow chart for the assignment of the ILV methyl resonances of Vps75$^{1–225}$.

The ILV methyl-group assignments were then transferred from Vps75$^{1–225}$ to the free full-length Vps75 in the NMR/SANS buffer and then to the Vps75 in the full complex following the CSPs in Vps75 titrations with Asf1–H3:H4 or Rtt109 (Figure 4.7.1).

The backbone amide resonances of the Vps75 C-terminal acidic domain (CTAD, residues 226–264) were assigned using HNCO, HNCACB and HN(CO)CACB experiments (Grzesiek & Bax, 1992; Wittekind & Mueller, 1993; Ikura *et al*, 1990a) recorded on ^{1}H,^{13}C,^{15}N-labeled full-length Vps75 at 0.6 mM in the NMR/SANS buffer.

Isoleucine methyl resonances of Rtt109 were assigned with a 3D (H)CCH HMQC-NOESY-HMQC spectrum recorded on perdeuterated, ^{13}C,^{1}H-methyl-I-labeled Rtt109 at ~0.6 mM in complex with perdeuterated Vps75 in 100% D$_2$O NMR/SANS buffer. Experimental NOEs were compared to those predicted from PDB entry 3q66 (Tang *et al*, 2011). Prediction of NOEs expected from the structure was done by selection of pairs of Ile, Leu and Val methyl groups which are closer than 6.5 Å to each other.

Assignment experiments for the histones were recorded on 0.66 mM ^{13}C,^{1}H-methyl-ILV-labeled H3:H4 in complex with perdeuterated Asf1 in 20 mM Tris–HCl pH 7.5, 1 M NaCl, 1 mM Na-EDTA, 5 mM BME in 100% D$_2$O. The methyl groups were assigned

with 3D (H)CCH-HMQC-NOESY-HMQC and 3D HCH HMQC-NOESY (Ikura *et al*, 1990b; Zuiderweg *et al*, 1990) spectra in combination with the NOEs predicted from PDB entry 2hue (English *et al*, 2006). To facilitate the assignment, methyl resonances belonging to H4 were identified from a ^{13}C-HMQC spectrum of ILV methyl-labeled H4 in the context of the Asf1–H3:H4 complex.

The Asf1 assignment was performed by Dr. L. Lercher and described in (Lercher *et al*, 2018).

Chemical shift perturbations (CSPs) were calculated according to:

$$CSP = \sqrt{\tfrac{1}{2}(\delta_H{}^2 + (0.3\,\delta_C)^2)} \text{ or } CSP = \sqrt{\tfrac{1}{2}(\delta_H{}^2 + (0.15\,\delta_N)^2)} \qquad (22)$$

Where δ_H is the distance between ^{1}H chemical shifts, δ_C is the distance between ^{13}C chemical shifts, δ_N is the distance between 15N chemical shifts; 0.3 and 0.15 are the scaling factors applied to weight δ_C and δ_N respectively (Williamson, 2013).

3.5. Paramagnetic relaxation enhancement (PRE) experiments and generation of the M3 distance restraints.

PRE datasets were acquired for eight paramagnetic tags located at different positions on the H3:H4 dimer to measure distance-restraints to Vps75- or Rtt109-methyl groups in the Asf1–H3:H4–Rtt109–Vps75$_2$ complex (H3 R63C, H3 Q76C, H4T30C, H4 R45C, H4 T82C, H4 Q93C to Vps75; H3 K115C and H4 Y98C to Rtt109).

Single-cysteine mutations were introduced in H4 or in H3-C110A, where the wild-type cysteine was substituted by alanine. Mutated histones were then purified and refolded as the wild-type proteins, followed by SEC in the refolding buffer without reducing agent (10 mM Tris–HCl pH 7.5, 2 M NaCl). The paramagnetic agent (3-(2-iodoacetoamido)-PROXYL radical (Sigma-Aldrich, cat. no. 253421) was added to the H3(C110A):H4 dimer containing the single cysteine mutation immediately after the SEC step. The coupling between the paramagnetic agent and the free cysteine was allowed to proceed overnight at 4 °C in the dark. The unreacted paramagnetic agent was removed during an additional SEC step in 10 mM Tris–HCl pH 7.5, 2 M NaCl, after which H3(C110A):H4 dimers, containing the paramagnetic tag, were used for reconstitution of the complex. Prior to complex reconstitution, performed as described in section 3.2, Rtt109, Vps75, and Asf1 were

exchanged into the NMR/SANS buffer with no reducing agent using a HiPrep 26/10 desalting column (GE Healthcare). The complexes were measured in 99% D_2O NMR/SANS buffer without BME on a Bruker Avance III 850 MHz spectrometer equipped with a TCI cryoprobe at 298 K. Sample concentrations varied between 30 and 90 µM.

^{1}H-^{13}C HMQC spectra were acquired for the paramagnetic and diamagnetic (reduced with 2 mM ascorbic acid) states of the complexes. In order to derive the peak intensity in the paramagnetic (I^{para}) and the diamagnetic (I^{dia}) states, the peaks were fitted with FuDA (Hansen Laboratory, https://www.ucl.ac.uk/hansen-lab/fuda/) assuming Lorentzian line-shapes (spectra were processed using exponential apodization in both dimensions). Overlapped peaks were fitted as groups.

The fitted volumes and line-widths from FuDA were converted into peak-heights using:

$$I = \frac{V}{Lw_H Lw_C} \qquad (23)$$

where I is the peak height, V is the peak volume and LW_H and LW_C are the fitted line-widths in the ^{1}H and ^{13}C dimensions, respectively. The peak-height ratios (I^{para}/I^{dia}) were calculated for each data set and the experimental errors for each I^{para}/I^{dia} ratio were derived by the error propagation rules from the standard deviation of the noise in the HMQC spectra.

In case of Vps75 only well-resolved peaks unambiguously corresponding to Vps75(A) or Vps75(B) were used to quantify distances between H3:H4 and ILV methyl-labeled Vps75. Additional peaks were fitted for the data-sets recorded with H3-Q76C and H4-T82C, where preliminary structural calculations, as well as the intensity ratios obtained for the well-resolved peaks of Vps75(B), indicated that the paramagnetic tags are far away from the Vps75(B) earmuff domain (> 40 Å). This allowed the selection of five additional peaks, for which only one of the two overlapped resonances of Vps75(A) and Vps75(B) was assumed to be affected by the paramagnetic tag. Hence, the (I^{para}/I^{dia}) height ratio of the portion of the peak corresponding to Vps75(B) was presumed to be equal to 1 and resulting restraints were set unambiguously to Vps75(A) and Vps75(B). In addition, the peaks for L159 were fitted for all datasets despite the overlap: the distance between the methyl groups of L159 in Vps75(A) and Vps75(B) was sufficient to assume that only one of these residues could be affected by the paramagnetic tag regardless of the tag position. In this case, the I^{para}/I^{dia} ratios were calculated as described above, but the resulting restraints were assigned ambiguously to either Vps75(A) or Vps75(B).

Diamagnetic transverse relaxation rates for the ^{1}H single-quantum coherence (R_2^{diaH}) and the ^{1}H-^{13}C double-quantum coherence (R_2^{diaHC}) were quantified using the pulse schemes reported in (Korzhnev *et al*, 2004; Tugarinov & Kay, 2006, 2013). Relaxation delays for the quantification of R_2^{diaHC} of both Vps75 and Rtt109 methyl resonances were 0, 2, 4, 7, 10, 13 & 20 ms. Relaxation delays for the quantification of R_2^{diaH} were 0, 2, 4, 6, 8, 10, 13 & 16 ms for Vps75 resonances, and 0, 3, 6, 9, 12, 16 & 20 ms for Rtt109 resonances. The peak-heights were fitted to a mono-exponential decay function to extract R_2^{diaH} and R_2^{diaHC}.

The resulting R_2^{diaH}, R_2^{diaHC} and the I^{para}/I^{dia} values were used to calculate the transverse PRE rates (Γ_2) according to equation (13), section 1.5:

$$\frac{I_{para}}{I_{dia}} = \frac{exp(-\Gamma_2 t_{hmqc})R_2^{diaHC}R_2^{diaH}}{(R_2^{diaHC}+\Gamma_2)(R_2^{diaH}+\Gamma_2)}$$

where t_{hmqc}, equal to 7.7 ms, is the total duration of the constant-time delays in the ^{13}C-HMQC pulse-sequence during which the magnetization is transverse on ^{1}H. The PRE rates were converted into distances using equation (12), section 1.5:

$$r = (\frac{K}{\Gamma_2}(4\tau_c + \frac{3\tau_c}{1+\omega^2\tau_c^2}))^{1/6}$$

where K is a constant (1.233×10^{-23} cm^{6}s^{-2}), ω is the proton Larmor frequency and τ_c is the correlation time of the electron–nucleus vector (81 ns). The correlation time was estimated by measuring the PRE values of Vps75 methyl groups in the presence of a paramagnetic tag on Vps75-E56C as part of the full complex. The measured PRE values were used as restraints for the simultaneous optimization of the tag conformations and τ_c via the PRE potential function in Xplor-NIH developed by (Iwahara *et al*, 2004) using a protocol written by Nick Anthis. The minimization was run for 20 structures with the "obsig" setting for error-weighting of the PREs, which yielded a τ_c of 81±7 ns.

The above-described protocol to derive the distances from the PRE measurements was developed in our group by Dr. J. Kirkpatrick, who also kindly provided the scripts for the extraction of the peak heights, PRE rates and τ_c.

The resulting distances were converted into M3 restraints as follows: the restraints were set between the Cβ of the tagged residues and the corresponding methyl groups; in order to account for the tag length and flexibility, an error of ± 6 Å was added to the experimental error. For I^{para}/I^{dia} ratios below 0.2 the distance was set between 2.0 and 18.0 Å; for I^{para}/I^{dia} ratios above 0.8 the distance was set between 24.0 and 99.0 Å. The distances

corresponding to the I^{para}/I^{dia} ratios below 0.2 and above 0.8 were estimated from the Vps75-E56C dataset.

3.6. Small-angle neutron scattering.

SANS experiments were performed on complexes reconstituted with different combinations of uniformly perdeuterated, 70% deuterated and protonated proteins were measured in the SANS buffer containing 0%, 42% or 100% D_2O. The complex concentrations varied between 11,5 and 36 µM (1.9 mg/mL and 6 mg/mL concentration of both scattering and contrast-matched proteins) (Appendix 1).

Seven measurements were carried out at the KWS-1 beamline (Frielinghaus *et al*, 2015) at the Juelich Centre for Neutron Sciences (JCNS outstation at MLZ, Garching, Germany). Two detector configurations of 1.5 m and 4 m were used for each sample with a collimation length of 4 m and a neutron wavelength of 5 Å. The remaining experiments were performed at the D22 instrument at the Institut Laue-Langevin (ILL, Grenoble, France) with a sample-detector distance of 4 m, collimation length of 4 m and a neutron wavelength of 6 Å.

All samples were measured at 298 K. Data reduction and radial integration were carried out according to the standard procedures with ILL and JCNS specific software. Buffer subtraction was done with PRIMUS (Konarev *et al*, 2003). Fits of the experimental data to the models were performed with CRYSON (Svergun *et al*, 1998) .

The *ab initio* models of the Rtt109−Vps75$_2$−Asf1−H3:H4, Rtt109−Vps75$_2$$^{1-225}$−Asf1−H3:H4 and Rtt109−Vps75$_2$$^{1-225}$−Asf1−H3$^{35-135}$:H4 complexes were generated with DAMMIF (Franke & Svergun, 2009). 20 independent *ab initio* models were calculated for each complex and averaged with DAMAVER. The averaged *ab initio* models were then filtered with DAMFILT.

3.7. Structure calculation.

The structural models of Rtt109−Vps75$_2$−Asf1−H3:H4 complexes were calculated with the M3 protocol (Karaca *et al*, 2017) employing HADDOCK 2.2.
The building blocks were defined as Asf1−H3:H4 and Rtt109−Vps75$_2$ sub-complexes. The structure for the Asf1–H3:H4 building block was extracted from the PDB entry 2hue (the structure included Asf1$^{1-164}$, H3$^{60-134}$, H4$^{20-101}$). The second building block was comprised of an ensemble of ten Rtt109–Vps75$_2$ structures (all structures included Rtt109$^{1-418}$, Vps75(A)$^{9-225}$,

Vps75(B)$^{2-231}$). The Rtt109–Vps75$_2$ conformers were generated by a 400-ns MD simulation (described in section 3.8) and selected with respect to the normalized consensus χ^2 between the calculated and experimental SANS curves of the complexes containing ^{2}H-Vps75$^{1–225}$ and ^{2}H-Rtt109/^{2}H-Vps75$^{1–225}$. The normalized χ^2 was calculated as:

$$\chi^2_{norm} = \frac{\chi^2 - \chi^2_{min}}{\chi^2_{max} - \chi^2_{min}} \qquad (24)$$

where χ^2_{norm} is the normalized χ^2, χ^2_{min} and χ^2_{max} are the minimal and maximal χ^2 for the dataset, respectively.

The M3 docking was driven by 145 unique PRE-derived distance restraints between Vps75/Rtt109 and H3/H4 (described in section 3.5). An additional restraint was added to loosely limit the distance between the Rtt109 C-terminus and Asf1 to ensure that the Rtt109 C-terminal tail can interact with Asf1 as reported in (Lercher *et al*, 2018). The complete restraint table can be found in the appendix 2.

The rigid body docking stage (it0) was performed with the default HADDOCK parameters, sampling of 180°-rotated molecules was disabled. As standard in HADDOCK, all residues were kept rigid at this stage. The resulting 5,000 structures, corresponding to 50,000 sampled conformers, were ranked using a restraint energy-based scoring. The E_{exp} violation energy represents the agreement between the experimental and back-calculated distance restraints. Only 15 structures with significantly lower ln(E_{exp}) values were identified as outliers in box-and-whisker plots with a whisker length of two times the interquartile range. Each of these 15 structures was submitted to ten independent runs of flexible docking (it1). The following parameters were changed to allow an extended search during it1: the temperature of the rigid-body torsion-angle-dynamics (TAD) search was increased to 5,000 K; the number of steps was increased to 20,000; the number of rigid-body cooling steps was increased to 20,000; the factor time step of annealing was decreased to 4. During this stage side-chains of residues at the interfaces, defined by HADDOCK based on the analysis of intermolecular contacts within a 5-Å cut-off, were flexible.

The resulting 150 structures were scored against the 6 SANS datasets using CRYSON. The χ^2 values for different datasets were individually normalized using rank-preserving normalization (equation 5). The fitness parameter represented the sum of six χ^2_{norm} values corresponding to the structure and evaluated the overall fit of the structures to the SANS data. The structures were then clustered by orientational RMSD (o-RMSD), which was

calculated using the scripts adapted from https://github.com/ezgikaraca/ISD-files (Karaca *et al*, 2017). The cluster containing the structure with the best fitness to the SANS data was defined as the best cluster (green in Fig. 4.4.6). 33 members were selected from this cluster to represent the final ensemble; the selected structures had better fitness than any member of the second best cluster. The structure with the lowest average backbone root-mean-square deviation (RMSD) to the 33 selected structures was chosen for further minimization in water (described in section 3.8) and subsequent deposition to the PDB (entry 6o22).

3.8. Molecular dynamics simulations.

MD simulations for the selection of the Rtt109–Vps75$_2$ building blocks were set up by Dr. L. Lercher and described in (Danilenko *et al*, 2019). The PDB entry 3q66 was selected as a starting structure. The simulation was performed at 298 K with Gromacs 5.1.1 and the AMBER99SB-ILDN force-field (Abraham *et al*, 2015; Lindorff-Larsen *et al*, 2010; Price & Brooks, 2004). The system consisted of a dodecahedral box, filled with explicit TIP3P water ((Abraham *et al*, 2015; Lindorff-Larsen *et al*, 2010; Price & Brooks, 2004)). The Na$^+$ and Cl$^-$ ions were added to neutralize the total charge. The system was minimized with the steepest descent method and then subjected to a two-phase solvent equilibration step: as a canonical ensemble (NVT) and as a isothermal-isobaric ensemble (NPT) for 1 ns each. The unrestrained MD simulation was then allowed to proceed for 400 ns. The Rtt109–Vps75$_2$ conformers were extracted from each 1 ns frame, resulting in 400 structures.

The MD simulation for the water refinement of the Rtt109−Vps75$_2$−Asf1−H3:H4 complex was performed at 300 K using AMBER 2018 with the AMBER14SB force-field (Maier *et al*, 2015; Case *et al*, 2018). The experimental structure of the complex, calculated as described in section 3.7, was used as a starting conformation. The system was simulated in explicit TIP3P water and consisted of a cubic box with dimensions of 138.22, 134.21 and 117.14 Å. The box encompassed a water-layer of 14 Å thickness to account for instabilities in the electrostatic potential. The total charge was neutralized by addition of Na$^+$ and Cl$^-$ ions, additional ions were also included to match the experimental NaCl concentration (150 mM). The system minimization was done in two stages. First, the solvent was equilibrated for 20,000 steps with NPT and heated to 300 K with NVT. Next, the complete system was minimized for 20,000 steps, and heated to 300 K with the same procedure. Finally, the system was relaxed with restraints on the protein heavy atoms with NPT for 0.5 ns at 300K.

The system was subjected to an additional 10 ns MD equilibration run and the resulting structure was deposited as PDB entry 6o22.

MD simulations to explore the dynamics of the H3$^{35-59}$ tail within the Rtt109−Vps75$_2$−Asf1−H3:H4 complex consisted of three stages: system minimization, equilibration and production run. The system minimization and equilibration were performed as described above for the Rtt109−Vps75$_2$−Asf1−H3:H4 complex. The starting structure was modified by the covalent addition of the H3$^{35-59}$ tail with Chimera Modeller(Pettersen *et al*, 2004), by K290 acetylation and by the non-covalent addition of acetyl coenzyme A (acCoA) with PYMOL (Schroedinger *et al*, 2015). The latter modification was achieved by alignment of the Rtt109 molecule from the complex model to the Rtt109 structure from PDB entry 3q35 and subsequent replacement of the coordinates. The replaced coordinates included residues 191–213, corresponding to the acCoA binding loop, and addition of acCoA. Partial charges and specific improper angles of acCoA and acetylated lysine were generated according to(Papamokos *et al*, 2012). The complex structure with H3-K56 placed in the Rtt109 catalytic pocket was obtained via a 100 ns MD production run with center-of-mass restraints on each protein and a restraint between H3-K56 and acCoA. Two resulting complex structures (with the extended H3$^{35-59}$ conformation or the H3$^{35-59}$ conformation with H3-K56 in the Rtt109 active center) were then subjected to 100 ns MD production runs with center-of-mass restraints on the proteins. These production runs were analyzed to study the conformational preference of the H3 N-terminal region and its dependence on the starting conformation.

Analysis of the trajectories was performed with CPPTRAJ and VMD v 1.9.4a29 (Humphrey *et al*, 1996).

The set up of the system (choice of the force-field, including parameters for the acetylated K and acCoA, optimisation of the system minimization and equilibration parameters) was done by Dr. L. Codutti.

3.9. Activity assays.

Acetylation activity of Rtt109 was quantified using the dot-blot technique. The Asf1–H3:H4 and Rtt109–Vps75$_2$ sub-complexes were reconstituted by mixing the individual components followed by SEC (Superdex S200 10/300; 10 mM HEPES, pH 8.0, 0.1 M NaCl, 5 mM BME for Rtt109–Vps75$_2$; 10 mM HEPES, pH 8.0, 0.5 M NaCl, 5 mM BME for Asf1−H3:H4). Acetylation reactions were performed using 0.2 µM Rtt109–Vps75$_2$ sub-complexes, 0.2 µM Asf1–H3:H4 and 2 µM Ac-CoA in 10 mM HEPES pH 8.0, 100 mM NaCl. Reactions were

stopped with 0.5 M sodium acetate, pH 5.5; 1.6 µL of the reaction mixtures were spotted on a nitrocellulose membrane. Membranes were air-dried for ca. 30 min, blocked for 30 min in 5% nonfat dry milk–TBST buffer (20 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20), and subsequently incubated for 1 h in a 1:2000 dilution of primary antibody in TBST: rabbit anti-histone H3 (acetyl K9) (ab4441 Abcam, lot no. lot no. GR304381-1 for the experiments in Fig. 4.3.4, GR3229436-1 for the remaining experiments) or rabbit anti-histone H3 (acetyl K56) (SAB5600015 Sigma, lot no. P1100739). The membranes were then washed three times with TBST buffer and incubated for 1 h with the secondary antibody, anti-rabbit IgG (whole molecule) alkaline phosphatase (A3687 Sigma, lot no. SLBK3154V), 1:10000 solution in TBST. After three washes with TBST buffer, the blots were developed with BCIP®/NBT Liquid Substrate System (B1911 Sigma, lot no. SLBS4876).

The intensities of the dots were quantified with ImageJ. After background subtraction the measured intensities were normalised relative to the negative control (reaction without Rtt109–Vps75$_2$). For the blots obtained with H3 (acetyl K9) lot no. GR304381-1 the intensities were normalised to the control reaction with 0.2 µM Asf1–H3:H4. For the remaining experiments (antibodies H3 (acetyl K9) lot no. GR3229436-1; H3 (acetyl K56) lot no. P1100739) the control reaction with 0.2 µM Asf1–H3:H4 did not give any signal above the background. An additional negative control with 6 µM Asf1–H3:H4, which corresponds to column "c" in the blots of Figs. 4.5.4, 4.6.3, 4.6.4, was used for the normalization of the dot density and comparison of the experimental repeats.

4. Results

4.1. Rtt109–Vps75 stoichiometry: the Vps75 dimer accepts one copy of Rtt109.

The available crystal structures of the Rtt109–Vps75 complex (Kolonko *et al*, 2010; Tang *et al*, 2011; Su *et al*, 2011) display two different stoichiometries, 1:2 and 2:2. In order to resolve this controversy and determine the correct stoichiometry of Rtt109:Vps75 in the full catalytic complex, I reconstituted the Rtt109–Vps75 complex *in vitro* from individual components expressed in *E. coli* (see section 3.1) and studied it in solution with NMR and SEC-MALS.

First, I attempted to reconstitute both 1:2 and 2:2 Rtt109–Vps75 complexes *in vitro* by mixing $Vps75_2$ with either one or two molar equivalent of Rtt109. While the 1:2 complex could be isolated by SEC, addition of more than one equivalent of Rtt109 to $Vps75_2$ resulted in either precipitation of excess Rtt109 or its elution in an unbound state, suggesting that only the 1:2 complex was formed under these conditions (20 mM Tris–HCl, 300 mM NaCl, 1mM DTT) (Fig. 4.1.1).

As X-ray structures were obtained from co-expression and co-purification of Rtt109 and Vps75, I also checked whether both 1:2 and 2:2 complexes (denoted as $Rtt109–Vps75_2$ and $Rtt109_2–Vps75_2$) could be assembled directly in *E. coli*. Rtt109 with a His_6-tag was co-expressed with untagged Vps75, to allow the enrichment of the purified material with either Rtt109 or any natively formed Rtt109–Vps75 complex. After co-purification by affinity chromatography (HisTrap), the sample eluted from the size-exclusion column as three species: high molecular weight aggregates, one peak containing both Rtt109 and Vps75 and one peak containing free Rtt109. The fractions containing the Rtt109–Vps75 complex were then loaded onto an analytical SEC column, where they eluted as a single peak with a retention time corresponding to the $Rtt109–Vps75_2$ complex reconstituted *in vitro* (Fig. 4.1.1A). This result indicates that also in *E.coli* Rtt109 binds Vps75 with a 1:2 stoichiometry.

Next, I monitored the formation of the Rtt109–Vps75 complex by NMR. In ^1H-^{13}C HMQC spectra of ILV-labelled $Vps75_2$ only one set of peaks for the methyl groups of chain A and B could be observed, meaning that Vps75 homodimer ($Vps75_2$) is symmetric (Fig. 4.1.1B). Addition of 1 equivalent of Rtt109 to $Vps75_2$ resulted in the splitting of some peaks into 2

sets of peaks: the peaks of the first set conserved the chemical shifts displayed in the absence of Rtt109, while the peaks of the second set adopted new chemical shifts. These chemical shift perturbations (CSPs) were assigned to the methyl groups of Vps75 located at the Rtt109 binding interface. Therefore, upon addition of one equivalent of Rtt109 to the Vps75 dimer, one Vps75 monomer, Vps75(A), remains equivalent to the free form, while the other one, Vps75(B)[2], binds Rtt109.
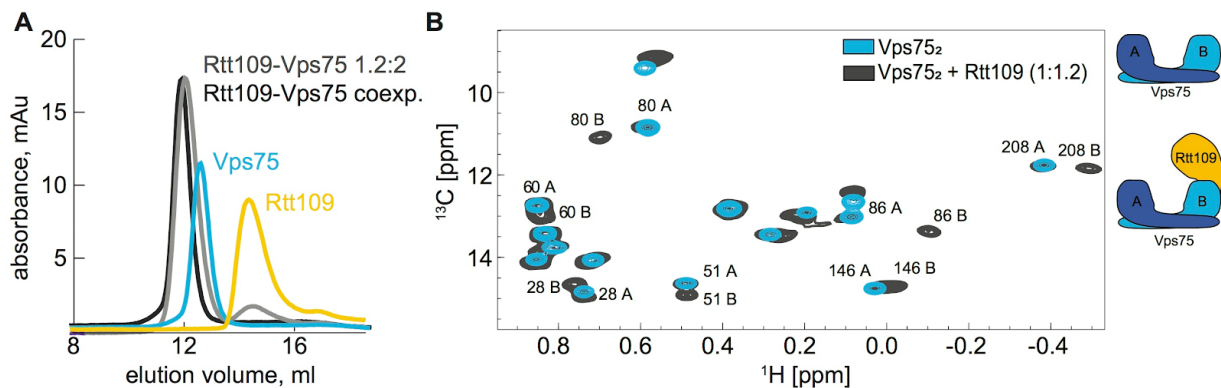
Figure 4.1.1. Vps75 forms a tight 2:1 complex with Rtt109. A. Size exclusion chromatography of Rtt109 (yellow), Vps75 (blue), Rtt109 mixed with Vps75 in a 1.2:2 molar ratio (light grey) and co-expressed Rtt109 and Vps75 (dark grey) in 20 mM Tris–HCl, 300 mM NaCl, 1mM DTT. In the 1.2:2 Rtt109:Vps75 mixture excess of Rtt109 elutes in the free form. The Rtt109–Vps75 complex obtained from co-expression of the proteins elutes at the same volume as the Rtt109–Vps75$_2$ complex. **B**. Overlay of ^{1}H-^{13}C HMQC spectra of 125 µM ILV methyl-labelled Vps75$_2$ in isolation (blue) and upon addition of 1 molar equivalent of Rtt109 (2:1 Vps75:Rtt109 molar ratio; grey). In isolated Vps75$_2$, methyl groups of Vps75(A) and Vps75(B) have identical chemical shifts. In the Rtt109–Vps75$_2$ complex, Vps75(A) remains unaffected, while Rtt109-bound Vps75(B) displays CSPs. Schematic representations of the models of Vps75$_2$ and Rtt109–Vps75$_2$ compatible with the NMR data are displayed on the right. The spectra were recorded in 25 mM Tris–HCl pH 7.5, 250 mM NaCl, 1 mM DTT at 298 K and 850 MHz.

Upon addition of the second equivalent of Rtt109, however, the NMR sample precipitated. The investigation of the 2:2 complex formation in these conditions was hindered by Rtt109 instability: Rtt109 is prone to degradation and aggregation due to the presence of the long unstructured α2β5 loop, which can be stabilised by the interaction with Vps75. The precipitation of the sample upon addition of the second equivalent of Rtt109 could thus have two reasons: (i) Vps75$_2$ accommodates only one Rtt109 copy and the excess of unbound Rtt109 precipitates at the concentration required for NMR experiments, (ii) added Rtt109

---

[2] Vps75(A) and Vps75(B), together forming the Vps75 dimer, are identical structurally and functionally. Rtt109 can bind on either side of the Vps75 dimer, interacting with the earmuff domain of one chain and the N-terminal part of the dimerization helix of the other chain. For convenience, the Vps75 monomer that is engaged by Rtt109 via its earmuff domain is always referred to as Vps75(B) or Rtt109-bound.

precipitates before it can form a stable complex with the Vps75$_2$–Rtt109 pre-formed complex.

To enable experiments with Rtt109 in isolation as well as the study of the "potentially-existing" 2:2 complex with Vps75, Dr. L. Lercher performed the thermofluor assays on Rtt109 (Kozak *et al*, 2016). The results revealed that Rtt109 has the highest stability at acidic pH (< 8.0) and at high concentration of NaCl (increase of the NaCl concentration up to 1 M correlated with the increase in the protein stability). However, we decided to exclude buffers with high ionic strength, as they are detrimental both for the interaction between Vps75 and histones (Bowman *et al*, 2011) and for the performance of the NMR experiments. Thus, we selected the following buffers: citrate (50 mM sodium citrate pH 5.5-pH 6.5, 150 mM NaCl, 5 mM BME), phosphate (25 mM phosphate pH 6.5, 200 mM NaCl, 5 mM BME) and HEPES (20 mM HEPES pH 7.5, 200-400 mM NaCl, 5 mM BME). Notably, the stability of Rtt109 in the citrate buffer depended less on the overall ionic strength than in the other buffers, allowing the reduction of the NaCl concentration to 150 mM nearly without any impact on the protein melting temperature.

NMR titration experiments with ILV-labelled Vps75 in the selected buffers demonstrated that in all cases addition of 1 molar equivalent of Rtt109 to Vps75$_2$ lead to the same CSPs as observed before (in Fig. 4.1.1B), indicating the formation of the Vps75$_2$–Rtt109 complex. However, the addition of the second equivalent of Rtt109 lead to different results depending on the conditions. In the phosphate buffer Rtt109 was not stable enough at the concentration of 100 μM required to reach 2 molar equivalents. Thus, in this buffer I monitored ILV-labelled Rtt109 at a concentration of 50 μM; Addition of 1 equivalent of Vps75 (2:2 Rtt109:Vps75 ratio) lead to two observable states of Rtt109, Vps75-bound and free (Fig. 4.1.2A). Upon titration of the second Vps75 equivalent, the peaks of Rtt109 corresponding to the free state of the protein disappeared, indicating the formation of the Rtt109–Vps75$_2$. In the HEPES buffer, addition of the second molar equivalent of Rtt109 to the Rtt109–Vps75$_2$ preformed complex yielded no changes in the spectrum of ILV-labelled Vps75, suggesting that only one Rtt109 can be bound by the Vps75 dimer in this condition (Fig. 4.1.2B). Thus, in HEPES and phosphate buffers, only the formation of the Rtt109–Vps75$_2$ complex was observed.

Differently, in citrate buffer addition of the second copy of Rtt109 caused broadening of the Vps75(A) peaks (which are unaltered in the Rtt109–Vps75$_2$ complex), while the peaks of the Vps75(B) (which binds Rtt109 in the Rtt109–Vps75$_2$ complex) increased in intensity (Fig. 4.1.2C). This suggested that in citrate buffer the Rtt109$_2$–Vps75$_2$ complex may exist in equilibrium with the Rtt109–Vps75$_2$, while complete conversion of Rtt109–Vps75$_2$ into the

Rtt109$_2$–Vps75$_2$ can be excluded based on the persistence of asymmetry in the spectrum of the Vps75(A)–Vps75(B) dimer.
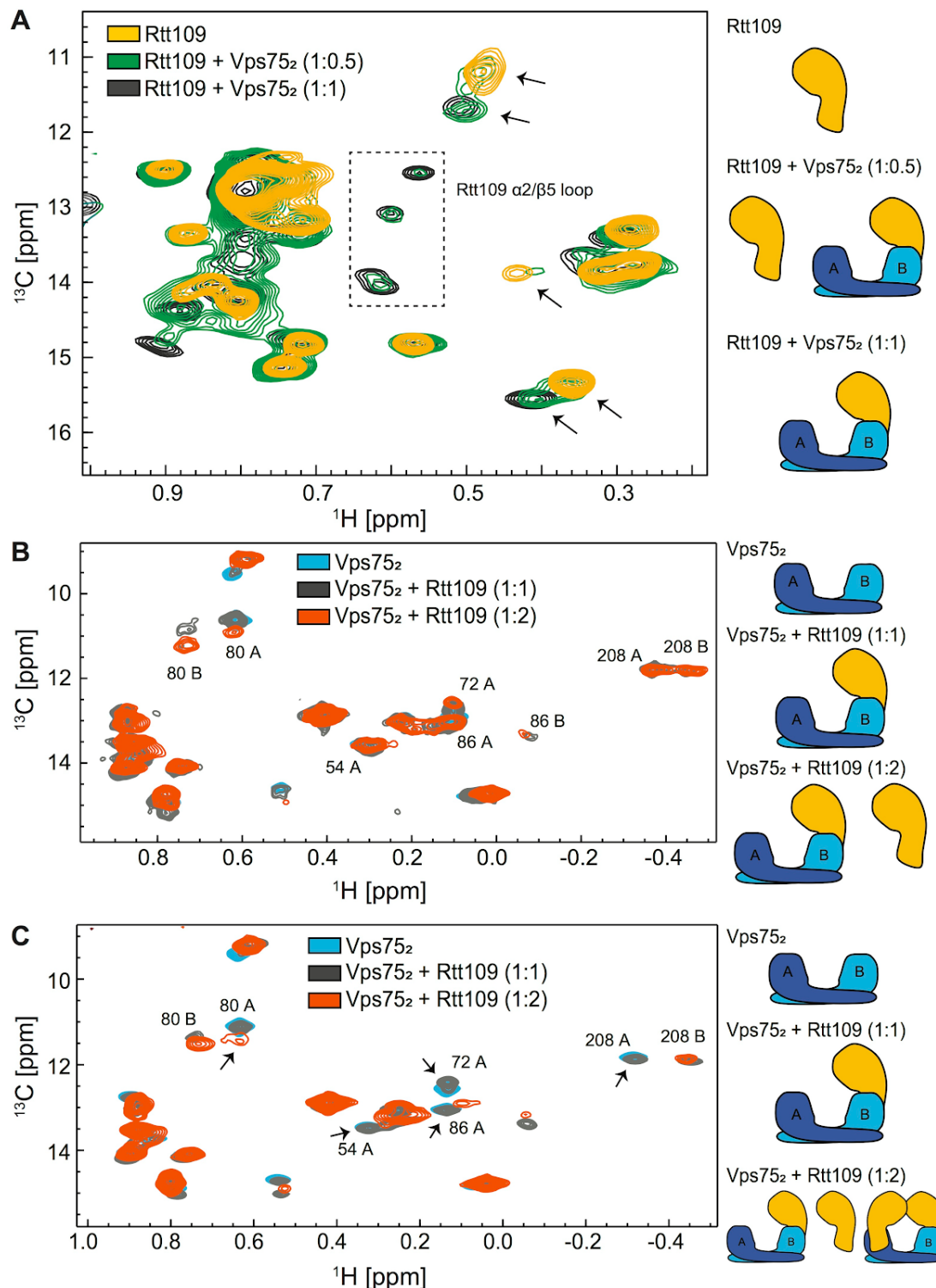
Figure 4.1.2. The formation of the Rtt109$_2$–Vps75$_2$ complex is observed only in the citrate buffer. **A**. Overlay of ^{1}H-^{13}C HMQC spectra of ILV methyl-labelled Rtt109 in isolation and upon addition of 1 or 2 molar equivalents of Vps75 in 25 mM phosphate pH 6.5, 200 mM NaCl, 5 mM BME at 50 µM Rtt109. Addition of 1 equivalent of Vps75 (or 0.5 equivalent of Vps75$_2$, green) leads to binding of only

50% of the available Rtt109 to Vps75. Upon addition of 2 Vps75 equivalents (black) Rtt109 is fully bound. Arrows indicate the peaks which are in slow exchange between the free Rtt109 and the Rtt109–Vps75$_2$ complex states. The dashed box shows the peaks located in the Rtt109 loop that becomes structured upon interaction with Vps75. **B-C**. Overlay of ^{1}H-^{13}C HMQC spectra of ILV methyl-labelled Vps75$_2$ in isolation and upon addition of 1 or 2 molar equivalents of Rtt109 in (**B**) 20 mM HEPES pH 8.0 (pH 7.5 for the 2:2 complex), 150 mM NaCl (300 mM NaCl for the 2:2 complex), 5 mM BME at 50–70 µM Vps75$_2$ (**C**) 50 mM sodium citrate pH 6.5 (pH 5.6 for the 2:2 complex), 150 mM NaCl, 5 mM BME at 50 µM Vps75$_2$. Addition of the second Rtt109 equivalent does not affect Vps75(A) in (**B**), while in (**C**) the Vps75(A) peaks indicated with an arrow broaden or disappear, suggesting that both Vps75 chains may bind to Rtt109. Spectra in (**A**) were recorded at 300K and 800 MHz. Spectra in (**B**-**C**) were recorded at 298 K and 850 MHz. Schematic representations of the complexes observed at each titration point are shown on the right (**A**-**C**). The data was produced jointly with Dr. L. Lercher.

To unambiguously determine the stoichiometry of the complex between Rtt109 and Vps75 in the citrate buffer, I performed SEC-MALS on the 2:2 complex (Fig. 4.1.3). SEC coupled with MALS allows the determination of the molecular weights of the eluted species independent on their shape and retention time. The Rtt109–Vps75 mixture, loaded in a 2:2 molar ratio, eluted as two peaks with the average molecular weights of 111,6 kDa and 54,4 kDa, matching the theoretical masses of 115,4 kDa and 50,1 kDa expected for the Rtt109–Vps75$_2$ complex and Rtt109 alone, respectively. The two peaks contained 72% and 20% of the overall injected mass, which is close to the expected distribution for the Rtt109–Vps75$_2$ complex and one free equivalent of Rtt109 (~70% and ~30%). The Rtt109$_2$–Vps75$_2$ complex could not be detected, indicating that this complex is too transient to be isolated by SEC.
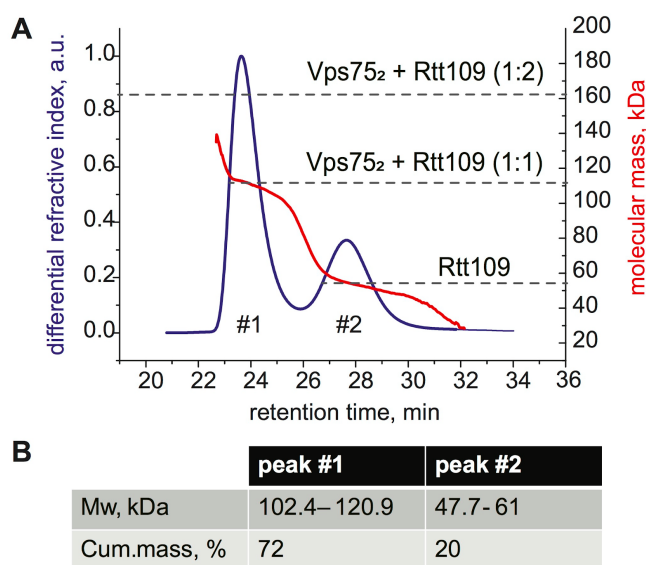
Figure 4.1.3. SEC-MALS analysis of the Rtt109–Vps75 2:2 mixture reveals the formation of the 1:2 complex exclusively. **A**. SEC-MALS analysis of the Rtt109–Vps75 2:2 mixture at 50 µM Rtt109 in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME. Red, molecular mass over the elution profile; the dashed grey lines indicate the theoretical masses of the corresponding complexes. **B**. Analysis of the peaks #1 and #2 of (**A**). Panel (**A**) is adapted from Danilenko *et al*, 2019 and originally produced by myself.

	peak #1	peak #2
Mw, kDa	102.4– 120.9	47.7- 61
Cum.mass, %	72	20

In summary, the data suggest that, although some amount of Rtt109$_2$–Vps75$_2$ complex can be observed in solution by NMR under certain conditions (citrate buffer, pH 5.6), this

complex is far less stable than the Rtt109–Vps75$_2$ complex.

4.2. Vps75$_2$ assembles an asymmetric complex with Rtt109 and Asf1–H3:H4.

After determining that Rtt109 binds Vps75 in a 1:2 ratio, I set out to determine with which stoichiometry Rtt109–Vps75$_2$ binds histones. To reveal how the substrate H3:H4 histones interact with the enzyme and Vps75 I used both NMR and SEC-MALS.

In vivo, the Rtt109–Vps75$_2$ sub-complex is thought to accept Asf1-bound histones as a substrate (Hammond *et al*, 2016). *In vitro*, the addition of the Asf1–H3:H4 complex to Rtt109–Vps75$_2$ leads to the formation of a stable complex composed of all five proteins, which could be isolated by SEC (Fig. 4.2.1). To verify the stoichiometry of the complex, the mixture of Rtt109, Vps75, Asf1 and H3:H4 in a 1:2:1:1 Rtt109:Vps75:Asf1:(H3:H4) ratio was loaded onto the SEC column coupled with MALS. The injected material formed a well-defined complex with a molecular mass of a 154 kDa; this value is similar to the molecular weight of 161 kDa calculated for the hexameric Rtt109–Vps75$_2$–Asf1–H3:H4 complex (later referred to as RVAH complex, by the first letters of the names of the proteins). The SDS-PAGE analysis of the peak confirmed that it contained all five proteins.
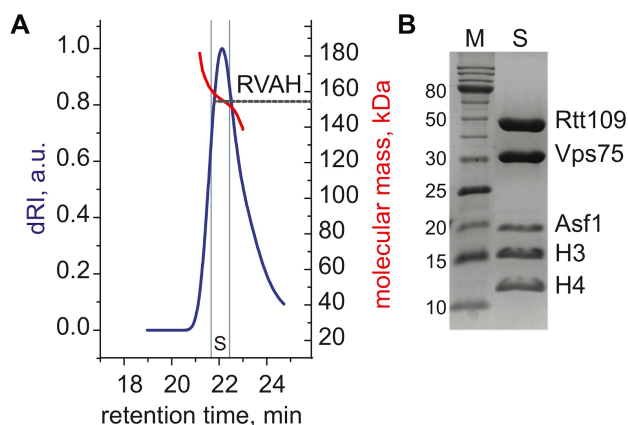
Figure 4.2.1. Rtt109 forms a stable complex with the histones and both chaperones. A. SEC-MALS analysis of the Rtt109:Vps75:Asf1:(H3:H4) mixture at 40 µM Rtt109 in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME. Red, molecular mass over the elution profile; the dashed grey line indicates the theoretical molecular mass of the RVAH complex. **B**. SDS-PAGE analysis of the fraction indicated in (**A**) as "S". Panel (**A**) is adapted from Danilenko *et al*, 2019 and originally produced by myself.

In the ^{1}H-^{13}C HMQC spectra of ILV-labelled Vps75, Vps75(B), displays the same chemical shifts indicative of Rtt109-binding as in the Rtt109–Vps75$_2$ sub-complex, while Vps75(A), which remained unbound in the Rtt109–Vps75$_2$ sub-complex, displays new chemical shifts, indicative of binding to the substrate Asf1–H3:H4 (Fig. 4.2.2). I concluded that the Rtt109:Vps75 1:2 stoichiometry is maintained upon addition of the Asf1–H3:H4; furthermore, additional ^{1}H-^{13}C spectra demonstrated that the formation of the RVAH complex is not

dependent on the order in which the protein are mixed and is not disrupted by the addition of Rtt109 in excess.
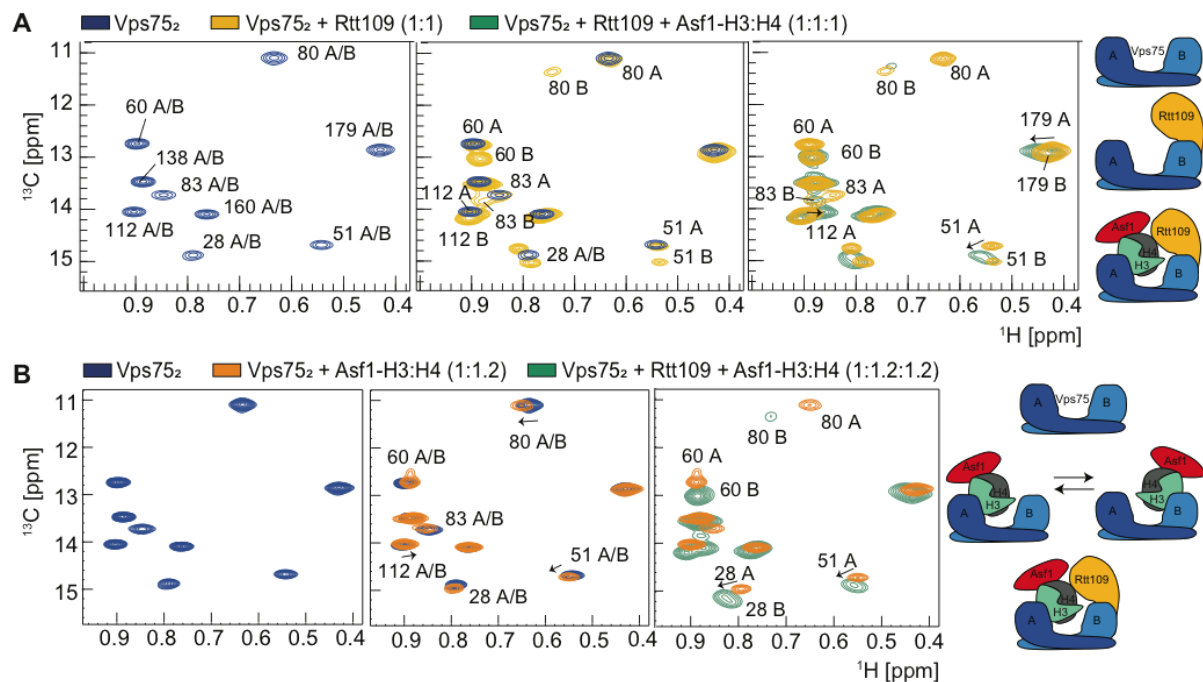
Figure 4.2.2. In the hexameric RVAH complex, Vps75(A) binds the substrate while Vps75(B) binds the enzymatic module. A. Overlay of ^{1}H-^{13}C HMQC spectra of ILV methyl-labelled Vps75 homodimer in isolation (blue); upon addition of Rtt109, Vps75(A) remains unaffected, while Vps75(B) is bound by Rtt109 (yellow) forming the Vps75$_2$–Rtt109 sub-complex; upon addition of Asf1–H3:H4 to the Rtt109–Vps75$_2$ sub-complex, the Vps75(B) peaks remain in their Rtt109-bound positions, while a subset of the Vps75(A) peaks either disappears or shows noticeable CSPs, indicative of binding to Asf1–H3:H4 (green). **B**. Overlay of ^{1}H-^{13}C HMQC spectra of the Vps75 homodimer in isolation (blue, assignments are the same as in (**A**)); addition of Asf1–H3:H4 results in the formation of a transient complex with Asf1–H3:H4, where the substrates exchange rapidly between Vps75(A) and Vps75(B) (orange); upon addition of Rtt109 to the Vps75$_2$–Asf1–H3:H4, a spectrum identical to the one in (**A**, right) is observed (green). All spectra were acquired with 60 μM Vps75$_2$ in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. Panel (**A**) is adapted from Danilenko *et al*, 2019 and originally produced by myself.

In order to better understand the interactions between the components of the complex and the function of both chaperones Asf1 and Vps75 in Rtt109 activation, I attempted to reconstitute several sub-complexes. In particular, I investigated whether Rtt109 can interact with histones in the presence of either Vps75 or Asf1.

Rtt109–Vps75$_2$–H3:H4 sub-complex. Previous studies of the interaction of histones with Vps75 and Rtt109 in the absence of Asf1 were plagued by aggregation(Bowman *et al*, 2014; Albaugh *et al*, 2010). Our attempts to study the interaction between Vps75$_2$ or Rtt109 or Rtt109–Vps75$_2$ and H3:H4 suffered from the same problem, with complexes without Rtt109 showing better solubility. To determine how histones interact with Vps75$_2$ I performed a

series of NMR titrations. Addition of one H3:H4 equivalent to the ILV methyl-labelled $Vps75_2$ yielded ^1H-^{13}C HMQC spectra with sufficiently good signal-to-noise ratio (Fig. 4.2.3A). However, increasing the H3:H4 concentration up to a 2:2 Vps75:(H3:H4) molar ratio lead to noticeable precipitation and poor quality of the spectra. Nevertheless, the observed CSPs could be mapped on the inner surface of the Vps75 dimer, particularly on the negatively charged surface of the Vps75 earmuff domain and dimerisation helix.

This CSP pattern was very similar to the one observed in the above-mentioned titration of $Vps75_2$ with Asf1–H3:H4, indicating that Vps75–(H3:H4) interaction is similar in the presence or absence of Asf1 (Fig. 4.2.3B).
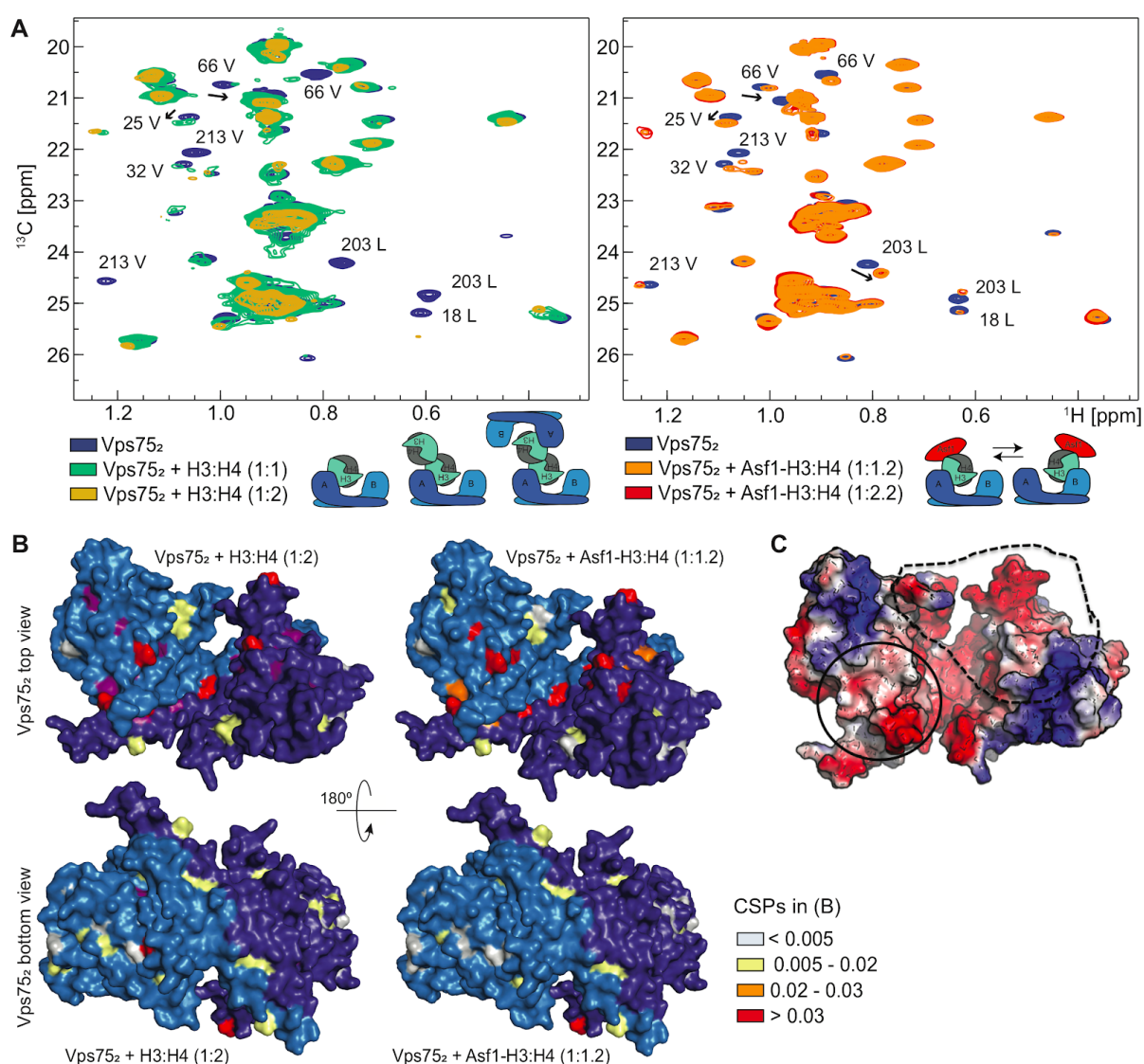
Figure 4.2.3. Vps75 interacts with H3:H4 via its inner surface and this interaction is compatible with Asf1 presence. A. Overlay of ^1H-^{13}C HMQC spectra of 50 µM ILV methyl-labelled $Vps75_2$ alone and upon addition of 1 or 2 molar equivalents of (left) H3:H4 in 10 mM TrisHCl pH 7.5, 100 mM NaCl, 1 mM DTT; (right) Asf1–H3:H4 in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 1 mM DTT. All spectra

were recorded at 298 K and 850 MHz. Cartoons below the spectra indicate the formed complexes (left, Vps75 dimer can interact with H3:H4 dimer, H3:H4 tetramer or two Vps75 dimers can bind one H3:H4 tetramer; right, Asf1 prevents H3:H4 tetramerization, Vps75 interacts with Asf1–H3:H4). **B**. CSPs observed in (**A**) mapped on the Vps75 structure (PDB entry 3q66). The disappearance of resonance is also indicative of a change in chemical environment and dynamics; as such, resonances that disappear upon addition of H3:H4 (**A, left**) are mapped in purple. The majority of CSPs/affected methyl groups cluster on the inner surface and cavity of $Vps75_2$, while the opposite side remains unaffected. **C**. Surface charge representation of $Vps75_2$ with positive, neutral and negative electrostatic potential indicated in blue, white and red, respectively. The Rtt109-binding surface of $Vps75_2$ is indicated by the dashed line. CSPs mapped in (**B**) correspond to the negatively charged part of the earmuff domain (encircled) and the bottom of the acidic cavity.

Vps75 and Asf1 do not interact in the absence of the histones, as indicated by the lack of CSPs upon addition of the unlabelled Asf1 to the ILV-labelled Vps75 and *vice versa* (Fig. 4.2.4, Lercher *et al.*, 2018). This supports the idea that the CSPs shown in Fig. 4.2.3B are a consequence of the Vps75-histone interaction.
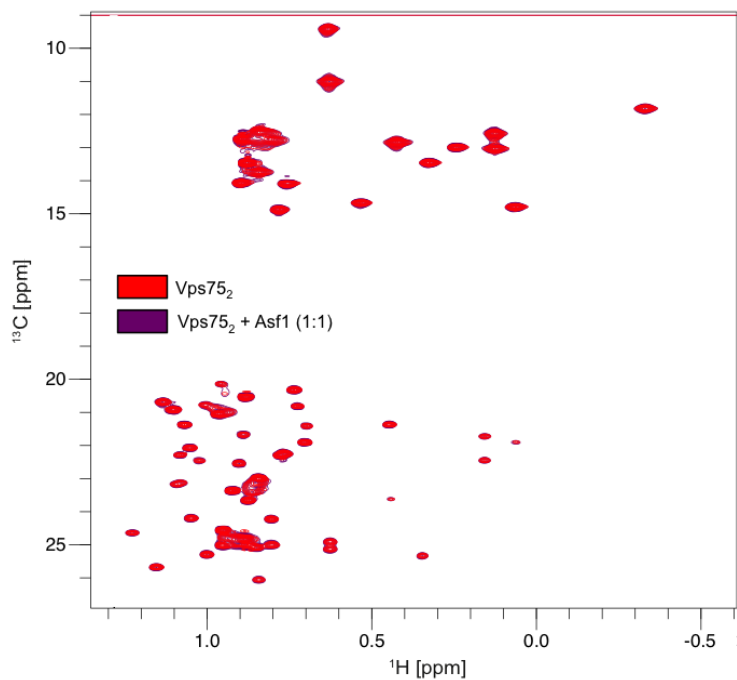
Figure 4.2.4. Vps75 does not interact with Asf1 in the absence of H3:H4. Overlay of ^{1}H-^{13}C HMQC spectra of 60 µM ILV methyl-labelled $Vps75_2$ alone and upon addition of 1 molar equivalent of Asf1. All spectra were recorded in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 1 mM DTT at 298 K and 850 MHz.

The spectra corresponding to the titration of $Vps75_2$ with either H3:H4 or the Asf1–H3:H4 complex show CSPs for the same residues; nevertheless the spectra are different. This can be explained by the inhomogeneity of the $Vps75_2$–H3:H4 complex, as $Vps75_2$ has been reported to bind both H3:H4 dimers and tetramers as well as to form Vps75 tetramers(Bowman *et al*, 2014; Hammond *et al*, 2016). The presence of mixed species (for example: $Vps75_4$, $Vps75_2$–H3:H4, $Vps75_2$–$(H3:H4)_2$, $Vps75_4$–$(H3:H4)_2$) would lead to aggregation and less clear CSP patterns. On the other hand, Asf1 is known to prevent the formation of the $(H3:H4)_2$ tetramer by interacting with the C-terminal region of H3(English *et*

al, 2006) and therefore promotes the formation of only one, better defined, $Vps75_2$–Asf1–H3:H4 complex.

Given the similarity of the CSPs of the Vps75 ILV-methyl groups in the $Vps75_2$–H3:H4 and $Vps75_2$–Asf1–H3:H4 complexes as well as the absence of a direct interaction between Vps75 and Asf1, it is plausible that the engagement of Asf1 with the H3:H4 tetramerization interface does not affect the Vps75–(H3:H4) interaction. In parallel to this work, the formation of the $Vps75_2$–Asf1–H3:H4 complex was reported by Hammond *et al* (Hammond *et al*, 2016), together with a low resolution model of the complex. Although the structural model was based on only five distance restraints, it suggested very few contacts between Asf1 and Vps75, which instead were modelled to interact with the opposite ends of H3:H4. The compatibility of the Asf1–H3:H4 and $Vps75_2$–H3:H4 interaction surfaces allows more links in the histone handover pathway during nucleosome assembly and further indicates that Asf1–H3:H4 is a likely substrate for Rtt109–$Vps75_2$ *in vivo*.

Rtt109–Asf1–H3:H4 sub-complex. While acetylation of H3 N-terminal lysines is carried out by the RVAH complex, the necessity of Vps75 for H3 K56 acetylation remained questionable. Investigation of the interaction of Asf1-bound histones with *S. cerevisiae* Rtt109 in the absence of $Vps75_2$ could shed a light on the role of each chaperone as well as on the interaction between Rtt109 and Asf1, which remained elusive. Recently, the Rtt109 from the pathogenic fungus *A. fumigatus* (*Af*Rtt109) was shown to form a functional complex with Asf1–H3:H4 *in vitro* (later referred to as *Af* RAH complex) (Zhang *et al*, 2018).

In order to characterise interaction of Rtt109 with Asf1–H3:H4, Dr. L. Lercher[3] tried to assemble the *S. cerevisiae* RAH complex by mixing Rtt109 and Asf1–H3:H4 in a 1:1 ratio in the citrate buffer. Both Rtt109 and Asf1–H3:H4 remained soluble; however, the complex was not stable enough to allow for isolation by SEC; a SEC run of the mixture resulted in Asf1–H3:H4 and Rtt109 eluting separately. I then attempted the reconstitution of the Rtt109–Asf1–H3:H4 complex with the protocol reported for the preparation of the *Af* RAH complex (Zhang *et al*, 2018), but this protocol lead to a complete precipitation of the *S. cerevisiae* (*Sc*) proteins. The difference in the biophysical properties of the RAH complex in *Af* and *Sc* can be explained by the differences in the secondary structure of *Sc*Rtt109[4] and *Af*Rtt109: *Af*Rtt109 lacks the α2β5 loop necessary for Vps75 binding and has a short loop inserted between the α6 and α7 helices, absent in *Sc*Rtt109 (Fig. 4.2.5). In addition, the

---

[3] Work on the Rtt109 interaction with Asf1–H3:H4 was done in the Carlomagno group by Dr. L. Lercher and myself.

[4] *S. cerevisiae* Rtt109 is referred to as "Rtt109" throughout the thesis. However, to avoid any confusion, in this paragraph and in the Fig. 4.2.5 is it named "*Sc*Rtt109".

surface charge of the two enzymes is different: *Sc*Rtt109 has positively charged areas, which are engaged in the interaction with Vps75$_2$. This surface charge distribution is not conserved in *Af*Rtt109. The difference in the isoelectric point of *Sc*Rtt109 and *Af*Rtt109 (9.41 and 6.05, respectively) might explain the solubility of *Sc*Rtt109 in citrate buffer, as opposed to the Tris-HCl buffer at pH 8.0 used for reconstitution of the *Af* RAH complex. The uncharacterised *Af* Vps75 homologue (*Af*Vps75) in turn does not interact with the *Af*Rtt109 *in vitro* and has an insertion in the region corresponding to the Vps75 α5 helix, suggesting a possible structural similarity to Nap1 instead. Altogether, these differences question the similarity of the H3 acetylation mechanisms in *A. fumigatus* and *S. cerevisiae*.
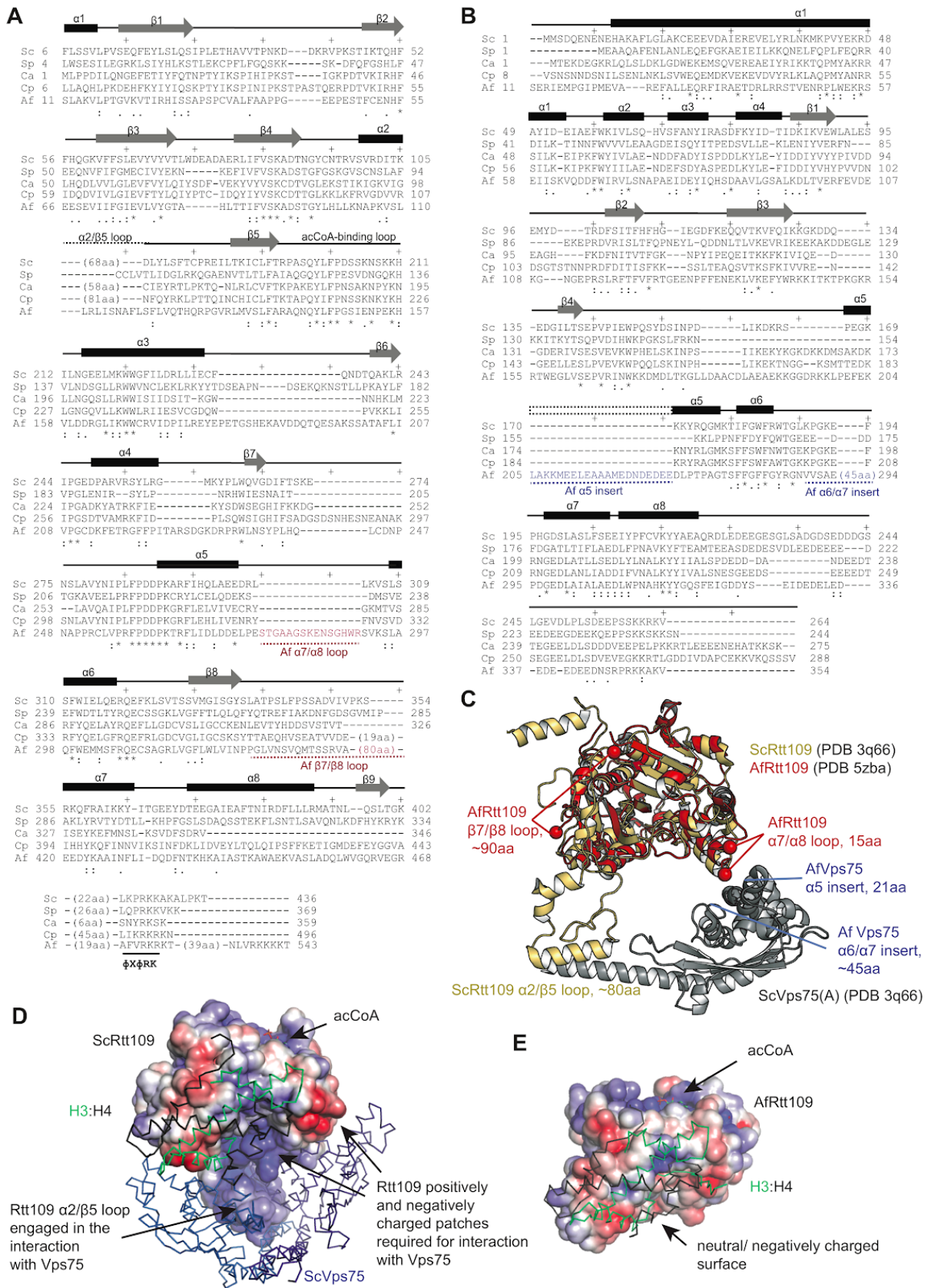
Figure 4.2.5. Comparison of Rtt109 from *S. cerevisiae* and *A. fumigatus*. A, B. Sequence alignments of Rtt109 (**A**) and Vps75 (**B**) in several fungal species (*Sc: S. cerevisiae; Af: A. fumigatus; Sp: Schizosaccharomyces pombe*; *Ca: Candida albicans*; *Cp: Candida parapsilosis*). Dots, colons and

asterisks indicate weakly similar, strongly similar and fully conserved residues, respectively. Secondary structure elements of *Sc* proteins are mapped above the alignments. The loops of *Af*Rtt109 are named according to the secondary structure of *Af*Rtt109, which differs from *Sc*Rtt109. **C.** Comparison of the structures of Rtt109 from *S. cerevisiae* and *A. fumigatus*. Red: *Af*Rtt109 structure (PDB entry 5zba); yellow: *Sc*Rtt109 structure in complex with Vps75$_2$ (PDB entry 3q66); grey: *Sc*Vps75 structure from 3q66 (only Vps75(A) is shown). The positions of amino acid insertions in the *Af*Vps75 sequence are indicated in blue on the Vps75 structure from PDB entry 3q66. The lack of the α2β5 loop in *Af*Rtt109, which in *Sc* is essential for interaction with Vps75, and the presence of insertions in both *Af*Rtt109 and *Af*Vps75 that map to the histone-binding surface of the *Sc* proteins, suggests that the acetylation of histones in *Af* may be regulated by a mechanism different to that in other fungi. **D, E.** Surface charge representations of *Sc*Rtt109 (**D**) and *Af*Rtt109 (**E**) with positive, neutral and negative electrostatic potential indicated in blue, white and red, respectively. The positively charged patch present in *Sc*Rtt109 is not conserved in *Af*Rtt109.

Due to the absence of a stable complex between Rtt109 and Asf1–H3:H4, we chose to study this by NMR, which is able to capture transient interactions. Comparison of ^{1}H-^{13}C HMQC spectra of ILV methyl-labelled Rtt109 alone and upon addition of Asf1–H3:H4 yielded small CSPs, which were confined to the Rtt109 unfolded region (Fig. 4.2.6A). A reverse NMR titration with ILV methyl-labelled H3:H4 in complex with Asf1 revealed no changes in the spectrum upon addition of Rtt109. The experiment did not reveal any interaction between Rtt109 and H3:H4, suggesting that the small CSPs seen in Fig 4.2.6A arise from an interaction between Rtt109 and Asf1. Interestingly, analogous experiments in the presence of Vps75$_2$ caused noticeable CSPs, indicative of interactions between the Rtt109 globular domain and the histones in the context of the RVAH complex (Fig. 4.2.6B, Fig. 4.5.3B).
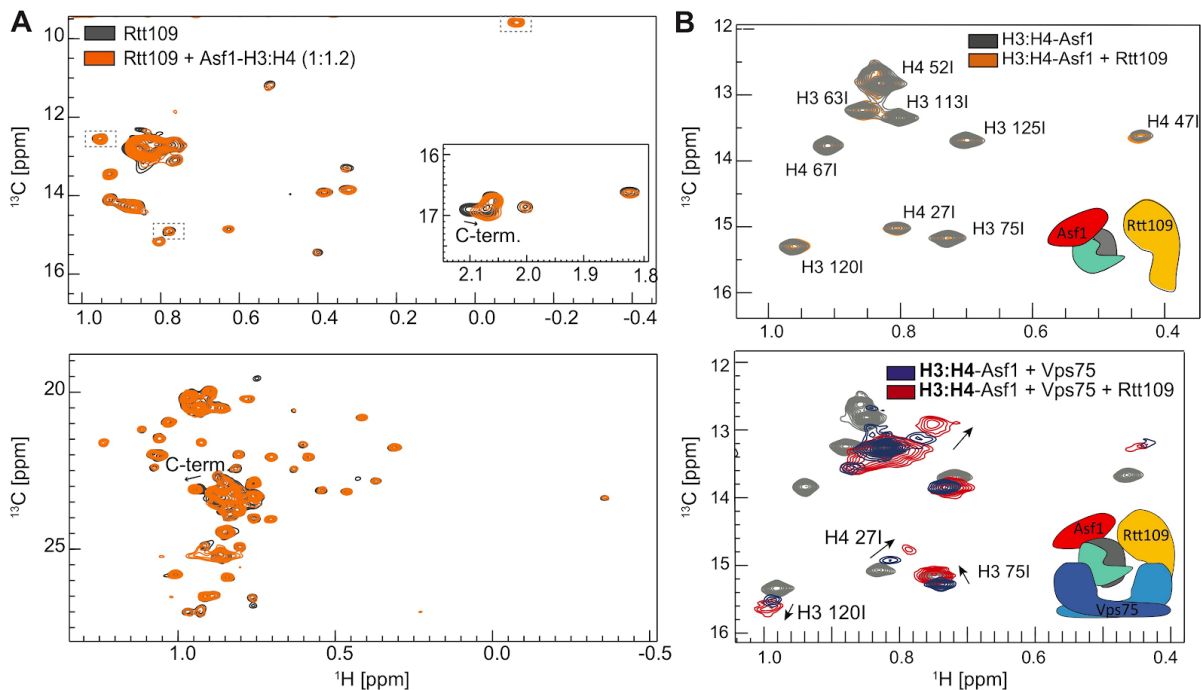
Fig. 4.2.6. Interaction of Rtt109 with Asf1–H3:H4 is altered/weaker in the absence of Vps75. A.

Overlay of ^{1}H-^{13}C HMQC spectra of 20 µM ILVM methyl-labelled Rtt109 in isolation and upon addition of Asf1–H3:H4. The IM and LV regions of the spectra are shown in the top and bottom panels, respectively. CSPs are observed for the Rtt109 C-terminus, while no CSPs are visible for the globular domain. The isoleucine residues indicated by the boxes with dashed borders belong to the Rtt109 globular domain and display CSPs upon addition of Asf1–H3:H4 to the ILV methyl-labelled Rtt109 in complex with Vps75$_2$ (Fig. 4.5.3). **B**. ^{1}H-^{13}C HMQC spectra of 50 µM ILV methyl-labelled H3:H4 in complex with unlabelled Asf1 upon addition of either Rtt109 or both Vps75$_2$ and Rtt109. No CSPs are observed upon addition of Rtt109 in the absence of Vps75$_2$, suggesting that Rtt109 interacts tightly with histones only in the RVAH complex.The spectra were recorded in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K.

Thus, Asf1 carrying the H3:H4 dimer can be directly bound by Vps75$_2$. The Vps75 homodimer can then act as an assembly platform for the full RVAH complex, binding and bringing together the enzyme and the substrate and facilitating their interaction. On the other hand, individual sub-complexes of Vps75$_2$–H3:H4 or Asf1–H3:H4 do not provide a viable pathway for the binding of the substrate to Rtt109 *in vitro*.

4.3. The Rtt109 C-terminal tail interacts with Asf1, releasing the H3 N-terminal tail for acetylation.

The NMR data detailed in section 4.2 hinted at the interaction between the unfolded region of Rtt109 and Asf1 (Fig 4.2.6). Fittingly, a previous *in vivo* study showed the importance of the unfolded Rtt109 -terminal region (residues 424-436) for Asf1-dependent acetylation of H3 (Radovani *et al*, 2013). Dr. L. Lercher hypothesized that the CSPs observed in Rtt109 titrations with Asf1–H3:H4 are the result of the interaction between the Rtt109 C-terminal tail and Asf1. Using NMR and Isothermal Titration Calorimetry (ITC) experiments, we demonstrated that the interaction between Rtt109 and Asf1 is confined to the short basic sequence at the C-terminus of Rtt109 (Rtt109$^{419–433}$) and solved the NMR structure of the Rtt109$^{419–433}$ in complex with Asf1 (deposited as PDB entry 6f0y). The investigation of the Rtt109$^{419-433}$ interaction with Asf1 as well as the structure determination were done by Dr. L. Lercher and are described in detail in (Lercher *et al*, 2018).

The Asf1-Rtt109$^{419–433}$ structure revealed that the Rtt109 peptide (419-433) interacts with the Asf1 surface on the opposite side of the histone-binding interface (Fig. 4.3.1), and that the Rtt109–Asf1 interaction is compatible with Asf1 binding H3:H4 (Lercher *et al*, 2018).
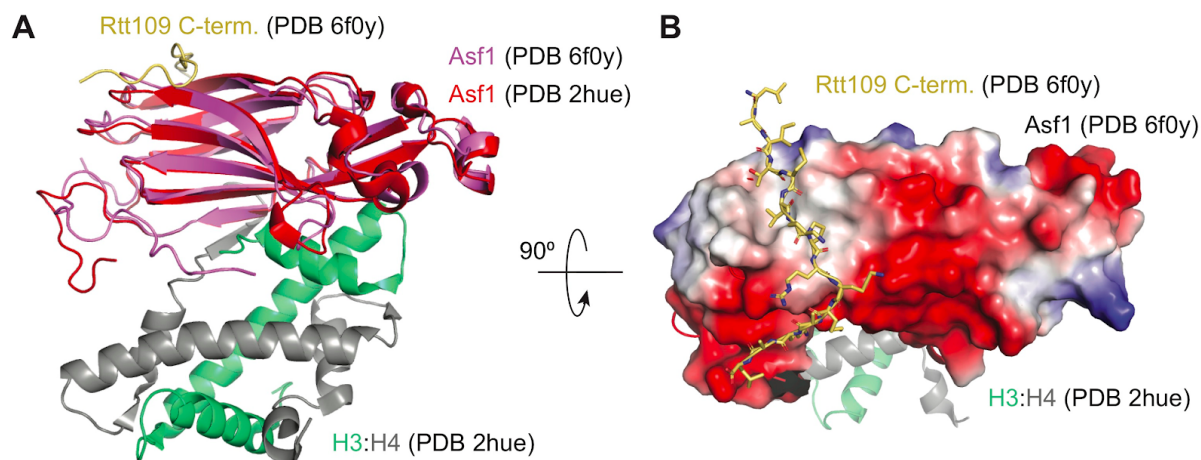
Fig. 4.3.1. Asf1 interaction with the Rtt109 C-terminus. A. Overlay of the Asf1-Rtt109$^{419–433}$ structure (PDB entry 6f0y) and Asf1–H3:H4 structure (PDB entry 2hue). Rtt109 C-terminus binds to Asf1 at the surface opposite to its histone-binding site and does not significantly alter Asf1 conformation. **B.** 90 degree rotation of the overlaid structures as in (A); Asf1 is shown in surface charge representation. Data used in the figure was produced by Dr. L. Lercher.

After the determination of the Asf1-Rtt109$^{419–433}$ structure, we addressed the question of whether this interaction is preserved in the RVAH complex and investigated its functional significance.

In order to characterise the Asf1–Rtt109 interaction in the context of the RVAH complex, we performed NMR titration experiments with ILV methyl-labelled Asf1. This allowed us to compare CSPs observed upon addition of Rtt109 to free Asf1 and to Asf1–H3:H4–Vps75$_2$. The Asf1 methyl groups affected by the addition of Rtt109 were identical in the two cases, indicating that the Rtt109$^{419–433}$–Asf1 interaction is preserved in the presence of histone and Vps75$_2$ (Lercher *et al*, 2018).

To obtain the Vps75$_2$–Asf1–H3:H4 sub-complex for this experiment, we first titrated unlabelled Vps75$_2$ to the ILV methyl-labelled Asf1 in complex with H3:H4. This yielded unexpected CSPs for the Asf1 residues mapping on the binding surface of the Rtt109 C-terminal tail, even prior to addition of Rtt109 (Fig. 4.3.2). As discussed previously, Vps75 and Asf1 do not interact directly (Fig. 4.2.4).
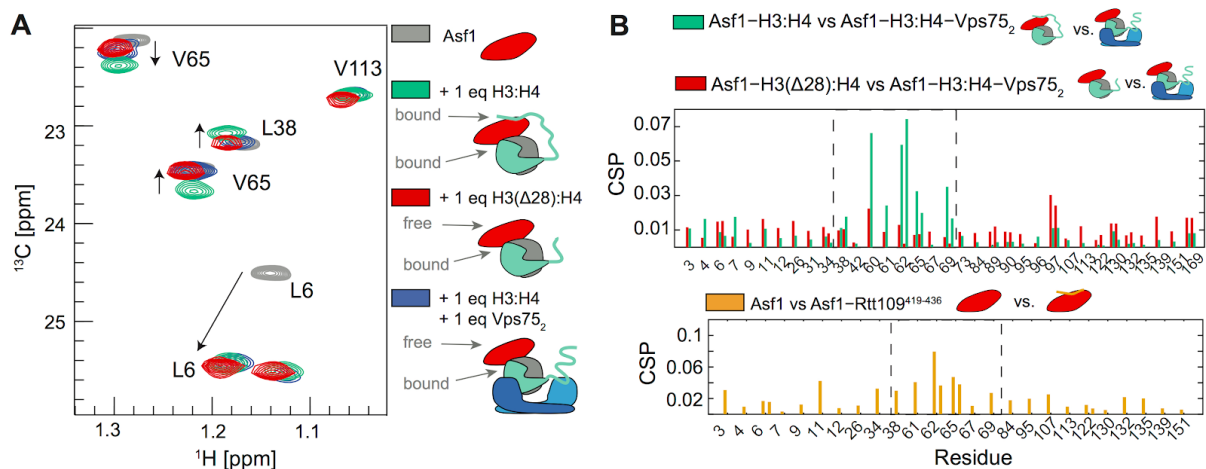
Figure 4.3.2. Asf1 binds H3 N-terminal tail via its Rtt109 C-terminus interaction interface. A. Overlay of ^{1}H-^{13}C HMQC spectra of 50 µM ILV methyl-labelled Asf1 alone (grey); upon titration of H3:H4 (green); upon titration of H3:H4 and Vps75$_2$ (green); upon titration of H3Δ28:H4 (red). Asf1 residues V65 and L38 are located at the Rtt109 C-terminus binding site, while residue L6 is located at the H3:H4 dimer interaction interface. While V65 and L38 are reverted to the state identical to the free Asf1 upon addition of Vps75$_2$ to Asf1–H3:H4, L6 remains H3:H4-bound. The cartoon schemes on the right represent the changes observed in Asf1 in each titration. "Bound" and "free" refer to the Asf1 surfaces that interact with Rtt109 C-terminus/H3 N-terminus (top part) and with the H3:H4 tetramerization interface (bottom part). **B.** CSPs plots for the spectra in (**A**) (top) and for the spectra of free Asf1 vs. Asf1 in complex with the Rtt109 C-terminus (bottom). The same residues of Asf1 that are involved in the interaction with Rtt109 C-terminus (indicated by the dashed box) are affected by the addition of Vps75$_2$ to Asf1–H3:H4. Figure is adapted from Lercher *et al*, 2018 and originally produced jointly by Dr. L. Lercher and myself; data was produced jointly by Dr. L. Lercher and myself.

The Asf1 peaks displaying CSPs after the addition of Vps75$_2$ to Asf1–H3:H4, shifted to the positions characteristic of free Asf1 (Fig 4.3.2A). At the same time, the peaks corresponding to the Asf1 residues involved in the interaction with the H3:H4 tetramerization interface remained in the histone-bound state. Altogether, this data suggests that Vps75 influences the interaction between Asf1 and histone regions other than the core, for example the long unstructured N-terminal tail of H3.

To test this hypothesis, we prepared a histone H3 protein carrying a 28-residue N-terminal deletion (H3Δ28) and reconstituted H3Δ28:H4 dimers. Spectra of Asf1-H3Δ28:H4 revealed that, although interaction between Asf1 and the H3:H4 tetramerization interface is retained, the Rtt109 binding surface of Asf1 had chemical shifts similar to those of free Asf1 or of Vps75$_2$–Asf1–H3:H4 (Fig. 4.3.2). This data suggested that in the Asf1–H3:H4 sub-complex residues 1-28 of the H3 N-terminal tail are bound by the same Asf1 surface that interacts with Rtt109$^{419-433}$, and that the addition of Vps75$_2$ to Asf1–H3:H4 results in the release of the H3 tail from Asf1.

To further test this idea, we monitored the H3 N-terminal tail by NMR. We reconstituted full-length H3:H4 dimers with ^{15}N labelled H3 and titrated Asf1, Asf1 and Vps75$_2$, or Asf1 and

Rtt109 onto H3:H4 (Fig. 4.3.3). Indeed, upon addition of Asf1 we observed CSPs for H3 residues 19-26, confirming the interaction between the H3 tail and Asf1. This interaction can be outcompeted either by binding of the Rtt109 C-terminal tail to the same Asf1 site, or by interaction of Vps75$_2$ with the H3 N-terminal tail.
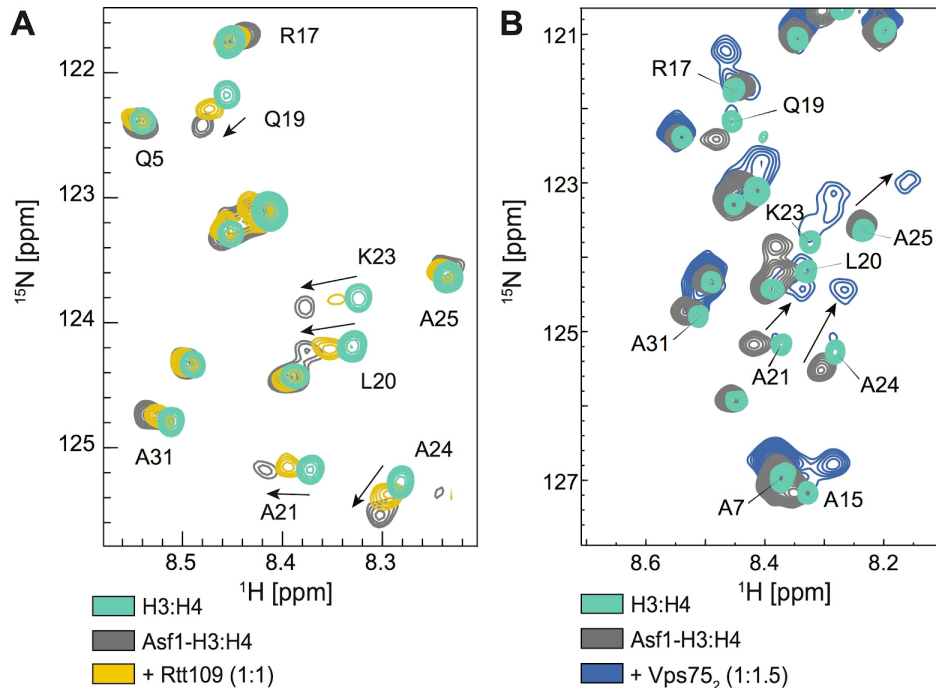
Figure 4.3.3. Rtt109 and Vps75$_2$ are able to prevent the interaction of the H3 N-terminal tail with Asf1. A. Overlay of ^{1}H-^{15}N HSQC spectra of 50 µM ^{15}N-labelled H3 in the H3:H4 complex (green), in the Asf1–H3:H4 complex (grey) and in the Asf1–H3:H4 complex upon addition of Rtt109 (yellow). In Asf1–H3:H4, peaks assigned to H3 19-26 display strong CSPs indicating binding of those residues by Asf1. Upon addition of Rtt109, peaks belonging to H3 19-26 are located between the Asf1-bound and free form, indicating that Rtt109 can outcompete the H3 tail for the interaction with Asf1. **B.** Overlay of ^{1}H-^{15}N HSQC spectra of 50 µM ^{15}N-labelled H3 in the H3:H4 complex (green), in the Asf1–H3:H4 complex (grey) and in the Asf1–H3:H4 complex upon addition of Vps75$_2$ (blue). Upon addition of Vps75$_2$, peaks belonging to H3 15-32 appear to be in neither in the Asf1-bound state, nor in the free state, indicating that Vps75$_2$ both releases the H3 tail from Asf1 and interacts with it. The figure is adapted from Lercher *et al*, 2018, Danilenko *et al*, 2019 and originally produced by myself; the data was produced jointly with Dr. L. Lercher.

Next, to test the functional relevance of the Asf1– Rtt109[419-433] interaction, I used *in vitro* dot blot assays to monitor the acetylation of K9 and K56 by Rtt109 and Rtt109[1-424], a truncated version of Rtt109 lacking the C-terminal residues crucial for the interaction with Asf1 (Fig. 4.3.4).
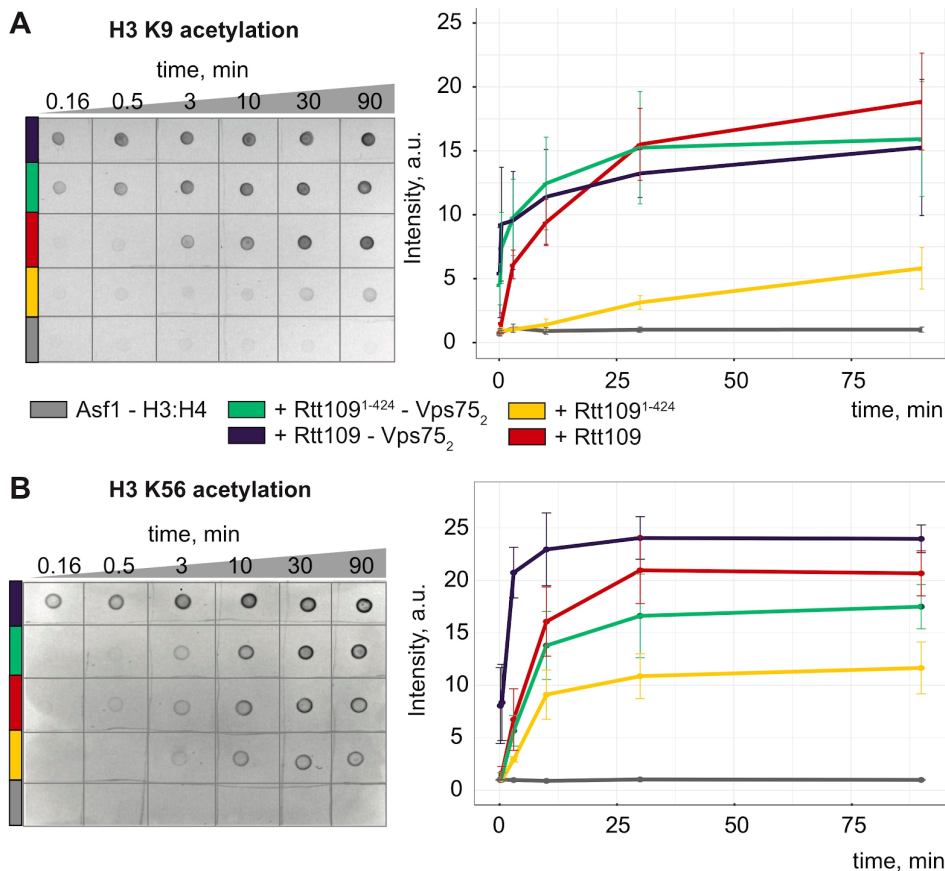
Figure 4.3.4. The C-terminal tail of Rtt109 is required for H3 K9 acetylation only in the absence of Vps75$_2$. A, B. Time-course of H3 acetylation in the Asf1–H3:H4 sub-complex by Rtt109–Vps75$_2$ (blue), Rtt109$^{1\text{-}424}$–Vps75$_2$ (green), Rtt109 (red), Rtt109$^{1\text{-}424}$ (yellow). Acetylation reactions were allowed to proceed for the indicated time, spotted on the membrane and stained with antibodies against acetylated H3 K9 (**A**) or H3 K56 (**B**). The dot blots (left panels) were quantified and the intensity of the dots, representing the amount of acetylated H3, were plotted (right panels). Reaction without Rtt109 (Asf1–H3:H4, grey) is a control for the antibody specificity towards acetylated H3. The experiments were repeated three times; the error bars represent the standard deviation. The figure is adapted from Lercher *et al*, 2018 and was originally produced by myself; the data was produced by myself.

Fig. 4.3.4A shows that in the absence of Vps75$_2$, H3 K9 acetylation is significantly reduced when the Rtt109 C-terminal residues are eliminated. However, we observed no difference in the acetylation activity of Rtt109 and Rtt109$^{1\text{-}424}$ in complexes containing Vps75$_2$. This data is in agreement with the results discussed in this section: the interaction between the enzyme and the Asf1–H3:H4 substrate relies on the Rtt109 C-terminal tail, which also releases the H3 N-terminal tail from Asf1, making it available for acetylation. Addition of Vps75$_2$ can functionally compensate for the role of Rtt109 C-terminal tail in K9 acetylation, because it can both free the H3 tail, containing K9, from Asf1 and facilitate the interaction between Rtt109 and the histones (Fig. 4.3.5). For an optimal level of K56 acetylation, however, the

presence of both Vps75$_2$ and Rtt109 C-terminus are required.
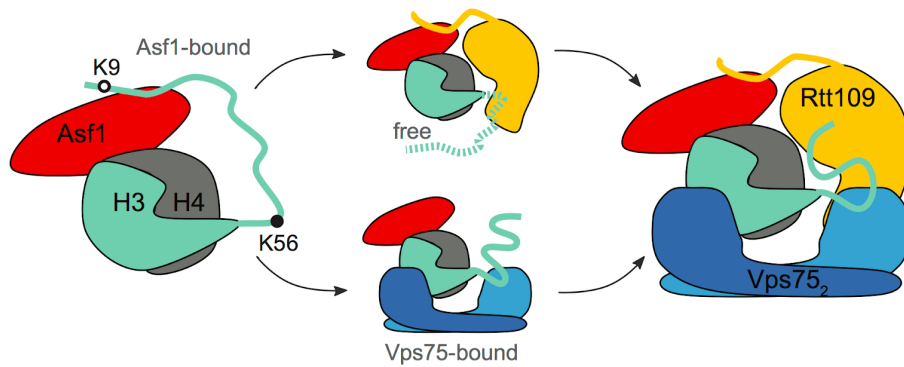
Figure 4.3.5. Scheme of the described interactions. H3 N-terminal tail is bound and protected from acetylation by Asf1 in Asf1–H3:H4, but is released by Rtt109 C-terminus via competition for the same Asf1 site or by Vps75$_2$ via interaction with Vps75$_2$. The figure is adapted from Lercher *et al*., 2018 and was originally produced by myself.

In summary, we found that Asf1 binds the H3 N-terminal tail, for example to protect it from nonspecific interactions or modifications. The Rtt109 C-terminal tail and Vps75$_2$ are required to overcome this Asf1-H3 interaction via different mechanisms. The Rtt109 C-terminal tail is solely responsible for the Rtt109 interaction with Asf1 (Lercher *et al*, 2018) and for the release of the H3 N-terminal tail from Asf1 in the absence of Vps75. *In vitro* Vps75$_2$ is able to fully replace the Rtt109 C-terminal tail function in K9 acetylation; however, *in vivo* both elements are required, suggesting the presence of additional factors or interactions impacting K9 acetylation (Radovani *et al*, 2013).

4.4. Determination of the Rtt109-Vps75$_2$-Asf1-H3:H4 complex structural model.

To gain further insights into the mechanism of Rtt109 activation by Asf1 and Vps75, I set out to solve the structure of the RVAH complex by an integrative structural biology approach, combining in-solution data from NMR and SANS.

Selection of the docking blocks: Vps75$_2$ adopts a new conformation in the RVAH complex to accommodate the enzyme and the histones.

To solve the structure of the RVAH complex I employed the M3 docking protocol described in section 3.7 (Fig. 4.4.7). M3 calculates complex structures from semi-flexible subunits (building blocks) using distance restraints between them. A low number of building blocks

simplifies the calculation and requires fewer distance restraints. The hexameric Rtt109–Vps75$_2$–Asf1–H3:H4 complex can be considered as a combination of two building blocks consisting of the X-ray structure of *S. cerevisiae* Asf1 in complex with the *Xenopus laevis* H3:H4 (PDB entry 2hue) and of the structure of *S. cerevisiae* Rtt109–Vps75$_2$ (PDB entry 3q66).

To be able to calculate the complex structure from the known structures of the building blocks, one needs to assume that the building blocks do not undergo significant conformational changes upon the formation of the full complex. To test this assumption in the case of the RVAH complex, I used both NMR and SANS to monitor the changes in the Asf1–H3:H4 and Rtt109–Vps75$_2$ structures upon formation of the RVAH complex.

As far as the Asf1–H3:H4 subunit is concerned (PDB entry 2hue), NMR data demonstrate that the interaction interface between Asf1 and H3:H4 is preserved in the RVAH complex (Fig. 4.4.1). The majority of the chemical shift differences observed between the free Asf1–H3:H4 and Asf1–H3:H4 in the context of the RVAH complex could be mapped to the Asf1 surface at the Rtt109[419-433] binding site. As discussed in section 4.3, this area is on the opposite side of the Asf1 H3:H4 binding site, thus the observed CSPs are likely a result of the Asf1 interaction with the Rtt109 C-terminus and release of the H3 N-terminal tail (absent in the PDB entry 2hue). As Asf1 residues engaged with the H3:H4 core did not display CSPs, I concluded that this interaction is preserved in the full RVAH complex. A further argument in favor of this conclusion is provided by the SANS data. The Asf1–H3:H4 structure (PDB entry 2hue) fits well to the SANS data reporting on the Asf1–H3:H4 shape in the context of the RVAH complex, further confirming that this structure can be used as a docking block (Fig. 4.4.1B).
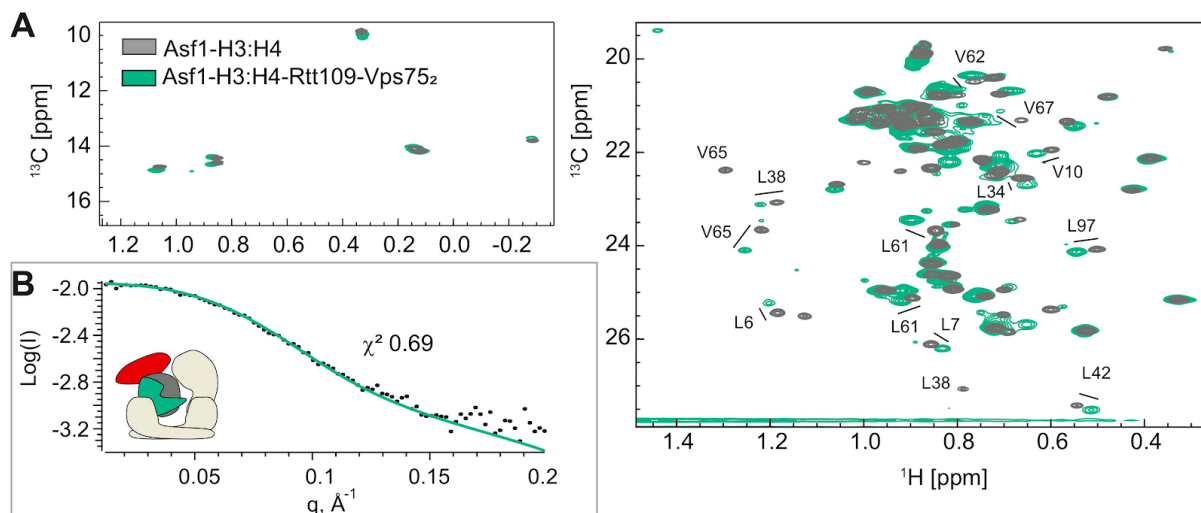
Figure 4.4.1. Asf1–H3:H4 structure (PDB entry 2hue) is preserved within the RVAH complex. A. Overlay of ^{1}H-^{13}C HMQC spectra of 50 μM ILV methyl-labeled Asf1 in complex with unlabelled H3:H4 or H3:H4–Rtt109–Vps75$_2$. Residues displaying chemical shift differences are marked. Spectra were recorded in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. **B**. Overlay of the experimental SANS curve (black, dotted) acquired for RVAH complex assembled with ^{1}H-Asf1–H3:H4 and ^{2}H(70%)-Rtt109–Vps75$_2$ in 100% D$_2$O buffer with the curve predicted from the structure of the RVAH complex (green) generated with PDB entry 2hue as the Asf1–H3:H4 building block. The schematic representation shows the proteins contributing to the scattering intensity (Asf1,H3:H4) in color, contrast-matched proteins (Rtt109, Vps75) in grey. Figure is adapted from Danilenko *et al*., 2019 and originally produced by myself.

On the other hand, the fit of the Rtt109–Vps75$_2$ structure (PDB entry 3q66) to the SANS data reporting on the Rtt109–Vps75$_2$ conformation in the RVAH complex was not satisfactory (Fig. 4.4.2B). Unlike the Asf1–H3:H4 structures, the known structures of Rtt109–Vps75$_2$ complexes and Vps75$_2$ dimers suggest a large conformational flexibility of the Vps75$_2$ dimer. As it can be seen from the superposition of these structures, the conformation of Vps75$_2$ ranges from "open" to "closed", with the degree of opening determined by the extent of the kink in the dimerisation helixes at residue P42 (Fig. 4.4.2A).
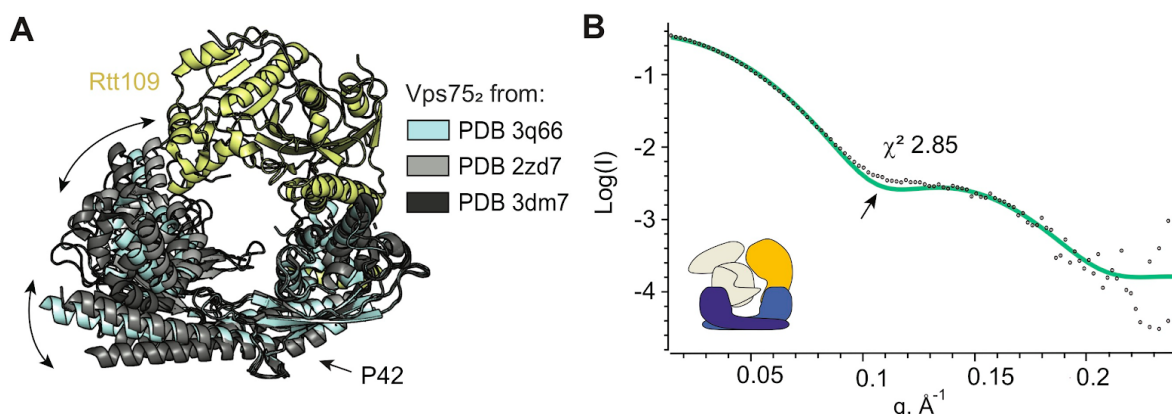
Figure. 4.4.2. The Rtt109–Vps75$_2$ conformation in the RVAH complex differs from the structure of PDB entry 3q66. A. The overlay of Rtt109–Vps75$_2$ and Vps75$_2$ structures aligned by Vps75(B) to PDB entry 3q66 shows that the Vps75$_2$ dimer can adopt more closed or more open conformations. **B.** Overlay of the experimental SANS curve (black, dotted) acquired for the RVAH complex assembled with ^{2}H-Rtt109–Vps75$_2$ and ^{1}H-Asf1–H3:H4 in 42% D$_2$O buffer with the curve predicted from the PDB entry 3q66 (green). The region of the mismatch is marked by an arrow. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

To determine the conformation adopted by Rtt109–Vps75$_2$ in the RVAH complex we used a combination of MD simulations (see section 3.8) and SANS. An unrestrained 400 ns MD simulation, starting from the Rtt109–Vps75$_2$ structure of PDB entry 3q66 was used by Dr. L. Lercher to explore the conformational space of the Rtt109–Vps75$_2$ complex. The 400 snapshot structures resulting from the MD simulation were scored against the SANS curves to identify the structures with the best fit to the experimental data. I recorded three SANS

datasets reporting on the conformation of Rtt109, $Vps75_2$ and Rtt109–$Vps75_2$ in the full complex (the ^{1}H-$Vps75_2^{1-225}$–Asf1–H3:H4 ^{2}H-Rtt109; ^{1}H-Rtt109–Asf1–H3:H4 ^{2}H-$Vps75_2^{1-225}$ and ^{1}H-Asf1–H3:H4 ^{2}H-Rtt109–$Vps75_2^{1-225}$ samples, respectively[5]) (Fig. 4.4.3A). The scoring against the first dataset yielded good fits and nearly identical χ^2 values between the experimental and back-calculated curves for all 400 structures. This indicated that (i) the shape of Rtt109 did not get significantly altered during the MD simulations, (ii) it adequately describes the Rtt109 conformation in the context of the RVAH complex.

The selection of the structure to use as building block was done by scoring against the curves reporting on the $Vps75_2$ and the Rtt109–$Vps75_2$ shapes (^{1}H-Rtt109–Asf1–H3:H4 ^{2}H-$Vps75_2^{1-225}$ and ^{1}H-Asf1–H3:H4 ^{2}H-Rtt109–$Vps75_2^{1-225}$ samples). The χ^2 values between the experimental SANS data and SANS data calculated from the MD structures were normalised from 0 to 1 for each of the two datasets. Subsequently, the fitness parameter, represented by the sum of the normalised χ^2 for each structure, was used to rank Rtt109–$Vps75_2$ structures according to their fitness to both SANS datasets.

Based on the fit to the experimental scattering curves, I selected an ensemble of 10 top-ranking Rtt109–$Vps75_2$ structures to represent the second building block in M3. Fig. 4.4.3D shows that all selected structures, despite belonging to different clusters, display a similar opening and rotation of Vps75(A) relative to Rtt109. This new $Vps75_2$ conformation results in a better accessibility of the Vps75 earmuff domain, which was previously identified as a potential binding site for Asf1–H3:H4 (section 4.2).

---
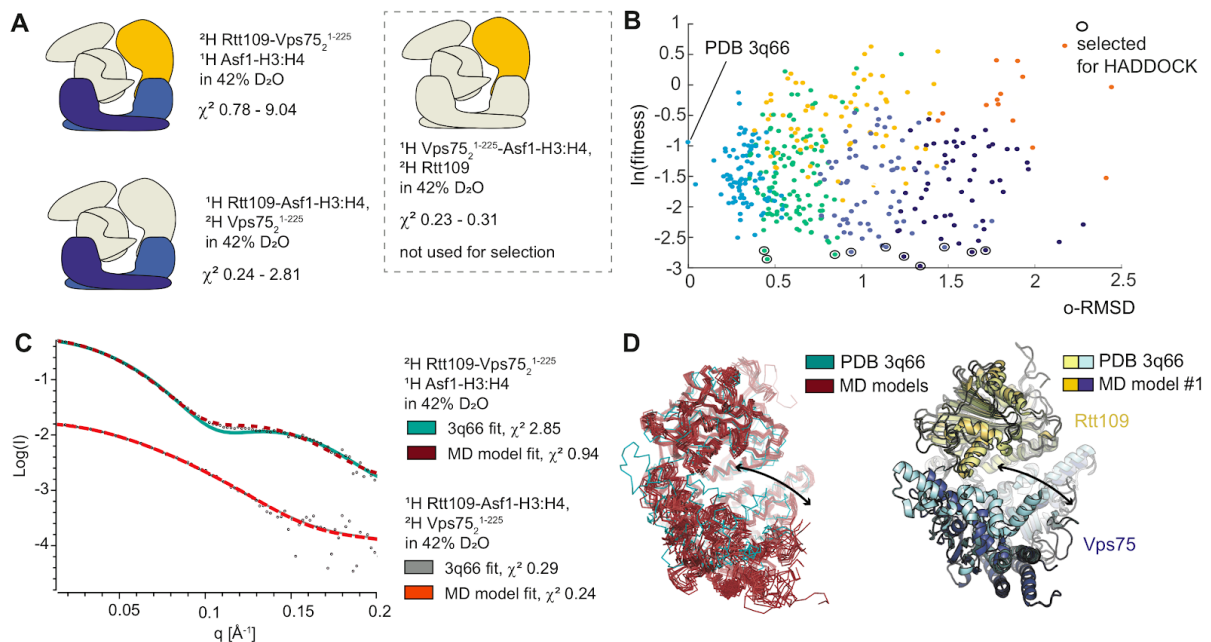
[5] $Vps75_2^{1-225}$ with the truncation of the 39 C-terminal residues was used for these experiments to match the number of Vps75 residues present in the PDB entry 3q66. The use of the truncated proteins for SANS samples is discussed below in the subsection "Distance restraints".

Figure 4.4.3. Selection of the Rtt109–Vps75$_2$ building block for M3. A. Schematic representation of SANS datasets used for the selection of the Rtt109–Vps75$_2$ conformation. The colored proteins contribute to the scattering-intensity, while the grey proteins are effectively invisible due to contrast-matching with the solvent. While the MD-generated structures can be discriminated by χ^2 values for the datasets on the left, they fitted equally well to the dataset reporting on the Rtt109 conformation in the RVAH complex. **B**. Plot of the normalized χ^2 versus structural similarity (orientational root-mean-square deviation, o-RMSD) to the PDB entry 3q66 for the MD-generated structures of the Rtt109– Vps75$_2$ complex. The o-RMSD was calculated for Vps75(A) relative to Rtt109. The normalized χ^2 was calculated for the two SANS curves used in the selection. The conformers are clustered by o-RMSD and the clusters are shown in different colors. Ten structures with lowest normalized χ^2, which contain representatives from the three best-scoring clusters, were selected as building blocks for M3. All selected structures have a better fitness to the experimental SANS data than the structure of PDB entry 3q66. **C**. Overlay of the experimental SANS curves used for the selection (black, dotted) with the curves predicted from the structure of PDB entry 3q66 (green and grey) and from the MD-generated structure with the best fitness to the SANS data (dark red, red). **D**. Overlay of the structure of PDB entry 3q66 and (**left**) 10 selected MD structures; (**right**) the selected MD structure with the best fitness. All selected models adopt the more open conformation, which results in a higher accessibility of the Vps75(A) earmuff domain for substrate binding. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself; the data was produced jointly with Dr. L. Lercher.

Distance restraints.

To gain sufficient information for the docking of the building blocks into the complex structure, I measured a number of PRE-derived distance restraints as well as a series of SANS curves with contrast matching.

Measurement of PRE restraints. The PRE experiments were conducted on the complexes reconstituted with the full length proteins as described in section 3.5. The distances between the building blocks were derived by placing paramagnetic tags on H3 or H4 and by monitoring the I/ILV methyl-groups of Rtt109/Vps75. The positions of the spin-labels as well as the overall number of extracted distances for each dataset are shown in Fig. 4.4.4.
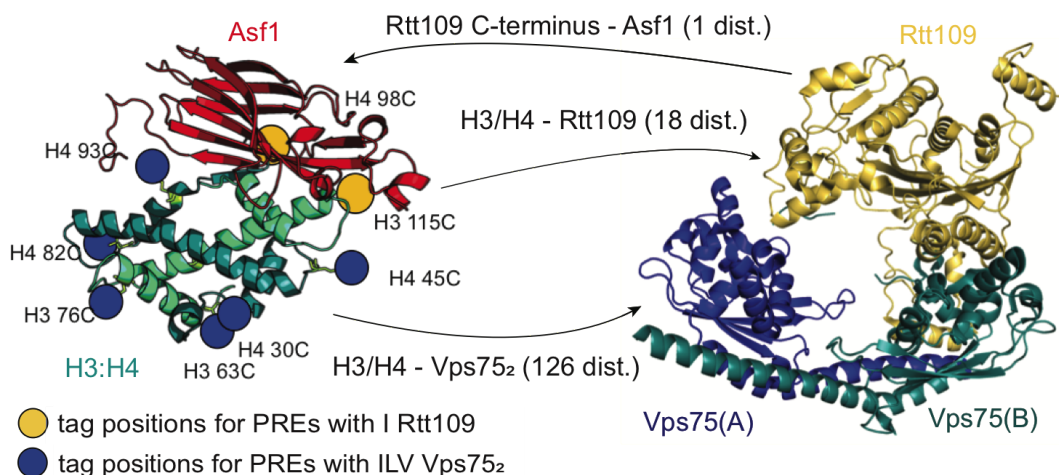
Figure 4.4.4. PRE-derived distance restraints. 8 PRE datasets were recorded with the paramagnetic tags attached to H3/H4 (positions are indicated by blue or yellow spheres); 2 of the

datasets reported on the distances between the tag and Rtt109 I methyl groups (yellow markers), the remaining datasets reported on the distances between the tag and Vps75$_2$ ILV methyl groups (blue). An additional restraint was added to loosely limit the distance between Rtt109 and Asf1 to guarantee the Rtt109[419-433]-Asf1 interaction.

Ambiguity in the Rtt109 methyl-group assignment (discussed in section 4.7) did not interfere with the generation of the PRE-derived restraints between Rtt109 and the histones. All PRE measurements from H3/H4 to the I methyl groups of Rtt109 returned so-called "negative" restraints, meaning that the paramagnetic tag was located too far from the methyl group to have an influence on the NMR signal. In this case a distance restraint greater than 24 Å was given between the methyl group and the Cβ of the tagged residue. Because all distance restraints were identical, they could be assigned to the Rtt109 I methyl groups irrespective of the assignment. This yielded 18 restraints between H3/H4 and Rtt109.

The generation of distance restraints from the experiments with the ILV methyl-labelled Vps75$_2$ was complicated by the presence of two chains. Initially, I attempted to generate restraints using the "or" statement, where distances were assigned to the methyl group of Vps75(A) or Vps75(B). Restraints generated by this approach proved to be too ambiguous, failing to produce a convergent structural ensemble. To solve this problem, I used only well-resolved peaks, which could be unambiguously assigned to either Vps75(A) or Vps75(B) methyl groups. In addition, when the distance between the corresponding methyl groups of chain A and B was two times larger than the radius of the sphere influenced by the paramagnetic tag, I generated additional distance restraints assuming that 0.5 of the methyl-group peak intensity was influenced by the presence of the paramagnetic tag and assigning this to the methyl-group of the chain A or B. This approach allowed to generate 126 distance restraints between H3/H4 and Vps75$_2$.

SANS-derived restraints. Vps75, H3, H4 and Rtt109 possess disordered tails: residues 225-264 in Vps75 (later referred to as the Vps75 C-terminal acidic domain, CTAD), 1-34 in H3, 418-436 in Rtt109, 1-19 in H4. These disordered sequences are absent in the X-ray structures serving as the building blocks. Presence of these residues in the SANS samples could compromise fitting of the RVAH complex structures, generated from the building blocks, to the SANS data. The Rtt109 and H4 parts missing in the building blocks comprise of only 18 residues each, while the longer Vps75 CTAD and H3 N-terminal tail (1-34) could affect the fitting to the SANS data more significantly.

To determine whether the RVAH complex could be assembled without these disordered tails, I used Vps75 and H3 constructs with truncated tails (Vps75$^{1–225}$ and H3$^{35–135}$) for reconstitution of the RVAH complex. I then monitored the complex assembly using ^{1}H-^{13}C

HMQC spectra of ILV methyl-labelled $Vps75_2$ or $Vps75_2^{1-225}$ (Fig. 4.4.5). The CSPs for full-length and truncated $Vps75_2$ upon subsequent addition of Asf1–H3/($H3^{35-135}$):H4 and Rtt109 followed the same pattern, indicating that the truncations did not affect the interactions between Vps75 and the remaining complex components. Moreover, I verified by SANS that the Asf1-Rtt109 conformation is preserved in the RVAH complex assembled with the $Vps75^{1-225}$ and $H3^{35-135}$. SANS samples of RVAH complexes reconstituted with ^2H Asf1–Rtt109 and either ^1H $Vps75_2^{1-225}$–Asf1–$H3^{35-135}$:H4 or ^1H $Vps75_2$–Asf1–H3:H4 resulted in the identical curves. In conclusion, truncation of the Vps75 CTAD and H3 N-terminus did not interfere with the assembly of the complex. Furthermore, the RVAH complex assembled with the truncated proteins could be isolated by SEC and the interactions between the complex components were preserved.
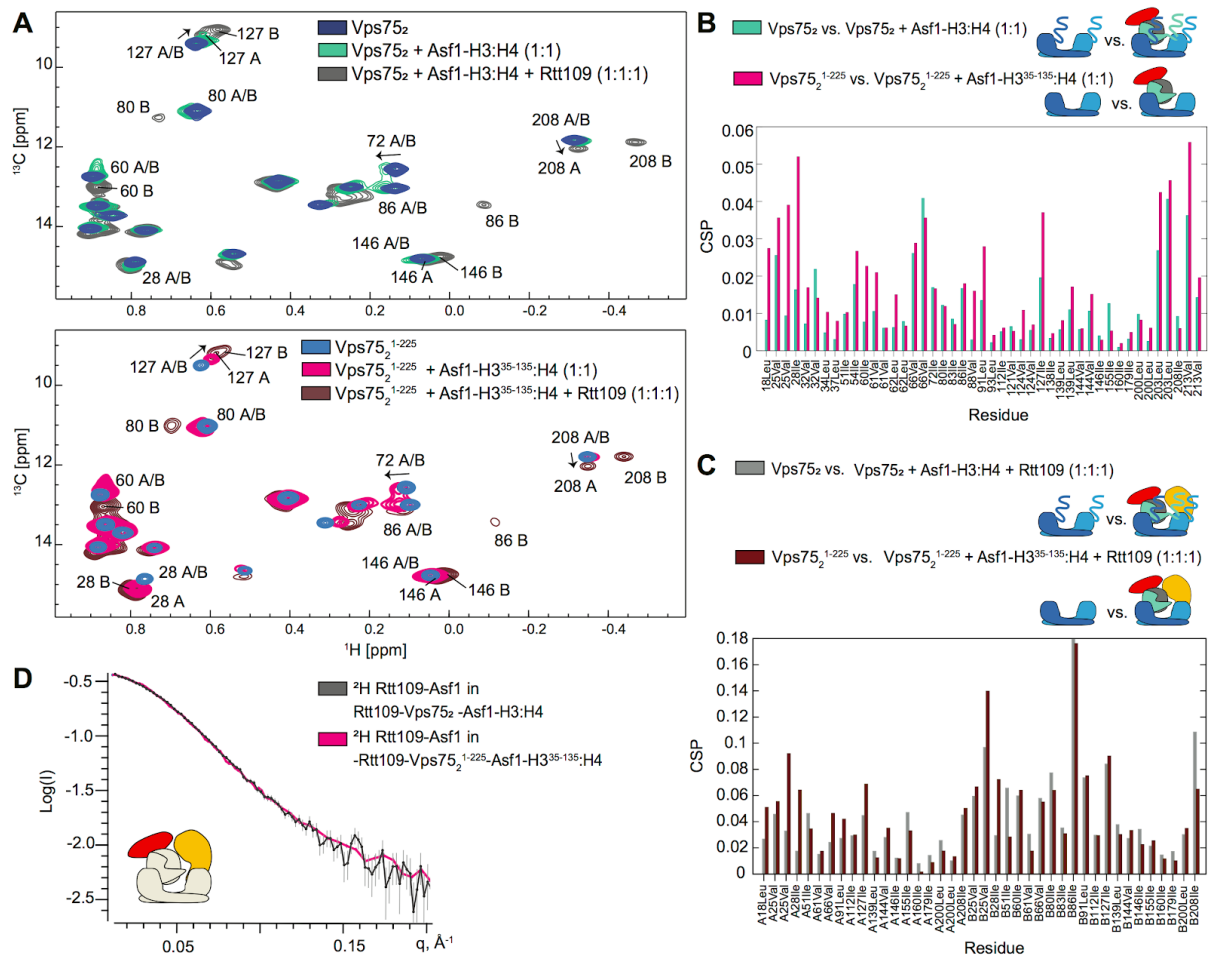
Figure 4.4.5. The assembly of the RVAH complex does not depend on the Vps75 CTAD and H3 N-terminal tail (residues 1-34). A. Overlay of ^1H-^{13}C HMQC spectra of 60 µM $Vps75_2$ in isolation and upon subsequent addition of Asf1–H3:H4 and Rtt109 (**top**); overlay of ^1H-^{13}C HMQC spectra of 60 µM $Vps75_2^{1-225}$ in isolation and upon subsequent addition of Asf1–$H3^{35-135}$:H4 and Rtt109 (**bottom**). The CSPs of the Vps75 follow the same pattern in both Asf1–H3:H4–Rtt109–$Vps75_2$ and Asf1–$H3^{35-135}$:H4–Rtt109–$Vps75_2^{1-225}$. **B-C.** CSP analysis of the spectra in (**A**). Spectra were recorded in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. Titration with $Vps75_2$ was done in a D$_2$O-based buffer, while the one for $Vps75_2^{1-225}$ was done in a H$_2$O-based buffer.

The differences in the CSPs values between the spectra are caused by the isotopic effects induced by the different solvents. **D**. Overlay of the experimental SANS curves acquired for the RVAH complex assembled with ^{2}H-Rtt109–Asf1 and ^{1}H-H3:H4–Vps75$_2$ in 42% D$_2$O (grey); ^{2}H-Rtt109–Asf1 and ^{1}H-H3$^{35-135}$:H4–Vps75$_2^{1-225}$ in 42% D$_2$O. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

Six SANS datasets on the RVAH complexes reconstituted with Vps75$^{1–225}$ and H3$^{35–135}$ were recorded to be used in the M3 protocol (Fig. 4.4.6). The labelling schemes were chosen in a way that allowed to obtain information on the shapes of the two docking blocks or their parts in relation to each other.
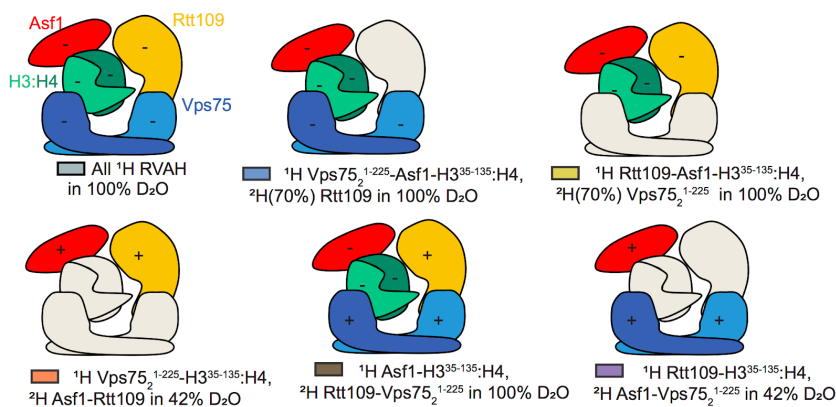
Figure 4.4.6. SANS restraint table. Schematic representation of the SANS samples measured for the M3 protocol. The colored proteins contribute to the scattering-intensity with either positive (+) or negative (−) contrast, while the grey proteins are effectively invisible due to contrast-matching with the solvent.

Structure calculation.

The scheme of the M3 protocol is shown in the Fig. 4.4.7. The PRE-derived distance restraints were used to drive the generation of the RVAH complex models.
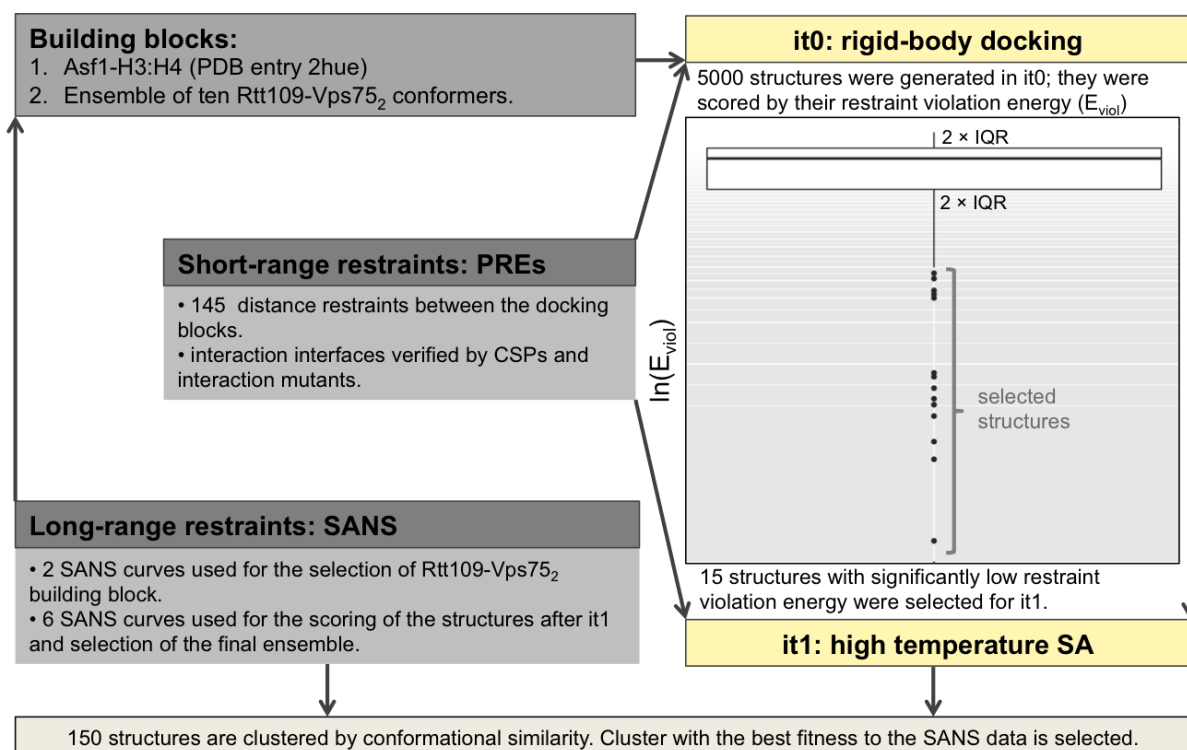
Figure 4.4.7. Scheme of the M3 calculation protocol. During the it0 stage, the building blocks are treated as rigid bodies and their docking is guided by the PRE-derived distance restraints. After it0, 15 structures with significantly better distance restraint violation energy were selected for the next stage (it1). During the high temperature simulated annealing (SA) in it1, the extended conformational search was driven again by the PRE-derived distance restraints; the side-chains of residues at the interaction interfaces were flexible. After it1 the structures were scored against the six SANS datasets and the cluster with the best fitness was selected. The protocol is described in detail in section 3.7.

The 150 structures resulting from the flexible refinement round were scored against the above-mentioned SANS datasets and clustered by their structural similarity (o-RMSD) to the structure with the best fitness to the SANS data. The best scoring cluster (green in Fig. 4.4.8) contained 33 structures with significantly better fitness to the SANS curves than the other structures; these structures were therefore selected as the final ensemble. The structure closest to the ensemble center (1.4 Å average RMSD to the other structures in the ensemble) was selected as the ensemble representative. This RVAH complex structure was refined in explicit water as described in section 3.8 and deposited as PDB entry 6o22.
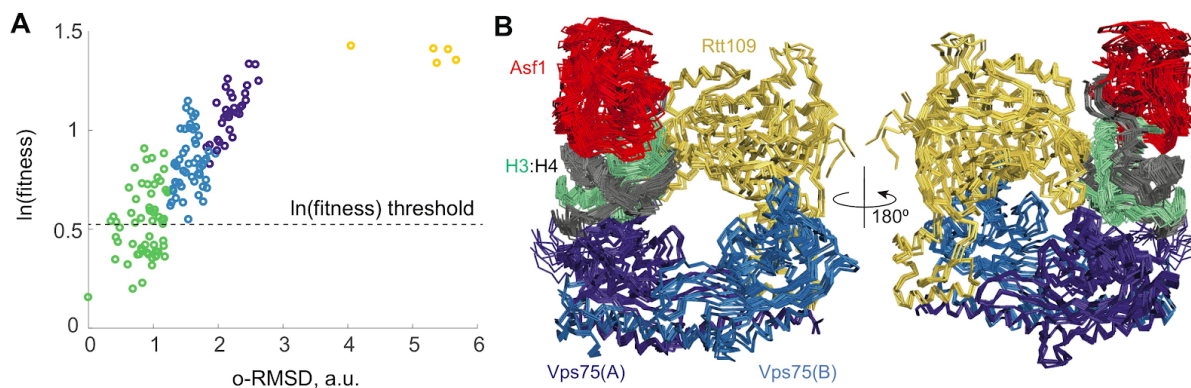
Figure 4.4.8. Cluster selection. A. Correlation between the o-RMSD of the RVAH structural models calculated by the M3 protocol and their agreement to the experimental SANS data (fitness). The o-RMSD was calculated with respect to the structure with the best fitness; the structures were clustered by o-RMSD and each cluster is represented by one color. Members of the cluster that best fit the experimental data (green cluster) with the fitness value below the indicated threshold were selected as the final ensemble (33 structures). **B.** Superposition of the 33 structures of the final ensemble. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

The deposited structure (PDB entry 6o22) fitted well to the experimental data, including both PRE measurements and SANS curves (Fig. 4.4.9, Fig. 4.4.10).
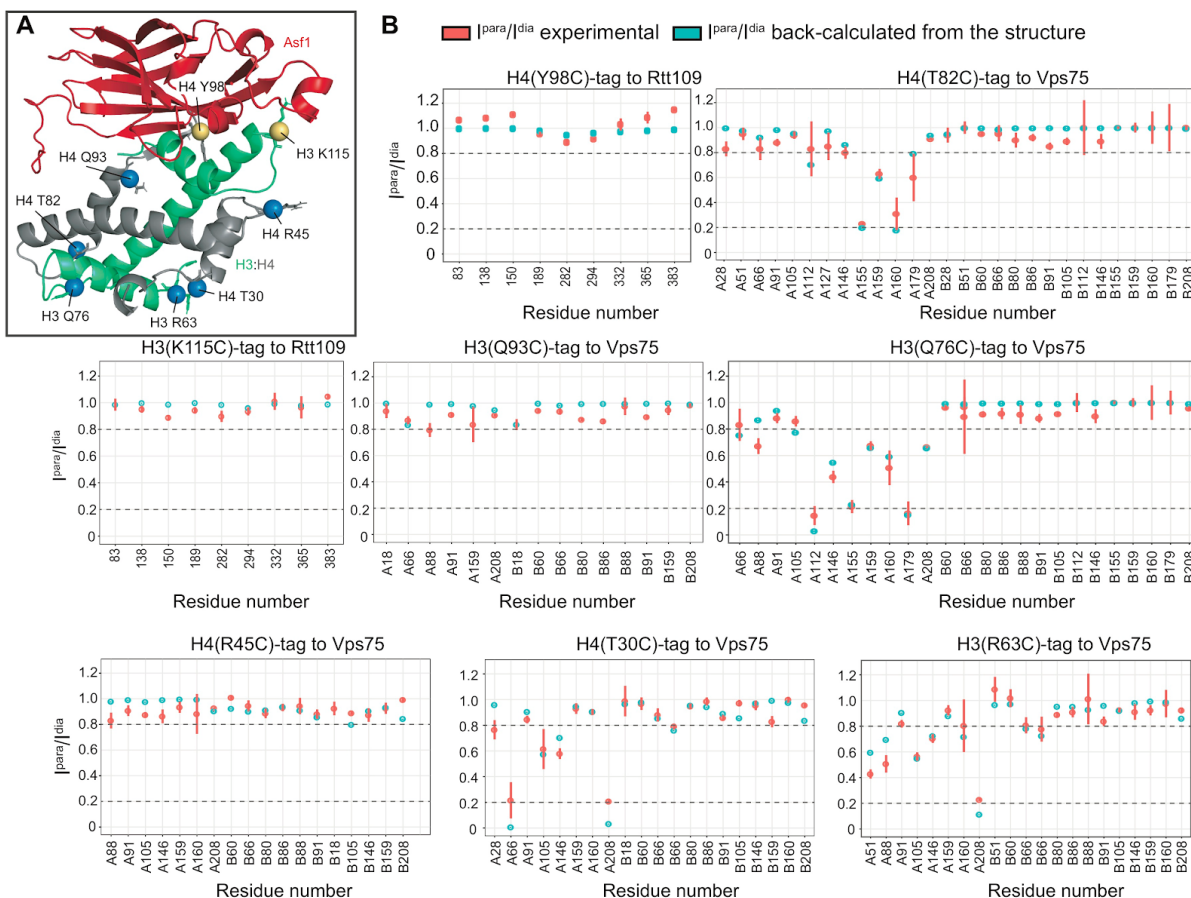
Figure 4.4.9. Fit of the structural model to the PRE-derived restraints. A. The positions of the paramagnetic tags used in the PRE measurements (color coding as in Fig. 4.4.4). **B.** Comparison of experimental (red) and back-calculated (light blue) ratios of peak intensities measured in the presence of the paramagnetic (I^{para}) or diamagnetic (I^{dia}) forms of the tags.. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.
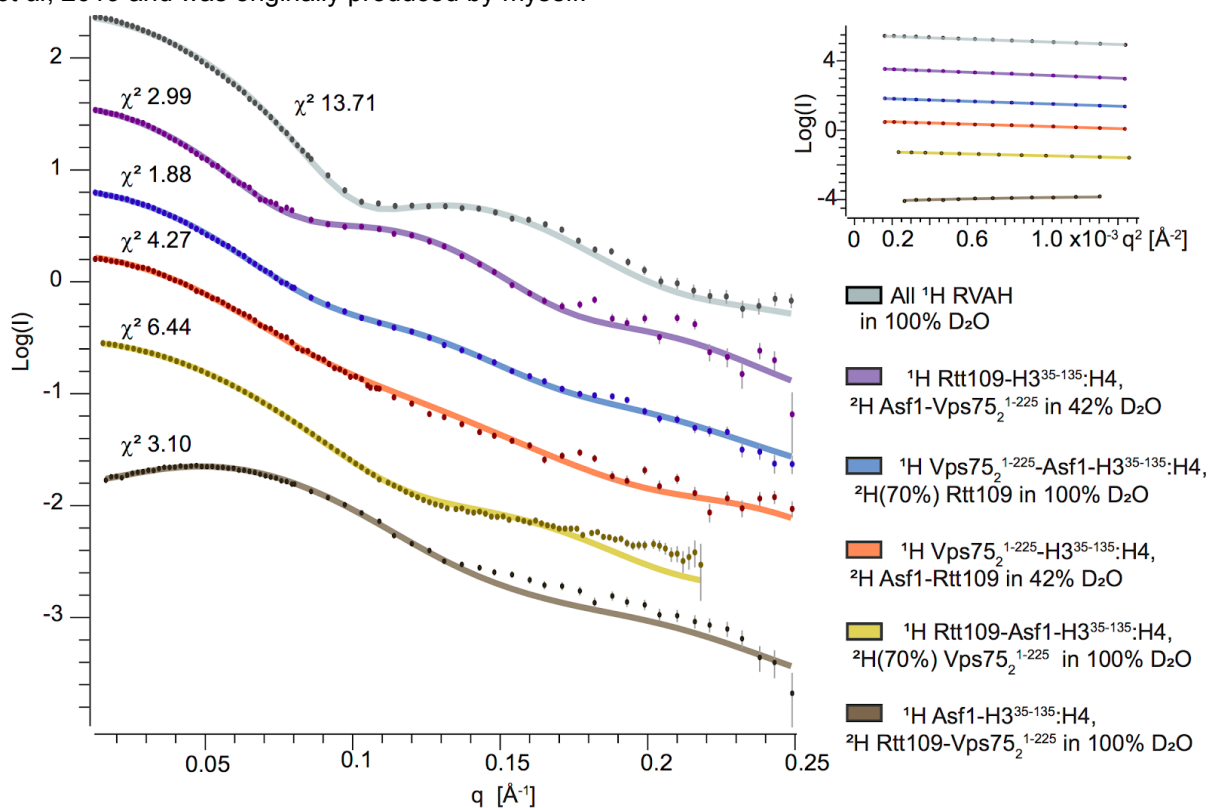
Figure 4.4.10. Fits of the curves predicted from the structure of the RVAH complex (PDB entry 6o22) to the experimental curves corresponding to the samples shown in Fig. 4.4.6. The experimental data is shown as dots, curves back-calculated from the PDB entry 6o22 are shown as lines. The plot on the right shows the Guinier region of the curves. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

4.5. The structure of the RVAH complex and its functional annotation.

The structural model of the RVAH complex obtained in this work provides several key insights into the acetylation mechanism of both the H3 N-terminal tail and H3 K56.

Overview of the RVAH structural model. The RVAH complex is characterised by a bagel-like shape with a central cavity of approximately 25 Å in diameter. The bottom half of the bagel is formed by the $Vps75_2$ dimer, while the top part is composed of Rtt109–Asf1–H3:H4 (Fig. 4.5.1). As expected from previous NMR titration experiments (section 4.2), Vps75(B) is occupied by Rtt109 and Vps75(A) binds histones via its earmuff

domain. In addition, the Vps75(A) α5 helix makes contacts with Rtt109, further facilitating the Rtt109-(H3:H4) interaction. The Vps75 acidic cavity builds the floor of the internal cavity of the bagel, while the top part is formed by the positively charged H3:H4 and Rtt109 surfaces (including the Rtt109 active center).
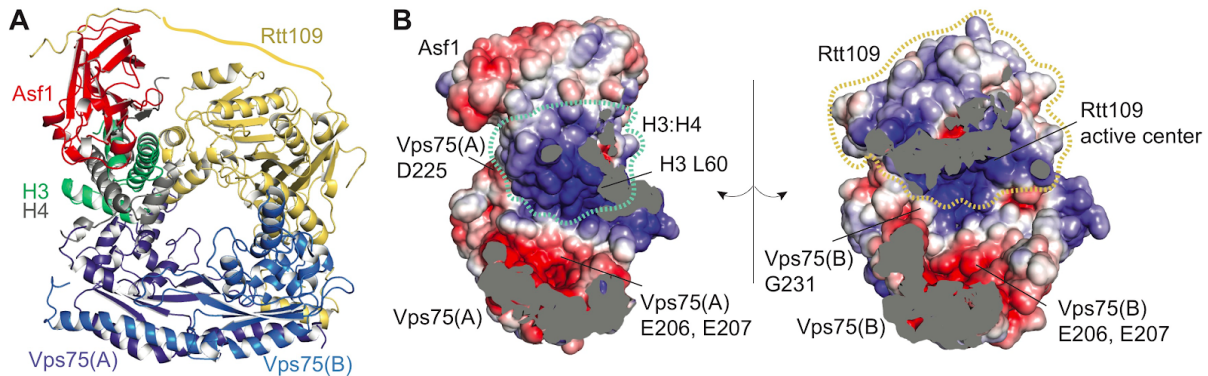
Figure 4.5.1. The RVAH complex adopts a bagel-like shape. A. The structure of the RVAH complex (PDB entry 6o22). The flat ribbon (yellow) connecting the Rtt109 core and the Rtt109[418-433] region replaces the part of the Rtt109 C-terminal tail for which no structural information is available. **B.** Electrostatic surface representation of the inner surface of the complex. D225 and G231 are the last structured amino acids of Vps75$_2$; L60 is the first amino acid of the H3 core; the Vps75-[206]EE[207] dyad is at the center of an acidic patch of the Vps75 earmuff domains. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

The Vps75 CTADs (residues 226-264) and the H3 N-terminal tail (residues 1-59) are not present in the structural model. The Vps75(A) and Vps75(B) CTADs start next to the basic surfaces of H3:H4 and Rtt109 respectively. The C-terminal end of the H3 N-terminal tail (H3 L60) faces the inner cavity of the complex and is located in the vicinity of the Rtt109 active center. The distance between H3 L60 and acCoA in our model is nearly identical to the one in the *Af* RAH structure, confirming that H3 K56 can reach into the Rtt109 active center in the RVAH complex (Fig. 4.5.2).
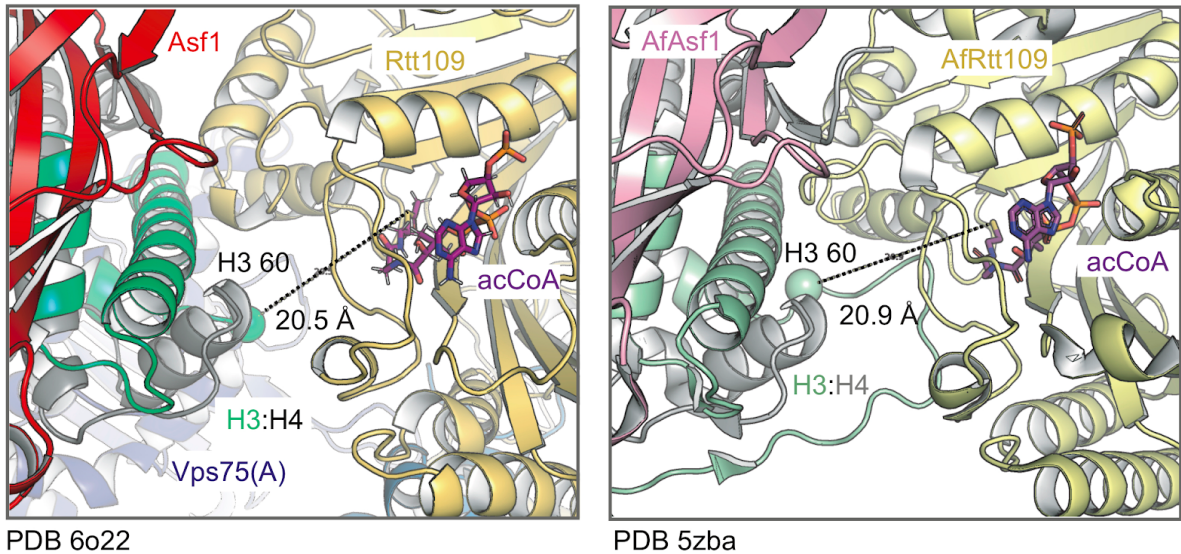
Figure 4.5.2. The RVAH structural model is compatible with Asf1 and Vps75 promoting H3 K56 acetylation. The distance between H3 L60 and the catalytic center of Rtt109 in the solution structure of the RVAH complex (PDB entry 6o22; left) is consistent with the corresponding distance in the 5zba crystal structure of the *S. cerevisiae* H3:H4 in complex with *Af*Asf1 and *Af*Rtt109 (PDB entry 5zba; right).

Intermolecular contacts within the RVAH complex. The contacts between Rtt109 and both histones are similar to the ones observed in the *Af* RAH structure (PDB entry 5zba). In particular, Rtt109 makes contacts with the long α2 helix of H3 via a salt bridge between Rtt109 R318 (R292)[6] and H3 E94, as well as between Rtt109 K323 and H3 E105 (Fig. 4.5.3, Fig. 4.5.4). The H3 E94-Rtt109 R318 (R292) interaction is formed by residues that are conserved in various fungi (Fig. 4.2.5). In agreement with this, the H3 E94C mutation with the subsequent attachment of the paramagnetic tag severely disturbed the complex formation (Fig. 4.5.3B). *In vivo* data on *S. cerevisiae* also demonstrates the importance of the H3 E94 for H3 acetylation (Zhang *et al*, 2018). Residues 299E, 300E, 301D of Rtt109 are engaged in electrostatic interactions and hydrogen bonds with both Vps75(A) and H3. Rtt109 299E and 301D interact with positively charged residues of the Vps75(A) α5 helix, while the conserved Rtt109 E300 interacts with H3 S85 and S86. These observations are supported by *in vivo* data, where Rtt109 [299]EED[301] mutation to [299]KKK[301] was demonstrated to reduce H3 K9 and K27 acetylation (Kolonko *et al*, 2010).

---

[6] A hydrogen bond between Rtt109 R292 and H3 E94 is observed in some of the structures from the final ensemble, but not in the PDB entry 6o22.
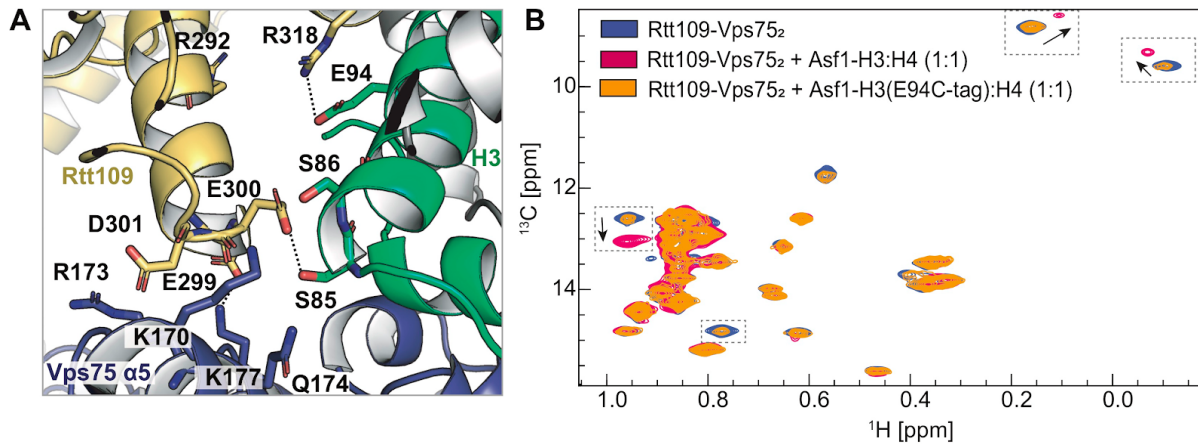
Figure 4.5.3. The contact between H3 E94 and Rtt109 is essential for the formation of the RVAH complex. A. Electrostatic contacts between Rtt109, H3 and Vps75(A) in the RVAH complex. **B**. Overlay of ^1H-^{13}C HMQC spectra of 65 µM I-labelled Rtt109 in complex with unlabelled Vps75$_2$ (blue); upon titration of Asf1–H3:H4 (pink); or Asf1–H3(E94C-tag):H4 (orange). H3(E94C-tag) contains E94C mutation followed by the attachment of paramagnetic tag coupled with the 94C residue. The peaks indicated by dashed boxes and arrows display noticeable CSPs upon addition of wild-type Asf1–H3:H4, but not Asf1–H3(E94C-tag):H4. The spectra were recorded in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. The figure is adapted from Danilenko *et al*, 2019 and originally was produced by myself.

In addition, Rtt109 interacts with the C-terminus of H4, where H4 R95 is engaged in a hydrogen bond with Rtt109 E317 (Fig. 4.5.4A). Interaction of the last residues of the H4 C-terminus with Rtt109 is supported by the ^1H-^{15}N HSQC spectra of H4: the C-terminal H4 G102 is not affected by the interaction of Asf1–H3:H4 with Vps75$_2$, but disappears upon addition of Rtt109 to the Asf1–H3:H4–Vps75$_2$ complex. *In vitro*, deletion of the H4 C-terminus (95-102) reduced H3 K56 acetylation nearly 2-fold, while H3 K9 acetylation was unaffected.

Deletion of the Rtt109 C-terminus, responsible for interaction with Asf1 and release of the H3 N-terminal tail in absence of Vps75$_2$, had a similar, albeit less dramatic effect on H3 K56 acetylation by the RVAH complex, while H3 K9 acetylation was less perturbed (Fig 4.3.4, section 4.3; Fig. 4.5.4).

The difference in the influence of these mutations on the acetylation of the different H3 sites may be explained by considering the position of these sites in the structure of the complex. H3 K56, as shown in the Fig. 4.5.2, is located near the folded core of H3:H4 and is able to fit into the Rtt109 active center only if the stretch between H3 L60 and K56 is unfolded. Any mutations influencing the interaction between Asf1–H3:H4 and Rtt109 might result in a change in the distance between the Rtt109 active center and H3 K56, physically limiting the access. On the other hand, acetylation of H3 K9, located towards the end of the long

unfolded H3 N-terminal tail, can tolerate a less close contact between Rtt109 and Asf1–H3:H4.
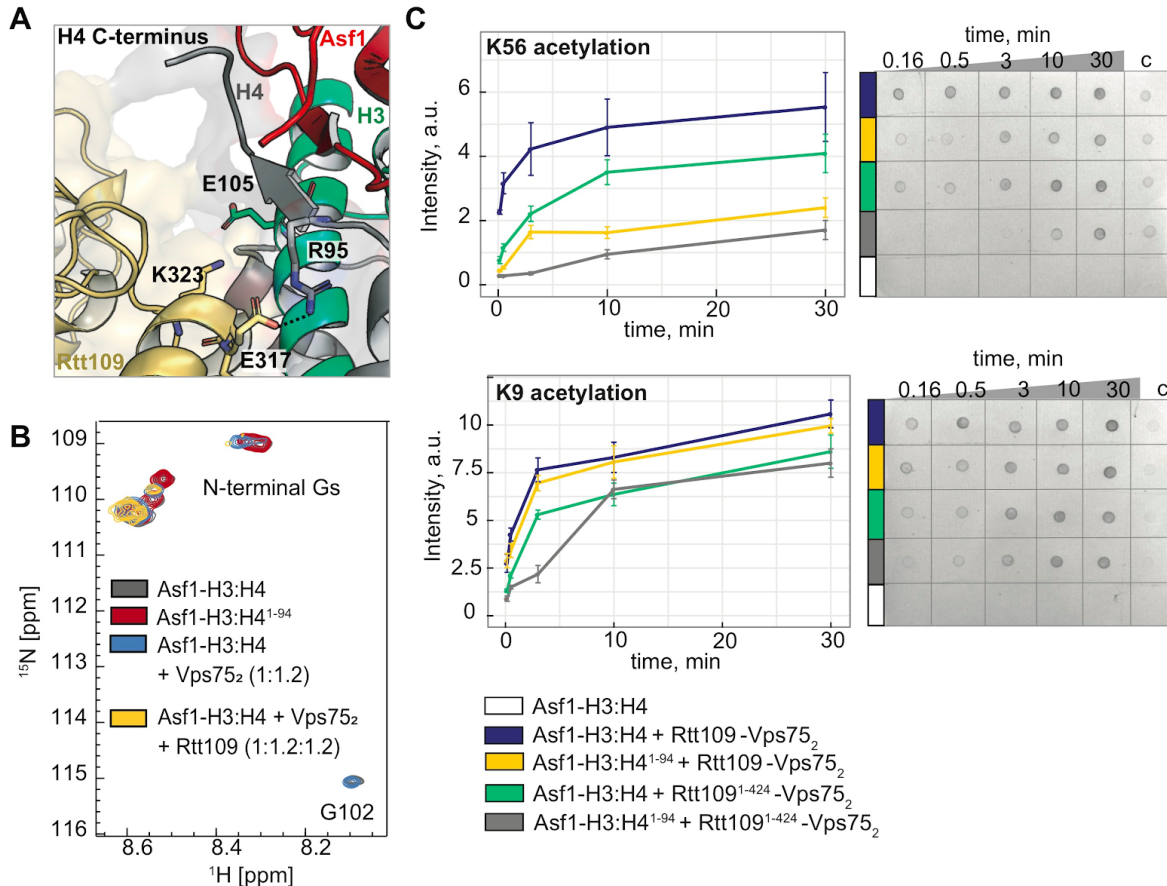
Fig. 4.5.4. The H4 C-terminus is required for H3 K56, but not H3 K9 acetylation. A. Contacts between Rtt109, the H4 C-terminus and H3 in the RVAH complex. **B**. Overlay of ^{1}H-^{15}N HSQC spectra of 50 μM ^{15}N-labelled H4 in the Asf1–H3:H4 complex (grey) upon titration of Vps75$_2$ (blue) and Rtt109 (yellow). Peaks missing in the spectra of ^{15}N-labelled H4$^{1-94}$ in the Asf1–H3:H4$^{1-94}$ complex (red) belong to the C-terminus of H4 (residues 95-102). Only peaks corresponding to the unfolded residues are visible due to the labelling scheme and the size of the complexes. The H4 G102 peak is not affected by Vps75, but disappears upon addition of Rtt109 indicating that it is no longer in the unbound unfolded state. The spectra were recorded in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 600 MHz and 298 K. **C**. Time-course of H3 K56 (top) and H3 K9 (bottom) acetylation. The dot blots (right panels) were quantified and the intensity of the dots, representing the amount of acetylated H3, were plotted (left panels). The reaction without Rtt109 (Asf1–H3:H4, white) is a control for the antibody specificity towards acetylated H3. The experiments were repeated four times; the error bars represent the standard error of the mean. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

In addition to the Vps75(A) α5 helix (Fig 4.5.3), histones H3:H4 interact with the Vps75(A) earmuff domain: H3 R69 is engaged in the stacking interaction with Vps75 F77, while H3 L65 and H4 L22 fit into the two hydrophobic pockets formed by either Vps75 residues 70-73 or Vps75 Q64 and A219, respectively (Fig. 4.5.5). Accordingly, the double H3 R69E,R72E mutation weakened the interaction of Vps75 with the histones.
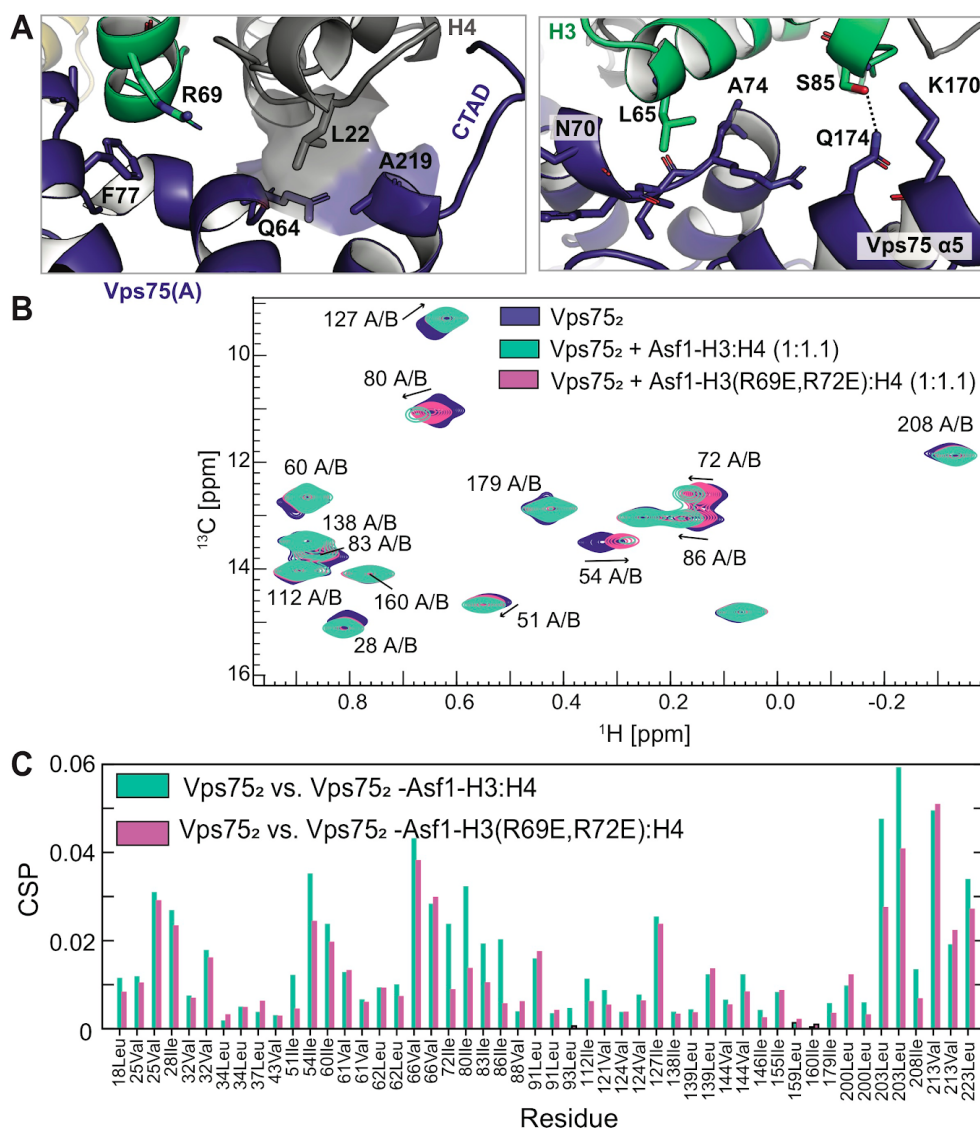
Figure 4.5.5. Interaction of Vps75(A) with H3:H4. A. Contacts between Vps75(A) and H3:H4 in the RVAH complex. **B.** Overlay of ^1H-^{13}C HMQC spectra of 60 μM ILV-methyl labelled Vps75$_2$ in the presence of Asf1–H3:H4 or Asf1–H3(R69E,R72E):H4. The H3 R69E,R72E mutation reduces the interaction between Vps75 and the histones, as evident by the reduction in CSPs. The spectra were recorded in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. **C.** CSP plots for the spectra shown in (**B**). The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

The Vps75 CTAD shields the DNA-binding surface of H3:H4. Although the Vps75 CTAD is not necessary for the RVAH complex formation (section 4.4), it was speculated that the Vps75 CTAD may contribute to H3 acetylation *in vivo*, independently of its nuclear localization signal (Keck & Pemberton, 2011; Park *et al*, 2008).

To investigate the function of the Vps75 CTAD and its potential interactions with the RVAH complex components, I recorded a set of NMR titration experiments. The Vps75 CTAD is unfolded, thus I could monitor it in ^1H-^{15}N HSQC spectra of ^{15}N-labelled protonated Vps75.

Fig. 4.5.6 shows that the Vps75 CTAD undergoes noticeable CSPs upon addition of Rtt109 and Asf1–H3:H4; however, the chemical shift dispersion indicates that the CTAD remains unfolded in all cases. The interaction with Rtt109 is confined to the last ten residues of the Vps75 CTAD, while the negatively charged part of the CTAD engages in interactions with Asf1–H3:H4.
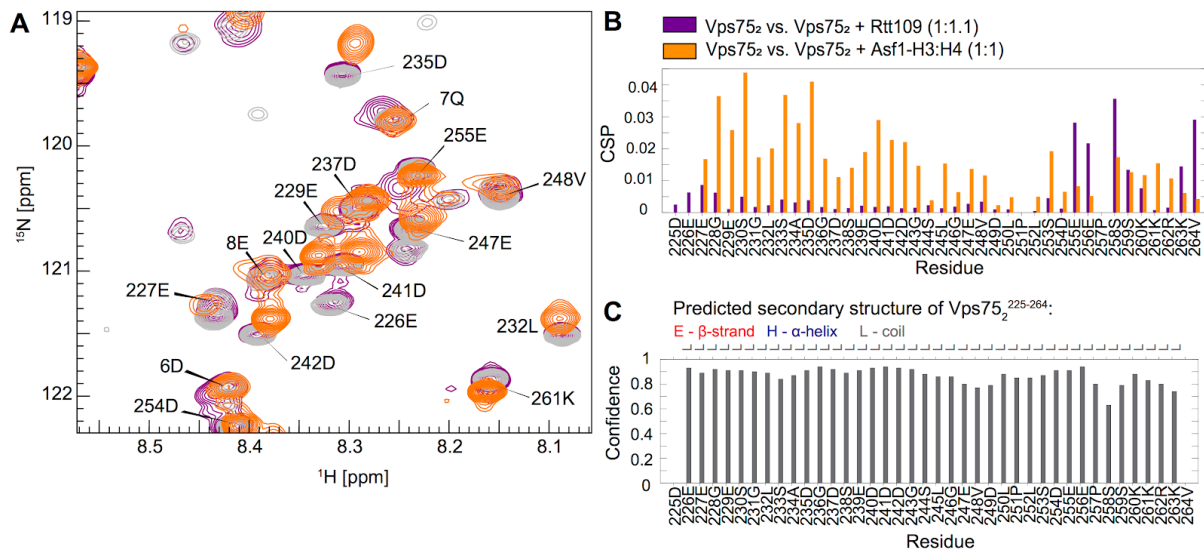
Figure 4.5.6. The Vps75 CTAD interacts with Asf1–H3:H4 but remains disordered. A. Overlay of ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled 50 μM Vps75$_2$ in isolation (grey), in complex with Rtt109 (purple) or Asf1–H3:H4 (orange). Only peaks corresponding to the unfolded residues of Vps75 are visible due to the labelling scheme and the size of the complexes. The chemical shift dispersion of the Vps75 226–264 stretch remains very narrow indicating that the Vps75 CTAD stays disordered in all complexes. Spectra were recorded in 50 mM sodium citrate pH 5.5, 150 mM NaCl, 5 mM BME at 600 MHz and 298 K. **B.** CSP analysis demonstrates that the interaction between the Vps75$_2$ CTADs and Asf1–H3:H4 mostly affects residues 227-242; these residues are not affected by the interaction with Rtt109. **C.** Chemical shift analysis of Vps75 (226–264) in the context of isolated Vps75$_2$ by TALOS-N demonstrates that the Vps75 tail is disordered. The chemical shift dispersion of Vps75 CTAD remains narrow in the RVAH complex. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

Despite its disordered nature, the Vps75 CTAD interacts with a specific surface of the Asf1–H3:H4 sub-complex, as demonstrated by comparison of the titrations of the ILV-methyl labelled H3:H4 with full-length and C-terminally truncated Vps75$_2$ (Vps75$_2$$^{1-225}$) (Fig. 4.5.7). Upon addition of full-length Vps75$_2$ to Asf1–H3:H4, The CSPs of ILV-methyl labelled H3:H4 mapped on the H3:H4 surface involved in the interaction with the earmuff domain (according to the structure of the RVAH complex) and on the H3:H4 DNA-binding surface, which in the RVAH complex is located near the start of the Vps75 CTAD. Truncated Vps75$_2$$^{1-225}$ could still bind H3:H4 via the earmuff domain, but did not interact with the H3:H4 DNA-binding surface (Fig. 4.5.7).
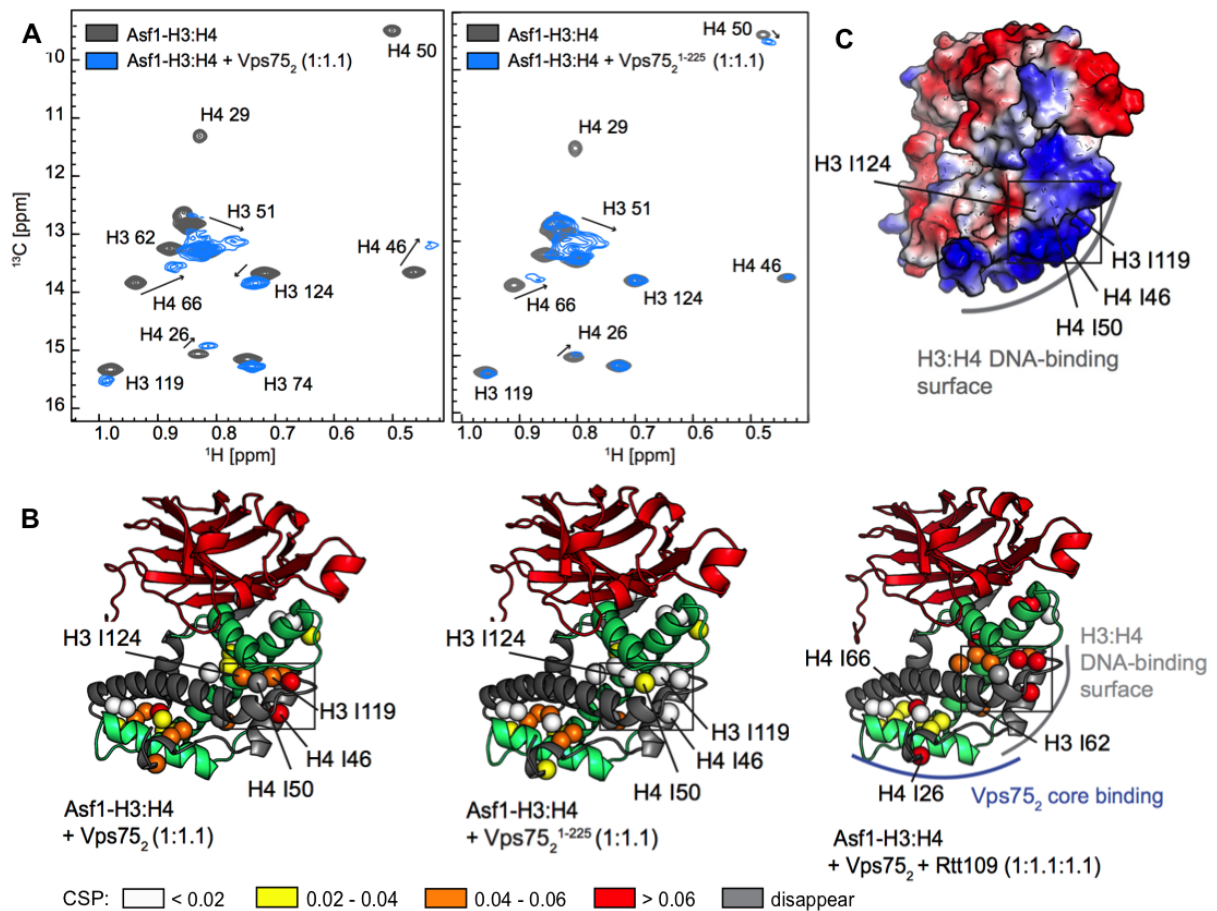
Figure 4.5.7. Vps75 CTAD H3:H4 DNA-binding. A. ^{1}H-^{13}C HMQC spectra of 60 µM ILV-labelled H3:H4 as part of the Asf1–H3:H4 sub-complex in the presence of 0.5 and 1.1 equivalents of Vps75$_2$ (left) or Vps75$_2$$^{1–225}$ (right). The spectra were recorded with in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. **B.** CSPs of the spectra in (**A**) mapped on the Asf1–H3:H4 structure (PDB entry 2hue; Asf1, H3 and H4 are shown in red, green and grey, respectively). The differences in CSPs between titration with the full-length Vps75$_2$ (left) and Vps75$_2$$^{1–225}$ (middle) are marked by a rectangle. The CSPs for the H3:H4 surface that interacts with the Vps75$_2$ earmuff domain (Vps75$_2$ folded core) are similar in both cases. The right panel shows the CSPs of Asf1–H3:H4 in the RVAH complex; the interaction with Vps75$_2$ (including the CTAD) is preserved. **C.** Electrostatic surface representation of Asf1–H3:H4. The surface bound by the Vps75 CTAD is indicated by a rectangle. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

These data suggested that the Vps75 CTAD covers the H3:H4 DNA-binding surface. Interestingly, the Vps75 CTAD does not appear to become structured in this interaction. The Vps75 CTAD lacks the motif (an acidic stretch with one or more aromatic residues, most often Y) identified previously as characteristic for the disordered domains of chaperons that wrap around the DNA-binding surface of histone dimers or octamers (Chen *et al*, 2015; Huang *et al*, 2015; Warren *et al*, 2017). This interaction of the Vps75 CTAD with H3:H4 could act both (i) to shield H3:H4 from nonspecific interactions, including spurious interaction with

DNA, (ii) to reduce the positive charge of the inner cavity of the RVAH complex in order to attract the H3 N-terminal tail to the Rtt109 active center.

In summary, the structural model of the RVAH complex revealed that both chaperones facilitate formation of a complex with an inner cavity, in which the Rtt109 active center and H3 K56 are placed next to each other. H3 K56 acetylation is promoted by Asf1 and Vps75 via a substrate presentation mechanism, where Vps75 aids the formation of the RVAH complex, while Asf1 interacts with Rtt109 C-terminus, and facilitates the interaction between Rtt109 and the H4 C-terminus. Both of these contacts were found to be important for the H3 K56 acetylation. However, the question remains open of how the RVAH complex promotes acetylation of H3 K9, K23 and K27 located in the long disordered and positively charged H3 N-terminal tail.

4.6. Vps75 promotes H3 N-terminal tail acetylation via two acidic structural elements.

Previous *in vivo* data as well as NMR data from this work identified two regions of Vps75 that could play a role in the H3 N-terminal tail acetylation: 1. The acidic patch composed of two conserved E (Vps75 E206, E207; later referred to as $^{206}EE^{207}$), which contributes to the negatively charged part of the RVAH complex cavity (Fig. 4.5.1); 2. The Vps75 CTAD.

A triple mutation of the $^{205}SEE^{207}$ to $^{205}AAA^{207}$ decreases H3 K9 acetylation by 10% *in vivo*[7], but does not affect the histone binding affinity of Vps75 (Berndsen *et al*, 2008). In agreement with this, in the RVAH structure this acidic patch is not involved in the interaction with the H3:H4 folded core. However, ILV residues next to it (I208, L203, L200) display CSPs upon addition of the Asf1–H3:H4, indicating a potential interaction with the H3 N-terminal tail.

As discussed in section 4.5, one of the two Vps75 CTADs of the Vps75 dimer interacts with the DNA-binding surface of H3:H4. However, it remains unfolded and potentially available for the interaction with the H3 N-terminal tail.

To test whether any of these structural elements are involved in the interaction with the H3 N-terminal tail, I used full-length Vps75$_2$, Vps75$_2$$^{1–225}$, and a Vps75$_2$ construct with E206 and E207 mutated to alanines (^{AA}Vps75$_2$) in NMR titration experiments with ^{15}N labelled H3 (Fig. 4.6.1).

---

[7] In *S. cerevisiae* H3 K9 acetylation is carried out by both Rtt109 and Gcn5. In the cited study, the influence of the Vps75 mutation on H3 K9 acetylation was monitored in a strain containing wild-type Gcn5, which may have masked the effect of the mutation on the Rtt109-dependent acetylation.

Figure 4.6.1. The Vps75 ^{206}EE^{207} motif and the CTAD are required for the chaperoning of the H3 N-terminal tail. A. Overlay of ^{1}H-^{15}N HSQC spectra of 60 μM H3 in the H3:H4 dimer, in the Asf1–H3:H4 sub-complex, and in the Asf1–H3:H4 sub-complex upon addition of Vps75$_2$, Vps75$_2$$^{1-225}$ or ^{AA}Vps75$_2$. **B.** CSP plots for the spectra in (**A**). Top panel, comparison of the CSPs between the free H3:H4 and the sub-complexes with Asf1 and Asf1 and Vps75 together. Bottom panel, comparison of the CSPs between the Asf1–H3:H4 and the sub-complexes with Vps75. The residues in the H3 N-terminal tail are in the free, unfolded state in the H3:H4 dimer; upon addition of Asf1 residues 19-28 interact with Asf1 in the Asf1–H3:H4 sub-complex; upon addition of Vps75$_2$ residues 15-32 are released from Asf1 and interact with Vps75 in the Asf1–H3:H4–Vps75$_2$ sub-complex. In the sub-complexes with ^{AA}Vps75$_2$ and Vps75$_2$$^{1-225}$, the mutant forms of Vps75 appear to be unable to bind the histone tail like the wild-type Vps75. The spectra were recorded with in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K.

Vps75 is able to release the H3 N-terminal tail from Asf1 and bind it (Fig. 4.3.2, 4.3.3). However, neither one of the Vps75 mutants is able to bind the H3 N-terminal tail like the wild-type chaperone (Fig. 4.6.1B). In the Asf1–H3:H4–Vps75$_2$$^{1-225}$ sub-complex, the chemical shifts of the H3 N-terminal tail and the intensity of the peaks are very similar to the ones observed in the spectrum for the free H3:H4. This suggests that although Vps75$_2$$^{1-225}$ is able to release the H3 N-terminal tail from Asf1, it fails to engage the H3 residues 15-32. In the Asf1–H3:H4–^{AA}Vps75$_2$ sub-complex, the resonances of the H3 N-terminal tail appear to be

only partially bound to Vps75. Taken together, the mutant data indicates that both the Vps75 CTAD and the $^{206}EE^{207}$ region play a role in H3 N-terminal tail acetylation.

In agreement with these data, both Vps75 mutants were unable to fully release the H3 tail from Asf1 (Fig. 4.6.2). The CSPs of the ILV-methyl groups of Asf1 displayed an equilibrium between the H3-bound and free states upon addition of $Vps75_2^{1-225}$ or $^{AA}Vps75_2$.

Figure 4.6.2. The Vps75 $^{206}EE^{207}$ and the CTAD are required for the release of the H3 N-terminal tail from Asf1. A. Overlay of ^1H-^{13}C HMQC spectra and of 60 μM ILV-methyl labelled Asf1 in the Asf1–H3$^{29-135}$:H4 sub-complex, Asf1–H3:H4 sub-complex in isolation and Asf1–H3:H4 sub-complex upon addition of wild-type $Vps75_2$, $Vps75_2^{1-225}$ or $^{AA}Vps75_2$. In the Asf1–H3$^{29-135}$:H4 sub-complex, assembled with H3 lacking the 1-28 residues of the N-terminal tail, Asf1 residues I60 and V62 are not bound by the H3 N-terminal tail. In the Asf1–H3:H4 sub-complex, addition of $Vps75_2$ leads to the release of the H3 N-terminal tail; chemical shifts for I60 and V62 are the same as in the Asf1–H3$^{29-135}$:H4 sub-complex. Addition of $Vps75_2^{1-225}$ and $^{AA}Vps75_2$ to the Asf1–H3:H4 sub-complex leads to a partial release of the H3 N-terminal tail; chemical shifts for I60 and V62 are between the H3 N-terminus bound state and free state. **B.** CSP plots for the spectra in (**A**). $Vps75_2^{1-225}$ and $^{AA}Vps75_2$ are able to only partially release the H3 N-terminal tail from Asf1, as indicated by lower CSPs for the Asf1 residues 60-69 (in comparison to the wild-type $Vps75_2$).

The inability of $Vps75_2^{1-225}$ and $^{AA}Vps75_2$ to completely release the H3 tail from Asf1 and engage it in interactions demonstrates that both the acidic patch $^{206}EE^{207}$ and the CTAD of Vps75 contribute to recruiting the H3 tail away from Asf1 towards the central cavity of the RVAH complex.

The difference in the $Vps75_2^{1-225}$ and $^{AA}Vps75_2$ interactions with the H3 N-terminus suggested that Vps75 CTAD is essential for binding the H3 N-terminal tail, while the acidic patch $^{206}EE^{207}$ could play a role in the release of the H3 N-terminal tail without a direct interaction (Fig. 4.6.1, H3 N-terminal tail residues 15-32 are bound by $^{AA}Vps75_2$, but not by

Vps75$_2$$^{1–225}$). As the acidic patch is located near the start of the H3 N-terminal tail (L60) in the RVAH complex (Fig. 4.5.1), I hypothesised that it may engage in the interaction with the H3 residues 35-59, not visible in the ^{1}H-^{15}N HSQC spectra of the 15N labelled H3. Through this interaction, the H3 N-terminal tail could be pulled into the Vps75$_2$ acidic cavity, disrupting its interaction with Asf1.

The role of the acidic patch was further investigated by molecular dynamics simulations of the RVAH complex. In the simulations, H3 residues 35-59 were modelled either in a random conformation or in a conformation where H3 K56 was placed directly into the Rtt109 active center (Fig. 4.6.3). Independently of the starting conformation, the H3 N-terminal tail engaged in a stable interaction with the ^{206}EE^{207} acidic patches on both Vps75 chains via H3 R52 and H3 R53. H3 R52 was predominantly bound by the Vps75(A) E207, while H3 R53 formed a hydrogen bond with the Vps75(B) E207. These interactions locked H3 K56 away from the active center of the enzyme. In agreement with this, H3 K56 acetylation was increased by 1.3-fold in the RVAH complex reconstituted with ^{AA}Vps75, while H3 K9 acetylation was unaffected.
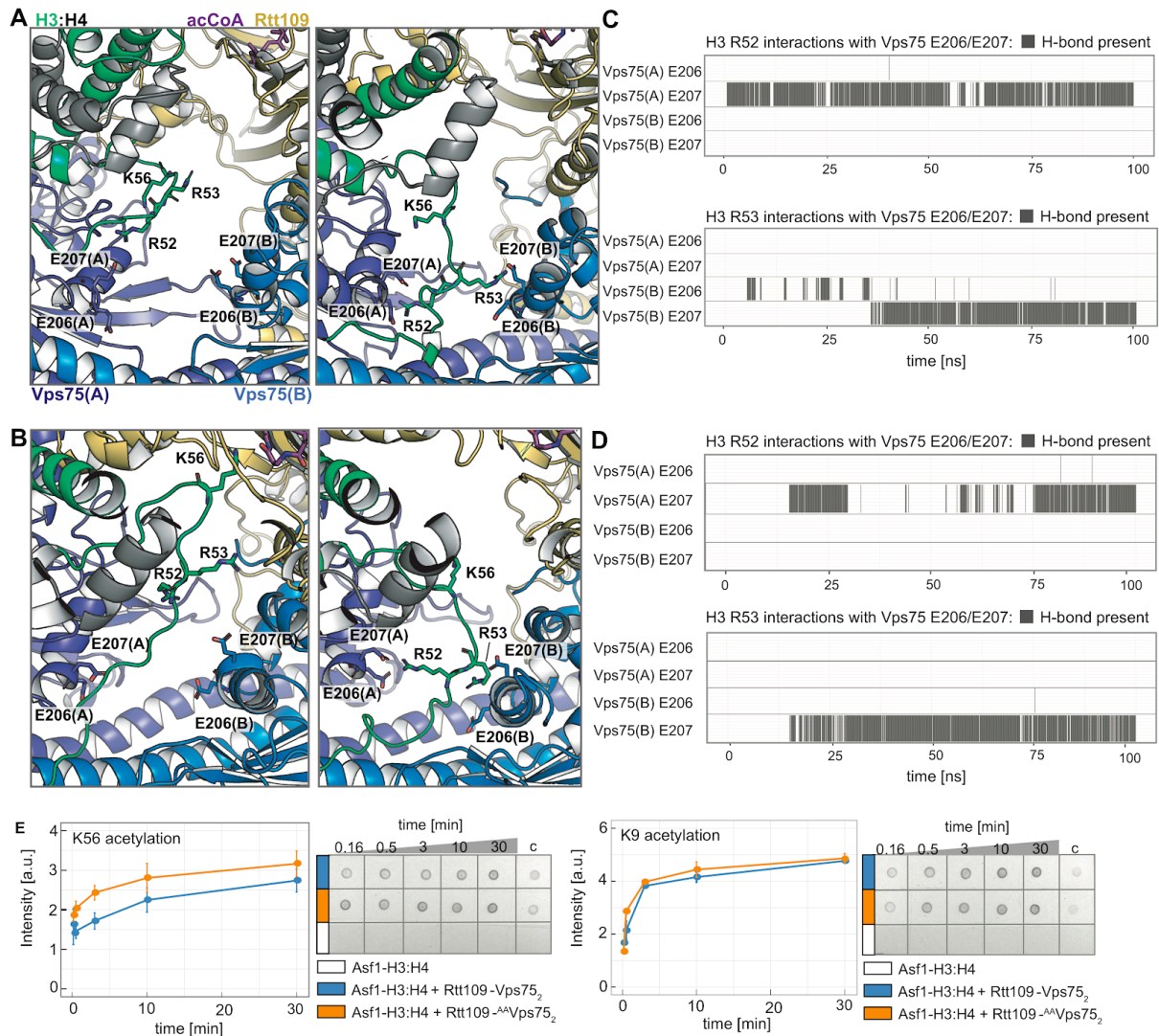
Figure 4.6.3. The Vps75 acidic patch captures H3 R52 and R53 and keeps H3 K56 away from the Rtt109 active center. A-B. Snapshots from the MD simulation trajectories at 0 and 100 ns. Left, starting conformation of the RVAH complex with H3 residues 35-59. Right, final conformation at 100 ns. Residues H3 35-59 were modelled without any restrains (**A**) or with H3 K56 placed into the Rtt109 active center in the starting conformation (**B**). **C-D.** Analysis of hydrogen bonds present between H3 R52, R53 and Vps75 E206, E207 in the 100 ns MD simulations of the RVAH complex from (A) and (B), respectively. Both MD simulations result in the formation of stable hydrogen-bonds (H-bonds) between H3 R52 and Vps75(A) E207 and between H3 R53 and Vps75(B) E207, which are incompatible with H3 K56 reaching the Rtt109 catalytic center. **E.** Time-course of H3 acetylation in Asf1–H3:H4 by Rtt109–Vps75$_2$ (blue) or Rtt109–^{AA}Vps75$_2$ (orange). Acetylation reactions were allowed to proceed for the indicated time, spotted on the membrane and stained with antibodies against acetylated H3 K56 (left) or H3 K9 (right). The dot blots were quantified and the intensity of the dots, representing the amount of acetylated H3, were plotted. The reaction without Rtt109 (Asf1–H3:H4, white) is a control for the antibody specificity towards the acetylated H3. The experiments were repeated four times; the error bars represent the standard error of the mean. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

I then investigated further the role of the Vps75 CTAD and its interaction with the H3 N-terminal tail. While this interaction is preserved in the RVAH complex, the intensity of

peaks of the H3 tail backbone amide resonances decreased in comparison to the spectra of the $Vps75_2$–Asf1–H3:H4 sub-complex (Fig. 4.6.4). This difference suggests that the H3 N-terminal tail has a lower mobility in the RVAH complex, possibly due to its confinement in the central cavity of the complex. To test the hypothesis that the Vps75 CTAD interacts with the H3 N-terminal tail, attracting it towards the RVAH complex cavity and hence the Rtt109 active center, I calculated the *ab initio* models of the RVAH complex from the SANS datasets recorded on complexes reconstituted with the full-length proteins, with $Vps75_2^{1–225}$, or with $Vps75_2^{1–225}$ and $H3^{35–135}$ (see section 3.6). The *ab initio* modelling of the RVAH complex with $Vps75_2^{1–225}$ and $H3^{35–135}$ returned an envelope with a large middle cavity of ~25 Å in diameter, similar to the structural model of the RVAH complex (PDB entry 6o22, Fig. 4.5.1). In the envelope generated for the RVAH complex with both the Vps75 CTAD and the H3 N-terminal tail, the area corresponding to the central cavity was filled with dummy atoms, indicating that the Vps75 CTAD and the H3 N-terminal tail occupy it. The fact that the *ab initio* model for the RVAH complex containing $Vps75_2^{1–225}$ had a cavity with dimensions identical to the envelope generated for the complex containing $Vps75_2^{1–225}$ and $H3^{35–135}$ demonstrates that the H3 N-terminal tail either does not localize to the central cavity of the RVAH complex in the absence of the Vps75 CTAD[8].

---

[8] The differences between the *ab initio* models could in principle be due to other factors: only Vps75 CTADs contribute to filling of cavity, hence their truncation has the same effect as the truncation of both Vps75 CTAD and the H3 N-terminal tail. However, NMR data of Fig. 4.6.4 together with the identity of the cavity parameters in the two models and the influence of Vps75 CTAD on H3 K9 acetylation support the model where the Vps75 CTAD attracts the H3 N-terminal tail to the cavity.
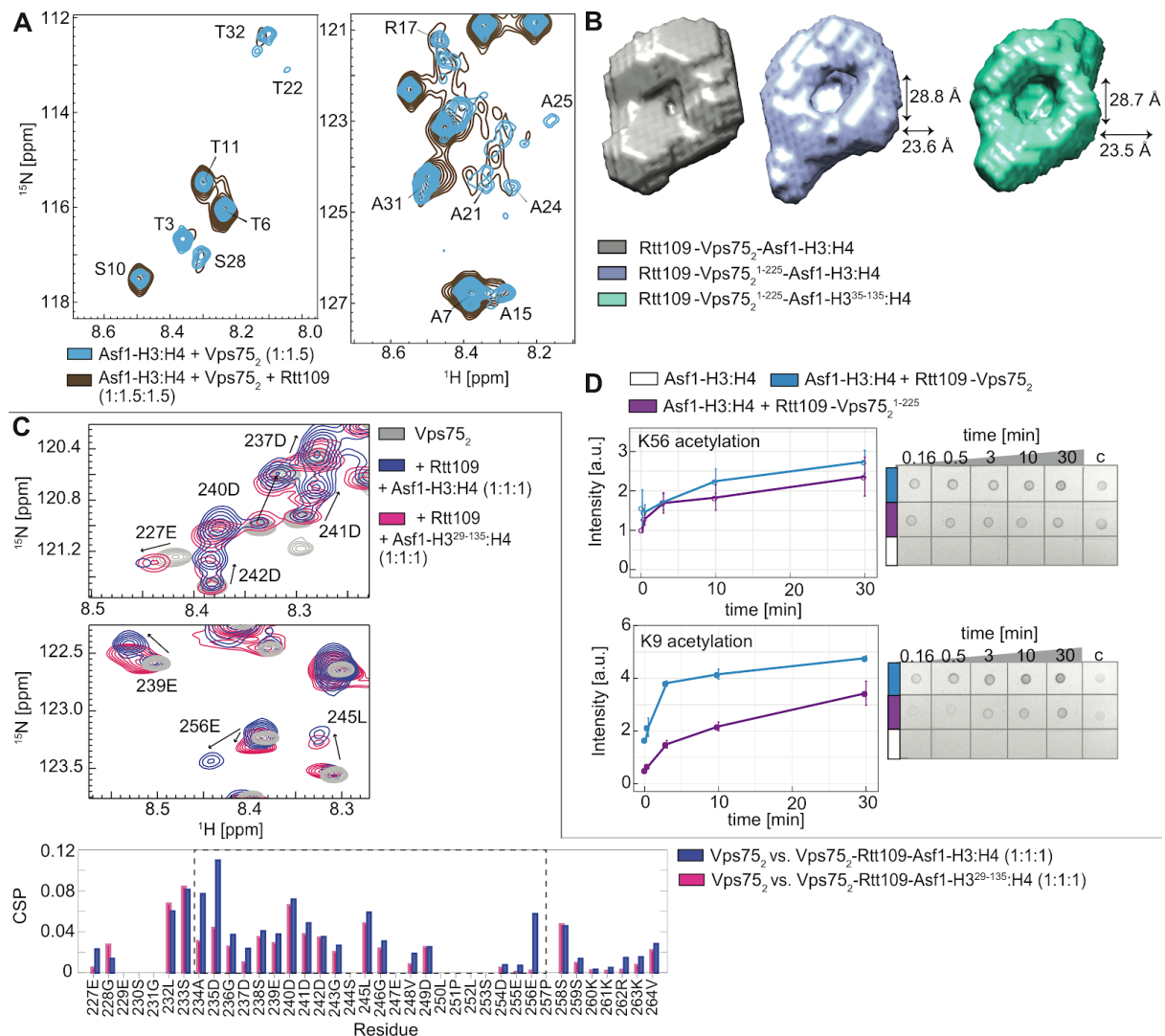
Figure 4.6.4. The Vps75 CTAD guides acetylation of the unfolded H3 N-terminal tail. A. Overlay of ^{1}H-^{15}N HSQC spectra of 60 µM H3 in the Asf1–H3:H4–Vps75$_2$ sub-complex in isolation and upon addition of Rtt109. The chemical shift of the H3 residues 1-32 are unaltered in the RVAH complex, but the intensity of the cross-peaks is decreased, suggesting a reduced mobility of the residues. The contour levels for the spectrum of H3 in the RVAH complex were lowered to make all peaks visible. **B.** SANS-derived *ab initio* envelopes of Rtt109–Vps75$_2$$^{1–225}$–Asf1–H3:H4 (middle) and Rtt109–Vps75$_2$$^{1–225}$–Asf1–H3$^{35–135}$:H4 (right) display an empty central cavity, which is filled by scattering units in the complex reconstituted with the full-length proteins (left). **C.** Overlay of ^{1}H-^{15}N HSQC spectra of 60 µM Vps75$_2$ in isolation (grey) and in the RVAH complex with either full-length H3 (blue) or H3$^{29-135}$ (pink). Only residues belonging to the Vps75 CTADs are visible due to the labelling schemes and the size of the complex. CTADs of Vps75(A) and Vps75(B) display different chemical shifts in the RVAH complex, causing the doubling of the resonances: the peaks of one of the two CTADs remain in the state similar to the free Vps75 CTAD, while those of the other CTAD move to new positions. The CSPs for the residues 234–256 are lower in the Rtt109–Vps75$_2$–Asf1–H3$^{29–135}$:H4 complex (CSP plot, bottom), indicating that these residues interact with H3 1-28. All spectra were recorded with in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. **D.** Time-course of H3 acetylation (top, K56; bottom, K9) in the Asf1–H3:H4 by Rtt109–Vps75$_2$ (blue) or Rtt109–Vps75$_2$$^{1–225}$ (purple). The experiments were repeated four times; the error bars represent the standard error of the mean. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

Finally, NMR titration experiments with the ^{15}N-labelled protonated Vps75 demonstrated that the CTAD residues 234-256 are involved in the interaction with the H3 N-terminal tail (Fig. 4.6.4C). Overall, regions of the unfolded Vps75 CTAD and the H3 N-terminal tail that were identified to interact with each other are composed of stretches of negatively or positively charged amino acids:

```
Vps75 (234-256):  ADGDSEDDDGSLGEVDLPLSDEE
H3 (15-32):           APRKQLATKAARKSAP
```

suggesting that "fuzzy" electrostatic interactions can occur between the two without folding. In agreement with the hypothesis that the Vps75 CTAD attracts the H3 N-terminal tail towards the central cavity via weak electrostatic interactions, the deletion of Vps75 CTAD reduced H3 K9 acetylation 2-fold *in vitro*. H3 K56 acetylation remained unaffected, indicating that the functional state of the RVAH complex was not altered by the truncation.

Thus, this work supports a model in which Vps75 acts as the Rtt109 acetylation specificity switch in a two-step mechanism. The model is summarized in the following scheme (Fig. 4.6.5).

Figure 4.6.5. Vps75 promotes H3 K9 acetylation by a two-step mechanism. Left, H3 K56 acetylation is promoted by a substrate presentation mechanism, where H3 K56 is placed next to the Rtt109 active center by the interactions between Rtt109 and Asf1/H4/H3. Right, the positively charged H3 N-terminal tail is confined in the proximity of the enzyme active center by two acidic regions of Vps75$_2$: the Vps75 acidic patch in the earmuff domain binds to H3 R52, R53; the Vps75 CTAD covers the positively charged surface of H3:H4 and engages in fuzzy electrostatic interactions with the H3 N-terminal tail. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

The H3 K56 acetylation is promoted by its placement near the enzyme catalytic pocket, in a classical substrate presentation mechanism. Contrarily, H3 K9, K23 and K27 are located in the unfolded and positively charged part of the H3 tail and cannot be easily localized to the active site because of their high mobility. In the absence of Vps75$_2$, these residues may

reach the Rtt109 active center as the H3 tail randomly samples the conformational space; however, in this case acetylation is inefficient. Contrarily, Vps75$_2$ facilitates acetylation of the H3 N-terminal tail by using its CTAD to confine the H3 tail to the RVAH cavity and thus in the proximity of the Rtt109 binding pocket. Both the H3 tail and the Vps75 CTAD remain flexible and thus preserve a large part of their entropy. In addition, the flexibility of the H3 tail allows acetylation of residues at multiple sites (H3 K9, K23 and K27). The process of H3 tail acetylation is aided by the Vps75 acidic patch, which interacts with H3 R52 and R53, disfavoring H3 K56 acetylation.

4.7. Assignment of the methyl groups.

In order to enable the above-mentioned NMR titration experiments as well as the PRE measurements, I assigned Vps75 (backbone and ILV-methyl groups), Rtt109 (I-methyl groups), H3:H4 (ILV methyl-groups and the backbone of the H3 N-terminal tail). Asf1 assignments were done by Dr. L. Lercher and described in Lercher *et al*, 2018.

Stereospecific assignment of Vps75 ILV-methyl groups in the free state and in the RVAH complex. The strategy and experimental details of the Vps75 assignment are described in section 3.4. The Vps75$^{1-225}$ construct, where the unfolded CTAD of Vps75 (residues 226–264) was truncated, was used for the assignment experiments to reduce the overlap of the cross-peaks. A total of 56% of backbone amide resonances were assigned (Fig. 4.7.1). ILV-methyl groups of Vps75$^{1-225}$ were assigned by connecting them to the backbone resonances (via Ile,Leu-(HM)CM(CGCBCA)NH and Val-(HM)CM(CBCA)NH experiments), point mutations, a 4D NOESY experiment (4D HCCH HMQC-NOESY-HMQC) and stereospecific labelling of the pro-S LV-methyl groups. In total, 17/17 I, 10/13 L, and 10/10 V residues were assigned (Fig. 4.7.2).

The ILV-methyl groups assignments were transferred from free Vps75$^{1-225}$ to the full-length Vps75 and then to the Vps75 in the context of the RVAH complex following the CSPs for each peak in the individual titrations of Vps75$_2$ with Asf1–H3:H4, Rtt109 and Rtt109–Asf1–H3:H4.

Figure 4.7.1. Backbone assignment of free Vps75$^{1-225}$. ^{1}H-^{15}N HSQC spectrum of 0.8 mM free Vps75$_2$$^{1-225}$ (0.8 mM) in 20 mM Tris–HCl pH 7.5, 200 mM NaCl, 1 mM DTT at 850 MHz and 300 K. The upper panel shows the full spectrum; the lower panel an excerpt thereof. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.
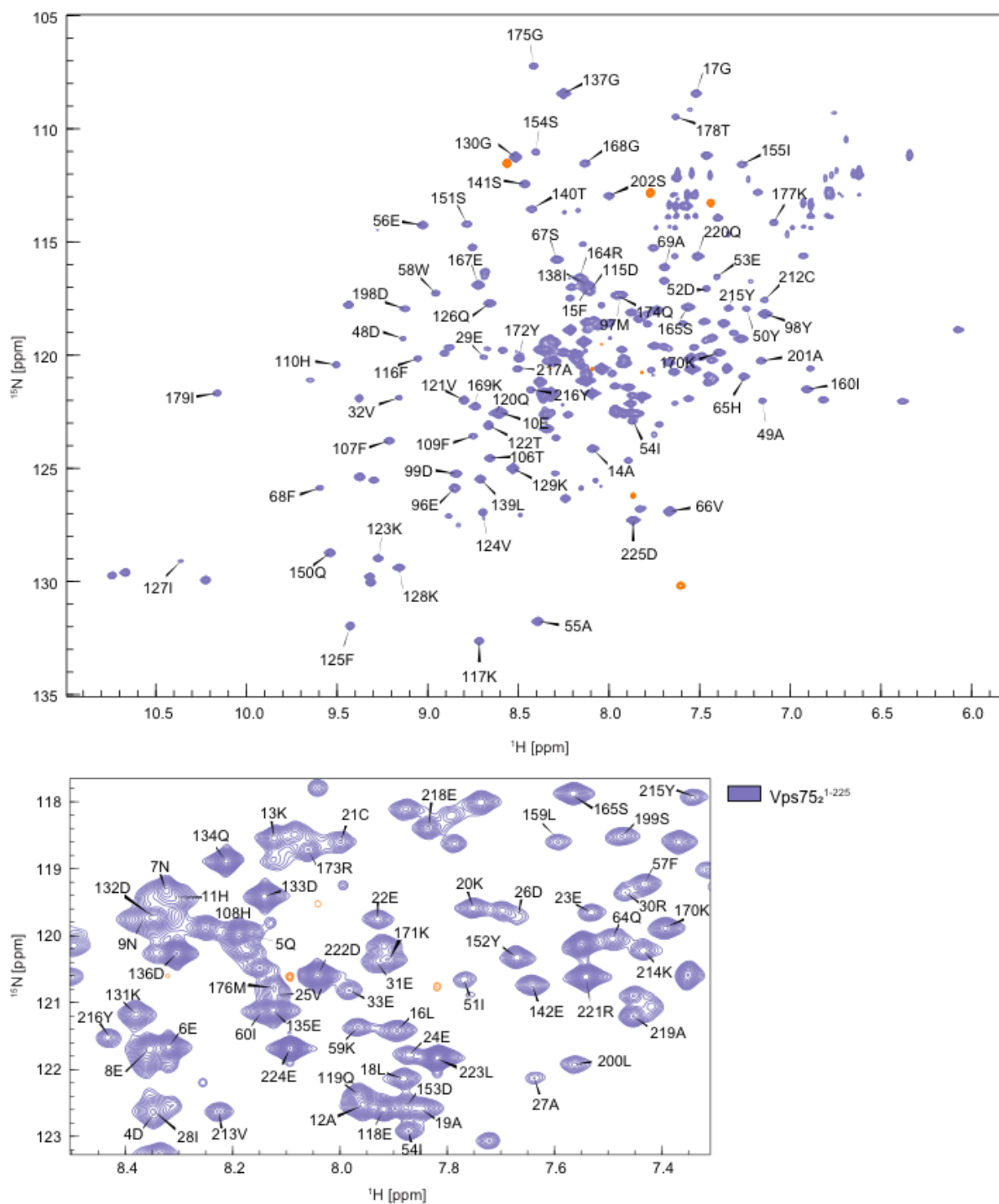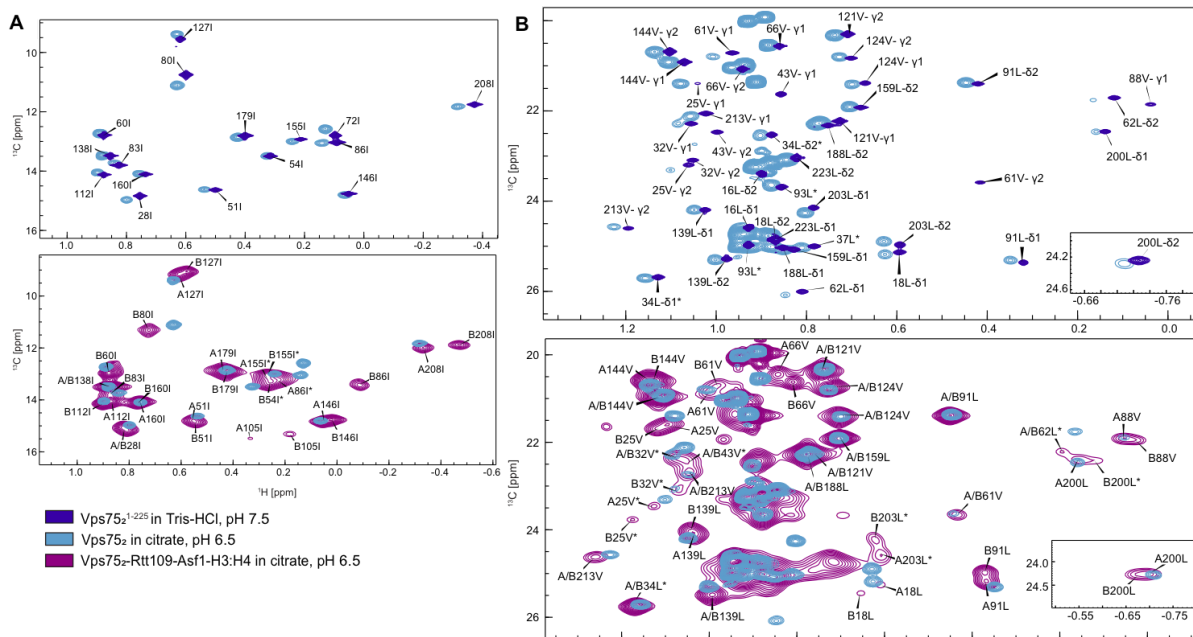
Figure 4.7.2. Assignment of Vps75 ILV-methyl resonances. A,B. Overlay of ^{1}H-^{13}C HMQC spectra of ILV-labelled Vps75$_2$$^{1-225}$ (dark blue) in 20 mM Tris–HCl pH 7.5, 200 mM NaCl, 1 mM DTT at 850 MHz and 300 K, Vps75$_2$ (light blue) and Vps75$_2$ as part of the Rtt109–Vps75$_2$–Asf1–H3:H4 complex (purple) in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K. (**A**) Ile region (**B**): Leu/Val region with stereospecific assignment. The assignments are shown for Vps75$_2$ and Vps75$_2$ in the Rtt109–Vps75$_2$–Asf1–H3:H4 complex. In the latter spectrum, A and B indicate the Vps75(A) and Vps75(B) chains, respectively. The asterisk marks tentative assignments. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

Assignment of Rtt109 I-methyl resonances. Due to the large size of Rtt109 (50 kDa versus the 32 kDa Vps75) and the poorer stability and solubility of the isolated protein, we chose to assign the methyl groups without the prior backbone assignment. Previously Dr. L. Lercher recorded a dataset for the ILV-methyl groups assignment on perdeuterated, ILV^{ProS}-methyl labelled Rtt109 sample, which yielded an insufficient number of NOE cross-peaks and no cross-peaks between I and LV. I then recorded a second dataset on perdeuterated, I-methyl labelled Rtt109 in complex with perdeuterated Vps75, where the presence of Vps75 improved sample stability, while labelling of only I residues allowed to reduce the spectral width.

In this dataset I observed four groups of resonances (three groups included just two resonances each, one group included three ones) that gave NOE cross-peaks among each other, indicating that these methyl groups are located in the vicinity of each other (up to ~6.5 Å distance). I then used the structure of Rtt109 from PDB entry 3q66 to predict all isoleucine residues pairs for which the methyl groups are within this distance from each other.

Comparison of the experimental NOESY data with the data expected from the Rtt109 structure allowed me to assign each of the groups to a defined I residues in Rtt109.

NOESY data: groups of I residues which δ1-methyl groups are close to each other:

1. Ile – Ile – Ile ← unique group of three
2. Ile – Ile ← located in α2β5 loop
3. Ile – Ile ←
4. Ile – Ile ← by PRE measurements and CSPs

X-ray data: groups of I residues which δ1-methyl groups are close to each other:

1. Ile 189 – Ile 332 – Ile 83
2. Ile 138 – Ile 150
3. Ile 365 – Ile 383
4. Ile 282 – Ile 294

While I could assign each group of residues (e.g. by characteristic chemical shifts, CSPs in complexes with Vps75 and RVAH), I could not specifically assign I resonances within each group. Thus, all resonances in the first group from the scheme above were assigned as I189/I332/I83 etc. Further assignment was complicated by the low expression and high instability of Rtt109 bearing point mutations of ILV residues. In total, I could assign 9/31 I-methyl group resonances.
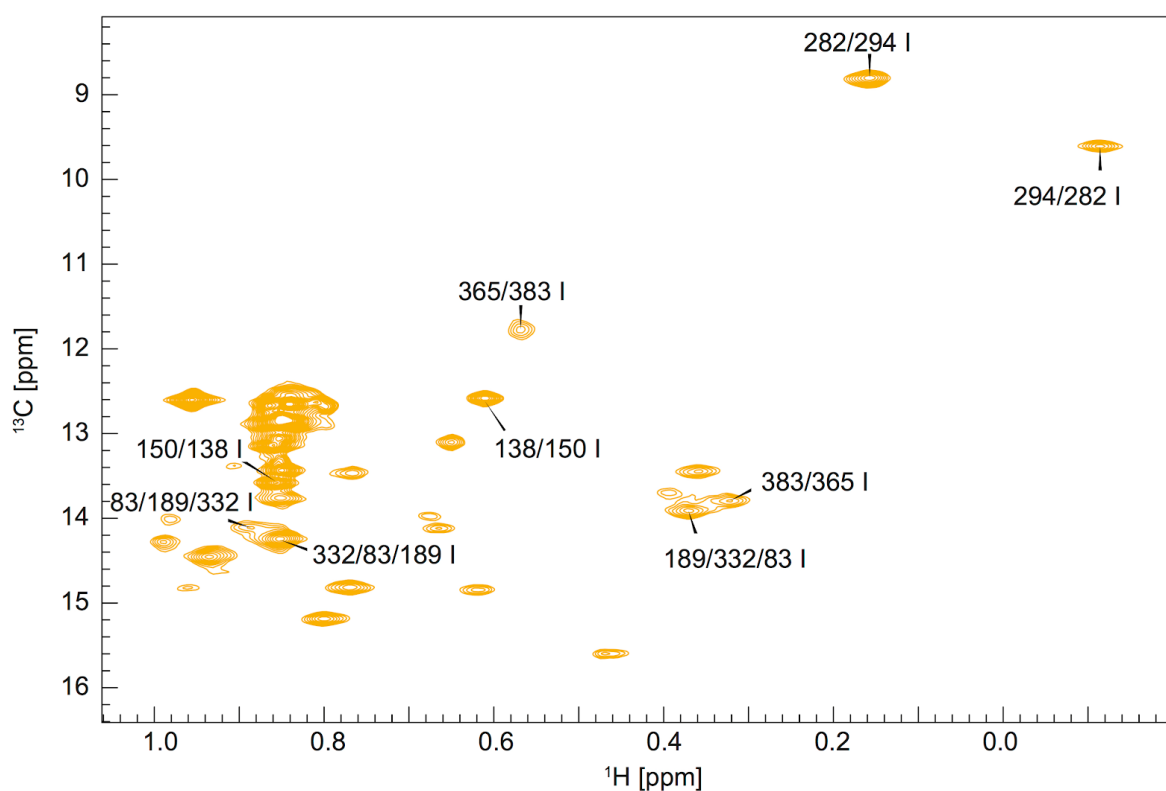
Figure 4.7.3. Assignment of Rtt109 I-methyl resonances. I-methyl labelled Rtt109 at ~0.6 mM in complex with perdeuterated Vps75 in 50 mM sodium citrate pH 6.5, 150 mM NaCl, 5 mM BME at 850 MHz and 298 K.

Assignment of H3:H4 ILV-methyl resonances. Assignment of H3 and H4 ILV-methyl groups was done similarly to Rtt109. Unambiguous assignment was possible due to a larger

number of unique combinations of residues that are close to each other in space, NMR data on ILV-methyl labelled H4, which allowed to differentiate the H4 residues from the H3 and comparison with the previously reported assignment of H3 and H4 in the context of the nucleosome (Kato *et al*, 2011). In total, I assigned 5/7 I, 4/12 L and 1/6 V residues of H3 and 6/6 I, 1/8 L and 4/9 V residues of H4.
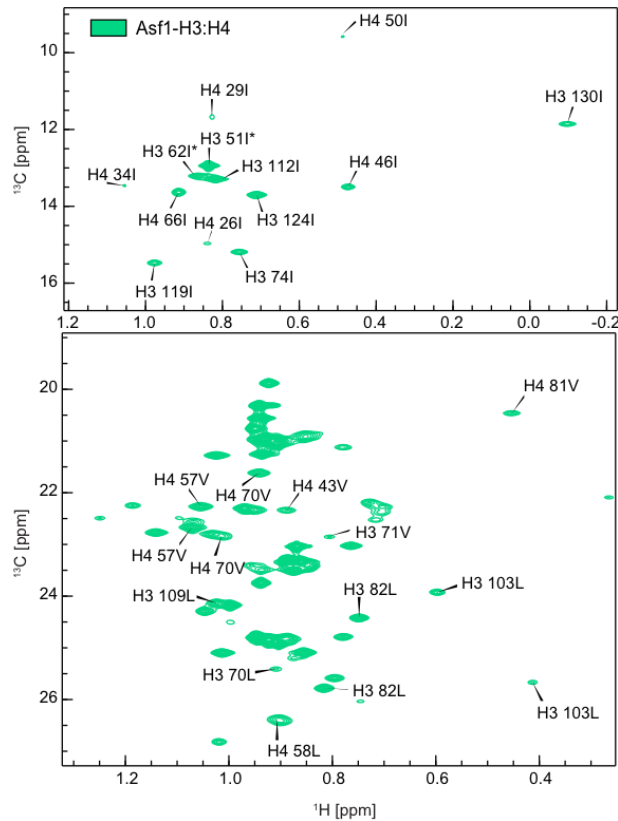
Figure 4.7.4. Assignment of Vps75 ILV-methyl resonances. ^{1}H-^{13}C HMQC spectrum of ILV-labelled H3:H4 as part of the Asf1–H3:H4 complex (0.66 mM) in 20 mM Tris– HCl pH 7.5, 1 M NaCl, 1 mM Na-EDTA, 5 mM BME in 100% D2O at 850 MHz and 298 K. The asterisk marks tentative assignments. The figure is adapted from Danilenko *et al*, 2019 and was originally produced by myself.

5. Discussion

5.1. Asf1 and Vps75 activate Rtt109 via substrate presentation mechanisms.

The structure of the RVAH complex presented in this work rationalizes the *in vivo* activity data, showing that both Asf1 and Vps75 are important for H3 K9 acetylation, but only Asf1 is essential for H3 K56 acetylation (Fig. 1.3.2, 1.4.2). The proposed mechanism of how the chaperones regulate Rtt109-dependent acetylation of H3 is summarized in Fig. 5.1.1.
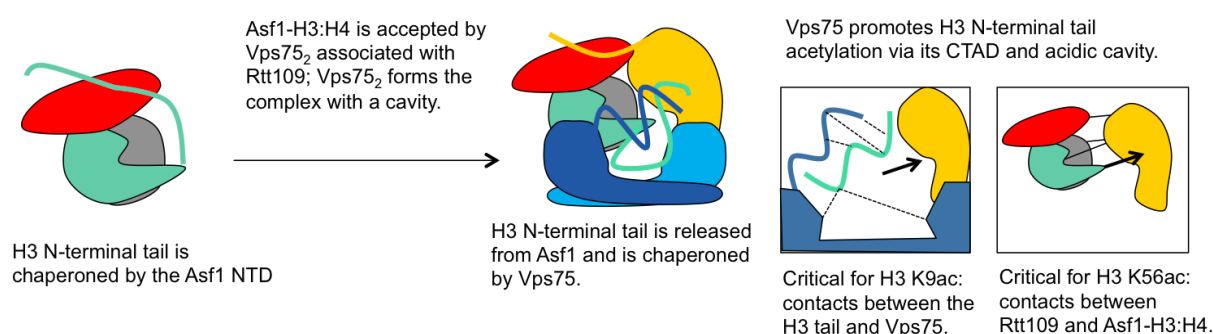
Figure 5.1.1. Regulation of Rtt109 activity by Asf1 and Vps75. Asf1 presents the H3:H4 dimer to Rtt109 in a correct orientation, where the H3 N-terminal tail faces towards the active center and H3 K56 is placed in the vicinity of the catalytic pocket. Vps75 facilitates the formation of the complex and attracts the disordered H3 N-terminal tail to the Rtt109 active center via both its folded and unstructured domains.

Mutations of H3 E94, which affect the formation of the RVAH complex, as shown by NMR data in Fig. 4.5.3, decrease both H3 K56 and K9 acetylation *in vivo* (Zhang et al, 2018). Mutations that affect local interactions of Rtt109 and the Asf1-bound histones, such as deletion of the H4 C-terminus (95-102) or the Rtt109 C-terminal tail, decrease the acetylation of H3 K56 but don't affect, or affect less significantly acetylation of H3 K9 by the RVAH complex *in vitro* (Fig. 4.5.4). On the other hand, deletion of the Vps75 CTAD decreases the acetylation of K9, but not K56 (Fig. 4.6.4). These results, taken together with the structure of the RVAH complex, are consistent with a model in which H3 K56 acetylation is promoted by Asf1, acting to position the H3 core with respect to Rtt109 in such a way that K56 faces the Rtt109 active site; Vop75 aids the formation of a stable complex, but does not exert an active function in catalysis. On the other hand, acetylation of H3 K9 is critically dependent on Vps75: the H3 N-terminal tail is first released from Asf1 by both Rtt109 and Vps75, and then is confined near the Rtt109 active center by the concerted action of both the acidic cavity

and the CTAD of Vps75. This is shown by the CSP data on the H3 N-terminal tail and the Vps75 CTAD as well as by the low resolution SANS envelopes (Fig. 4.5.7 and section 4.6).

The $^{206}EE^{207}$ acidic patch located in the Vps75 cavity further aids H3 N-terminal tail acetylation by capturing R52 and R53 of H3 and mildly inhibiting H3 K56 acetylation, as demonstrated by molecular dynamics simulations and *in vitro* activity assays (Fig. 4.6.3). The H3 $^{52}RR^{53}$ is the only double arginine patch in the H3 N-terminal tail, despite the fact that all of the H3 lysine residues that can be acetylated have an arginine residue in the vicinity. However, H3 $^{52}RR^{53}$ might be engaged by the Vps75 cavity more efficiently due to both its ability to make a higher number of hydrogen bonds and its proximity the $^{206}EE^{207}$. It is important to note that although Vps75 $^{206}EE^{207}$ has a mild inhibitory effect on H3 K56ac, the overall acetylation of H3 K56 is not inhibited but rather promoted by the presence of Vps75 *in vitro* (Fig. 4.3.4). This is most likely due to the role of Vps75 in the formation of the tight complex between Rtt109 and Asf1–H3:H4.

In vitro, CSP experiments and SEC did not detect a tight complex between Rtt109 and Asf1–H3:H4 in the absence of Vps75 (Fig. 4.2.6). However, Rtt109–Asf1–H3:H4 can acetylate H3 K56 both *in vitro* and *in vivo*. One explanation of this is that a transient short-lived interaction between Rtt109 and Asf1-bound histones is sufficient for the acetylation of H3 K56; on the other hand, efficient acetylation of the disordered N-terminal tail, which is not actively guided towards the active center in the absence of Vps75, requires a longer lifetime of the complex. In addition to this, the Rtt109–Asf1–H3:H4 stability and activity *in vivo* might be influenced by such factors as PTMs of H3 and the long CTAD of Asf1.

Some Rtt109 homologs lack the ability to bind to their corresponding Vps75 (Chen *et al*, 2019b; Zhang *et al*, 2018). The conservation of the residues involved in the interaction between the Asf1–H3:H4 and Rtt109 in *A. fumigatus* and *S. cerevisiae* together with the role of characterized Rtt109 homologs in the DNA damage response, suggest that the mechanism of H3 K56 acetylation by Rtt109 is conserved throughout fungi. The degree of the H3 N-terminal tail acetylation by Rtt109 homologs as well as the role of Vps75 homologs appears to be more variable.

While the structural model of the RVAH complex explains the preference of Rtt109 for the acetylation of H3, and how Vps75 promotes acetylation of the H3 N-terminal tail, it does not explain the selectivity for specific lysine residues in the H3 tail. It appears that the selectivity

of Rtt109 towards H3 K9, H3 K14[9], K23, K27 and K56, but not K18, K36 or K37 is determined by the structure of the active center of the acetyltransferase rather than by histone chaperones. All lysine residues targeted by Rtt109 have a similar surrounding amino-acid sequences (Fig. 5.1.2A).
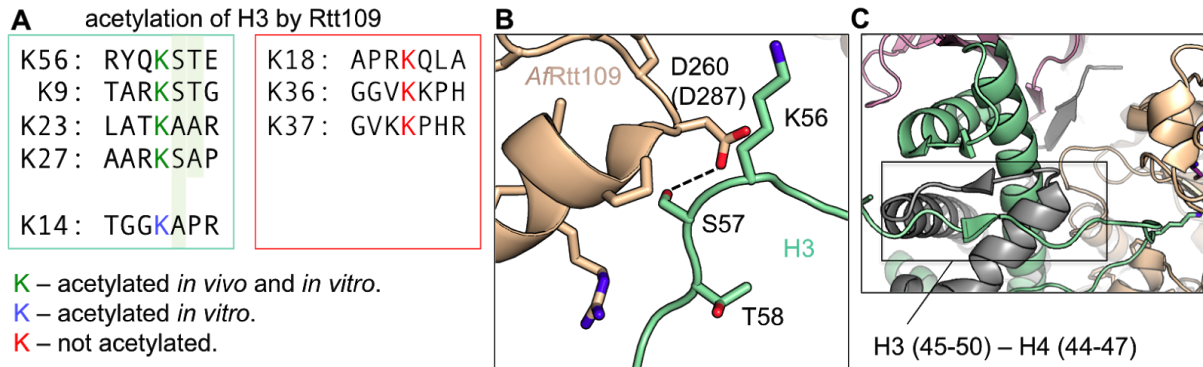
Figure 5.1.2. Residues following H3 lysines determine the acetylation by Rtt109. A. Lysines which are acetylated by Rtt109 are followed by small residues (similar residues are highlighted by a green background). **B,C**. Particular views of the *A. fumigatus* Rtt109–Asf1–H3:H4 structure (PDB entry 5zba). Rtt109 is shown in beige, Asf1 in pink, H3 in green and H4 in grey.

As an example of how the neighbouring amino acids influence lysine acetylation, the H3 S57Q mutation abolishes K56 *in vivo* without affecting K9, K23 or K27 (Zhang *et al*, 2018). In the X-ray structure of the *A. fumigatus* Rtt109–Asf1–H3:H4 from Zhang *et al*, H3 S57 is engaged in a hydrogen bond with D260 (corresponding to D287 in *S. cerevisiae*), which could explain its importance (Fig. 5.1.2B). However, H3 K23 and K14, which can be acetylated by Rtt109, are followed by alanines, which are similar to serines in size but cannot form hydrogen bonds with their side chain. This, together with the effect of the mutation of a small S to a large Q, suggests that the size of the neighbouring amino acids is the defining factor. From the crystal structure it is unclear why a large residue could not be accommodated in this position.

In addition to the placement of H3 K56 in the catalytic pocket, the Zhang *et al* structure shows that a part of the H3 tail (residues ~45-51) is wrapped around the H3:H4 dimer via interactions with H4 (residues 44-47) mediated mostly by the backbone atoms (Fig. 5.1.2C). Interestingly, our CSP data indicates that this H4 surface is affected by the presence of the Vps75 CTAD (Fig. 4.5.7). While I hypothesize that these CSPs arise from a direct interaction between the Vps75 CTAD and H3:H4, they may also be caused by the modulation of the interaction of the H3 tail with H3:H4.

---

[9] Rtt109 is able to acetylate H3 K14 *in vitro*, however *in vivo* Gcn5 is the major HAT responsible for the H3 K14 acetylation. It is possible that H3 K14ac is introduced by Gcn5 prior to the H3 acetylation by Rtt109.

It would be interesting to obtain a more detailed molecular view of the H3 N-terminal tail and its interactions within the complex, including the residues surrounding the catalytic pocket of Rtt109, the Vps75 CTAD and the complex cavity.

5.2. Vps75 dimer accepts one copy of Rtt109 to form a functional complex.

This work addresses the long-standing debate on the stoichiometry of the Rtt109–Vps75 complex, which originated from the detection of 1:2 and 2:2 complexes by X-ray crystallography (Tang *et al*, 2011; Su *et al*, 2011). The NMR titration experiments and MALS data of section 4.1 show that the Rtt109$_2$–Vps75$_2$ complex is detectable in solution *in vitro* only in certain buffer conditions and is less stable than the Rtt109–Vps75$_2$. Although conformational flexibility of the Vps75 dimer can allow the formation of the 2:2 complex, from our functional and structural data it is unclear which functions Rtt109$_2$–Vps75$_2$ could perform *in vivo*, if this complex is present in the cell.

As it can be seen from the CSP and SEC-MALS data (Fig. 4.2.1, Fig. 4.2.2), formation of the complex of Rtt109–Vps75 with the substrate Asf1–H3:H4 requires a free earmuff domain of one of the Vps75 monomers. It is unlikely that Rtt109$_2$–Vps75$_2$ could accept Asf1–H3:H4 as a substrate, as the surfaces of Rtt109 and Vps75 that interact with the histones in the RVAH complex are largely inaccessible in the Rtt109$_2$–Vps75$_2$ complex. Thus, if Rtt109$_2$–Vps75$_2$ has an acetylation function, it would most likely acetylate H3 in a non-Asf1-bound state. In disagreement with this hypothesis, it has been demonstrated that Asf1 is necessary for the Rtt109-dependent acetylation of both K9 and K56 *in vivo* (Fig. 1.4.2). Thus, it is unlikely that Rtt109$_2$–Vps75$_2$ takes part in H3 acetylation at all.

The interaction of the Rtt109–Vps75$_2$ with a second Rtt109 copy could potentially play a role in a storage of Rtt109, which is not stable *in vivo* in the absence of Vps75. However, expression of Rtt109 peaks during the S-phase, while Vps75 is expressed constantly; hence, it is more conceivable that Vps75 is stored as a self-tetramer outside of the S-phase, during which the requirement of the newly synthesized and acetylated H3 is high (Gown *et al*, 1996; Bowman *et al*, 2014).

In summary, data of this thesis work suggest that acetylation of H3 *in vivo* is performed by Rtt109 bound to Vps75 in a 1:2 stoichiometry, while the 2:2 complex may have other cellular functions or be an artifact observed *in vitro*.

5.3. Asf1 regulates the accessibility of the H3 N-terminal tail.

The high conservation of both the H3 N-terminal tail and the Asf1 surface it engages suggests that the mechanism by which Asf1 chaperones the N-terminal tail might be conserved throughout eukaryotes. In support of this hypothesis, the *Xenopus laevis* H3 N-terminal tail interacts with the *H. sapiens* Asf1b in a similar manner as with *Sc* Asf1, as seen in NMR titration experiments (Lercher *et al*, 2018). The NMR experiments in sections 4.3 and 4.6 show that this binding of the H3 N-terminal tail to Asf1 can be disrupted by different mechanisms. One is a competition for the same Asf1 surface by other binding partners. CSP data show that the C-terminal tail of Rtt109 is able to displace the H3 N-terminal tail; it is conceivable that other Asf1 binding partners, such as the HIRA and CAF-1 complexes, utilize the same mechanism *in vivo.* Alternatively, the interaction between Asf1 and the H3 N-terminal tail can be disrupted by binding partners of the tail itself; our CSP data show that Vps75 can engage different regions of the H3 N-terminal tail by its CTAD and its folded acidic cavity, thereby releasing the H3 tail from Asf1.

In vivo data demonstrates that both Rtt109 C-terminal tail and Vps75 are required for efficient H3 K9 acetylation (Radovani *et al*, 2013); *in vitro*, however, Vps75 can fully rescue the C-terminally truncated Rtt109 (Fig. 4.3.4). PTMs of H3 could be one of the factors which contribute to this difference and add additional complexity to the acetylation regulation *in vivo*. In agreement with this hypothesis, it was recently shown that H3 carrying the K14ac mark promotes the acetylation of H3 K56 over K9 by the RVAH complex *in vitro* (Cote *et al*, 2019). Sensing of the acetylation state of K14 is mediated by Asf1 E105, as demonstrated by the Asf1 E105A mutant. In addition to the E105 mutation, mutations of residues located on the Asf1 surface that is involved in binding of either Rtt109[419–433] or the H3 N-terminal tail (D37, E39, D58), slightly alter the selectivity of Rtt109 acetylation *in vitro* (Cote *et al*, 2019). In particular, some of these mutations decrease the preference for H3 K56ac promoted by the presence of H3 K14ac. The structural data presented in section 4.3 and Lercher *et al* are consistent with the hypothesis that the underlying mechanism of this alteration is the change in the H3 N-terminal tail-Asf1 interaction: while H3 K14ac might contribute to stabilization of the Asf1-bound state of the H3 N-terminal tail, mutations of the negatively charged Asf1 residues engaged with H3 ~19-32 could weaken this interaction, rendering the H3 tail more accessible for acetylation.

The chaperoning of the H3 N-terminal tail by Asf1 could play a role in contexts other than Rtt109-dependent acetylation of H3. Asf1 is central to various pathways that include H3:H4, such as replication, transcription and DNA-damage response. Thus, the Asf1-H3 N-terminal tail interaction, which can be regulated by protein partners as well as H3 PTMs, could provide additional links in the histone biogenesis and recycling pathways.

5.4. Structurally disordered protein-protein interactions.

The majority of structurally characterised protein-protein interactions involve either binding of folded domains to each other, or folding of unstructured regions upon binding. Only recently the term "fuzzy" interactions was introduced to describe the interactions where disorder is preserved upon binding (Tompa & Fuxreiter, 2008). In the electrostatically-driven interaction of the H3 N-terminal tail with the Vps75 CTAD, both binding partners remain disordered (judging from the secondary structure prediction, Fig. 4.5.6C, and the preservation of a narrow chemical shift dispersion for both Vps75 CTAD and H3 tail in the RVAH complex). This binding mode is advantageous for the role of Vps75 CTAD in H3 tail acetylation, as it provides fast dissociation rates and accessibility of all N-terminal lysines of H3 to Rtt109; furthermore, this interaction mode is not associated with a large loss of conformational entropy which would occur in a folding-upon-binding mechanism.

Histone chaperones from diverse families, such as NAP, FACT, NPM, ASF and CAF, contain intrinsically disordered acidic regions (acidic IDRs) (Warren & Shechter, 2017). For several of them, including Spt16 (FACT), Chz1, SWR1, Npm1 and Nap1, the Vps75 homolog in *S. cerevisiae*, it was shown that these acidic IDRs shield the DNA-binding surfaces of the H2A:H2B, thus preventing nonspecific interactions and aggregation (Kemble *et al*, 2015; Warren *et al*, 2017). The sequence motif of these IDRs includes a variable negatively charged patch with one or more aromatic residues, typically tyrosines, which aid the "anchoring" to the histone surface. The disordered domains of MCM2 and Spt2, that wrap the H3:H4, display similar features (Chen *et al*, 2015; Huang *et al*, 2015). On the other hand, Vps75 CTAD engages the H3:H4 DNA-binding surface, despite the fact that its sequence lacks a tyrosine or any other aromatic residue.

While the sequence of the Vps75 CTAD is not conserved in fungi, fungal Vps75 CTADs share an overall similar negative charge (Fig. 1.3.1), which further suggests that the interaction between the Vps75 CTAD and H3 N-terminal tail relies on sequence-nonspecific electrostatic contacts based on the charge complementarity between the two disordered

regions. Recently, other examples of fuzzy electrostatic interactions between disordered domains have been reported. In the complex of chaperone ProTα and H1, the complex is highly mobile but the interaction between the two proteins is tight (pM to sub-µM) (Borgia *et al*, 2018; Feng *et al*, 2018); this aids the role of ProTα in chromatin remodelling, allowing it to access the DNA-bound H1 and compete for the interaction. In the intra-molecular interaction of two oppositely charged sequences within the C-terminal tail of the histone chaperone Npm1 (Warren *et al*, 2017), the self-chaperoning of the tail regulates the accessibility of the acidic patch that engages the H2A:H2B DNA-binding surface, as discussed above. In both cases, the interaction is based on charge complementarity, and the complexes are highly dynamic, as shown by the models generated with FRET, PRE, NMR and SAS data and molecular dynamics simulations. Thus, this interaction mode could be an as yet over-looked mechanism by which IDRs interact with each other.

Our understanding of the function of IDRs in protein-protein interactions, and their specific mode of action remains poor despite these regions being widespread in eukaryotes, where 25-30% of all proteins are predicted to be mostly disordered and ~40% are expected to contain disordered sequences longer than 40 amino acids (Dunker *et al*, 2000; Oldfield *et al*, 2005). IDRs are frequently identified in transcription factors and are reported to have a role in multiple diseases (Minezaki *et al*, 2006; Anbo *et al*, 2019). Overall, this understudied interaction mode would have wide-ranging implications for our understanding of eukaryotic protein-protein interactions.

Appendix 1. SANS data

The SANS data obtained during the thesis work. The highlighted entries contain the samples used for 1) the scoring of the structural models in M3 (blue, the labelling schemes are shown in the Fig. 4.4.6); 2) the selection of the docking blocks (red); 3) for the controls and for the generation of the *ab initio* envelopes shown in the Fig. 4.6.4 (green). These datasets were deposited to SASDBD, the accession numbers are noted below the sample name.

The proteins are denoted by the first letters of their names (R - Rtt109; V - Vps75$_2$, VΔ225 - Vps75$_2$$^{1-225}$; A - Asf1; H - H3:H4, HΔ35 - H3$^{35-135}$:H4). The concentration (conc.*) is the concentration of the scattering subunits, which is lower or equal to the concentration of the full RVAH complex. "+" and "-" signs refer to the scattering contrast in case of the different, non-zero, contrasts of the two parts of the complex.

June 2016, ILL, D22.

sample	scattering subunits	conc.*, mg/ml	Theor. I(0), cm^{-1}	Exp. I(0) SANS, cm^{-1}	R$_g$ Guinier, Å	R$_g$ GNOM, Å
all ^{1}H, 0% D$_2$O	RVAH	6.08	0.5	0.42	34.92	34.6
70%-^{2}H AV, ^{1}H RH, 100% D$_2$O	RH	1.99	0.15	0.15	27.57	27.55
70%-^{2}H AR, ^{1}H VH 100% D$_2$O	VH	1.5	0.12	0.14	29.59	29.34
^{1}H AH, ^{2}H RV, 0% D$_2$O	RV++ AH+	2.92 & 1.16	n.d.	n.d.	n.d.	n.d.
^{1}H AH, ^{2}H RV, 42% D$_2$O	RV	2.96	0.53	0.56	33.95	33.83
^{1}H AH, ^{2}H RV, 100% D$_2$O	RV+ AH-	3 & 1.2	n.d.	n.d.	n.d.	n.d.
^{1}H AH, ^{2}H RV, 42% D$_2$O	RV	1.12	0.2	0.22	34.47	34.53
^{1}H RAH, ^{2}H V, 42% D$_2$O	V	2.22	0.22	0.20	29.79	29.84
^{1}H VAH, ^{2}H R, 42% D$_2$O	R	1.62	0.13	0.16	25.87	25.98
^{1}H RH, ^{2}H AV, 42% D$_2$O	AV	1.9	0.25	0.23	37.23	37.54
^{1}H VH, ^{2}H AR, 42% D$_2$O (SASDFQ7)	AR	1.345	0.15	0.17	30.82	30.88
all ^{1}H, 100% D$_2$O (SASDFP7)	RVAH	2.73	0.42	0.38	34.96	34.53

November 2016, ILL, D22.

sample	scattering subunits	conc.*, mg/ml	Theor. I(0), cm^{-1}	Exp. I(0) SANS, cm^{-1}	R$_g$ Guinier, Å	R$_g$ GNOM, Å
all ^{1}H, 0% D$_2$O	all	3.75	0.31	0.26	35	34.83
70%-^{2}H RV, ^{1}H AH, 100% D$_2$O (SASDFK7)	AH	0.73	0.029	0.02	19.43	20.65
70%-^{2}H V, ^{1}H RAH, 100% D$_2$O	RAH	2.45	0.22	0.17	27.9	28.05
^{2}H RVΔ225, ^{1}H AH, 42% D$_2$O (SASDFL7)	RVΔ225	2.7	0.55	0.54	33.65	33.24
^{2}H VΔ225, ^{1}H RAH, 42% D$_2$O (SASDFM7)	VΔ225	1.75	0.15	0.083	27.41	27.81
^{2}H R, ^{1}H VΔ225AH, 42% D$_2$O	R	1.15	0.094	0.12	26.17	26.23

March 2017, MLZ, KWS-1.

sample	scattering subunits	conc.*, mg/ml	Theor. I(0), cm⁻¹	Exp. I(0) SANS, cm⁻¹	R_g Guinier, Å	R_g GNOM, Å
all ¹H RVΔ225AHΔ35, 100% D₂O (SASDFL3)	RVΔ225AHΔ35	2.32	0.33	0.31	35.26	35.32
70%-²H V, ¹H RAHΔ35, 100% D₂O	RAHΔ35	2.35	0.22	0.18	29.8	29.5
70%-²H R, ¹H VΔ225AHΔ35, 100% D₂O (SASDFN3)	VΔ225AHΔ35	2.15	0.2	0.16	32.82	32.86
²H RVΔ225, ¹H AHΔ35, 100% D₂O (SASDFM3)	RVΔ225+ AHΔ35-	-	n.d.	n.d.	n.d.	n.d.
all ¹H RVAH 100% D₂O	RVAH	2.2	0.34	0.38	34.8	34.2
²H AR, ¹H VΔ225HΔ35, 42% D₂O (SASDFR3)	AR	2.54	0.29	0.29	30.60	31.03
²H AVΔ225, ¹H RHΔ35, 42% D₂O (SASDFQ3)	VΔ225A	2.4	0.28	0.25	34.99	36.02

April 2018, ILL, D22.

sample	scattering subunits	conc.*, mg/ml	Theor. I(0), cm⁻¹	Exp. I(0) SANS, cm⁻¹	R_g Guinier, Å	R_g GNOM, Å
70%-²H RV, ¹H AH, 100% D₂O	AH	0.73	0.029	0.02	19.7	20.1
70%-²H V, ¹H RAH, 100% D₂O	RAH	2.45	0.22	0.17	27.9	28.0
¹H AH, 100% D₂O (not in the RVAH complex, aggregation)	AH	0.41	0.018	0.048	29.2	~30
70%-²H A, ¹H H, 100% D₂O (not in the RVAH complex, aggregation)	H	1.28	0.028	0.042	~30	n.d.
70%-²H RVA, ¹H H, 100% D₂O	H	1.83	0.045	n.d.	n.d.	n.d.

May 2018, ILL, D22.

sample	scattering subunits	conc.*, mg/ml	Theor. I(0), cm⁻¹	Exp. I(0) SANS, cm⁻¹	R_g Guinier, Å	R_g GNOM, Å
all ¹H RVΔ225AH, 100% D₂O (SASDFN7)	all	1.92	0.28	0.27	35.01	34.31
70%-²H VΔ225, ¹H RAHΔ35, 100% D₂O (SASDFP3)	RAHΔ35	3.13	0.25	0.26	28.35	28.40
70%-²H VΔ225, ¹H RAH, 100% D₂O	RAH	2.51	0.21	0.23	29.08	29.25
70%-²H R, ¹H VΔ225AH, 100% D₂O	VΔ225AH	1.40	0.12	0.10	33.26	32.95

Appendix 2. Restraints used for the structure calculation by M3

The table of unambiguous restraints used in M3 during both it0 and it1. The building block A is presented by 10 Rtt109–Vps75$_2$ structures, the building block B is the Asf1–H3:H4 structure (PDB entry 2hue). The residues within the building blocks were renumbered to avoid duplicating numbers. In the building block A (segid A) Vps75(B) numbering is preserved; the Vps75(A) and Rtt109 are renumbered by addition of 1000 or 2000, respectively, to the original residue number. In the building block B (segid B) Asf1 numbering is preserved, H3 and H4 are renumbered by addition of 1000 or 2000, respectively.
The distances are in Å, the distance restraint boundaries range from [distance1 + distance3] to [distance1 - distance2].

```
assign (resid 1063 and name CB and segid B) ((resid  159 or resid 1159) and name CD2 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid   66 and name  CG1 and segid A) 27.2 7.2 7.9
assign (resid 1063 and name CB and segid B) (resid   91 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 1146 and name  CD1 and segid A) 25.4 6.5 6.7
assign (resid 1063 and name CB and segid B) (resid   60 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 1018 and name  CD1 and segid A) 25.1 9.6 73.9
assign (resid 1063 and name CB and segid B) (resid 1105 and name  CD1 and segid A) 21.5 8 13
assign (resid 1063 and name CB and segid B) (resid  105 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 1091 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 1208 and name  CD1 and segid A) 19.2 6.1 6.1
assign (resid 1063 and name CB and segid B) (resid  208 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid   80 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid   86 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid  146 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) ((resid  159 or resid 1159) and name CD2 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid   66 and name  CG1 and segid A) 29.8 6.6 6.6
assign (resid 2030 and name CB and segid B) (resid   91 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid 1146 and name  CD1 and segid A) 25.0 6.6 6.6
assign (resid 2030 and name CB and segid B) (resid   60 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid 1018 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid 1105 and name  CD1 and segid A) 22.7 8.4 20.6
assign (resid 2030 and name CB and segid B) (resid  105 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid 1066 and name  CG1 and segid A) 19.2 7.5 9.4
assign (resid 2030 and name CB and segid B) (resid   18 and name  CD1 and segid A) 18.0 16.0 8.0
assign (resid 2030 and name CB and segid B) (resid 1091 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid 1208 and name  CD1 and segid A) 19.2 6.1 6.2
assign (resid 2030 and name CB and segid B) (resid  208 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid   80 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid   86 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid  146 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) ((resid  159 or resid 1159) and name CD2 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid   66 and name  CG1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid   91 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid 1146 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid   60 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid 1018 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid 1105 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid  105 and name  CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid 1091 and name  CD1 and segid A) 99.0 75.0 0.0
```

assign (resid 2045 and name CB and segid B) (resid 1208 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid 208 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid   80 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid   86 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid 146 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) ((resid 159 or resid 1159) and name CD2 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid   66 and name CG1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid   91 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid   60 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid 1018 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid 1066 and name CG1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid   18 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid 1091 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid 1208 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid 208 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid   80 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid   86 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 71 and name N and segid B) (resid 2401 and name N and segid A) 50.0 45.0 10.0
assign (resid 1076 and name CB and segid B) (resid 1179 and name CD1 and segid A) 18.0 16.0 8.0
assign (resid 1076 and name CB and segid B) (resid 1155 and name CD1 and segid A) 21.5 6.1 6.1
assign (resid 1076 and name CB and segid B) (resid 1160 and name CD1 and segid A) 24.7 8.1 11.5
assign (resid 1076 and name CB and segid B) (resid 1159 and name CD2 and segid A) 25.6 6.5 6.6
assign (resid 1076 and name CB and segid B) (resid 1112 and name CD1 and segid A) 18.0 16.0 8.0
assign (resid 1076 and name CB and segid B) (resid 179 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 155 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 160 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 159 and name CD2 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 112 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid   66 and name CG1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid   91 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 1146 and name CD1 and segid A) 21.9 6.5 6.5
assign (resid 1076 and name CB and segid B) (resid   60 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 1105 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 105 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 1066 and name CG1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 1091 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 1208 and name CD1 and segid A) 24.2 6.1 6.1
assign (resid 1076 and name CB and segid B) (resid 208 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid   80 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid   86 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 146 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1179 and name CD1 and segid A) 23.6 8.4 19.2
assign (resid 2082 and name CB and segid B) (resid 1155 and name CD1 and segid A) 22.2 6.1 6.1
assign (resid 2082 and name CB and segid B) ((resid 1160 and name CD1 and segid A) 22.6 9.4 76.4
assign (resid 2082 and name CB and segid B) (resid 1159 and name CD2 and segid A) 26.1 6.7 6.7
assign (resid 2082 and name CB and segid B) (resid 1112 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 179 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 155 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 160 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 159 and name CD2 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 112 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   66 and name CG1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   91 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1146 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   60 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1018 and name CD1 and segid A) 26.3 10.3 72.7
assign (resid 2082 and name CB and segid B) (resid 1105 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 105 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1066 and name CG1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   18 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1091 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1208 and name CD1 and segid A) 99.0 75.0 0.0

assign (resid 2082 and name CB and segid B) (resid 208 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   80 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   86 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 146 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2294 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2282 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2332 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2083 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2189 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2383 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2138 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2150 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1115 and name CB and segid B) (resid 2365 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2294 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2282 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2332 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2083 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2189 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2383 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2138 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2150 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2098 and name CB and segid B) (resid 2365 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 160 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 1051 and name CD1 and segid A) 21.3 6.5 6.7
assign (resid 1063 and name CB and segid B) (resid   51 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid   88 and name CG2 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 1160 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1063 and name CB and segid B) (resid 1088 and name CG2 and segid A) 21.5 6.9 7.4
assign (resid 2030 and name CB and segid B) (resid 160 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid 1028 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2030 and name CB and segid B) (resid   28 and name CD1 and segid A) 27.6 7 7.2
assign (resid 2030 and name CB and segid B) (resid 1160 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid   88 and name CG2 and segid A) 99.0 75.0 0.0
assign (resid 1076 and name CB and segid B) (resid 1088 and name CG2 and segid A) 25.6 6.7 6.9
assign (resid 2082 and name CB and segid B) (resid 1051 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   51 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1028 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid   28 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 1127 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2082 and name CB and segid B) (resid 127 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid   88 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2093 and name CB and segid B) (resid 1088 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid   88 and name CD1 and segid A) 99.0 75.0 0.0
assign (resid 2045 and name CB and segid B) (resid 1088 and name CD1 and segid A) 99.0 75.0 0.0

Appendix 3. Assignment of Vps75 and H3:H4

Stereospecific assignment of ILV methyl groups of Vps75 in the RVAH complex. A - Vps75(A) chain, earmuff domain of which is histone-bound; B - Vps75(B) chain, earmuff domain of which is Rtt109-bound; (A/B) - overlapping peaks of chain A and B. "?" mark is a tentative assignment.

^{1}H [ppm]	^{13}C [ppm]	Assign F2
0.95251	24.66024	(B/A)16LeuCd1
0.43426	23.65727	(B/A)61ValCg2
0.19384	22.22572	(A/B)62LeuCd2
0.74305	20.37418	(B/A)121ValCg2
0.70339	21.42008	(A/B)124ValCg1
0.72507	20.74785	(B/A)124ValCg2
0.87300	13.49436	(B/A)138IleCd1
0.70123	21.90810	(B/A)159LeuCd2
0.87561	25.01719	(B/A)188LeuCd1
0.93975	21.01286	(B/A)248?ValCg1
0.93339	20.04826	(B/A)248?ValCg2
0.93860	24.83885	(B/A)252?LeuCd1
0.87912	23.23720	(B/A)252?LeuCd2
0.90658	21.37492	(B/A)264?ValCg1
0.91196	23.29527	(A/B)16LeuCd2
0.81836	15.12098	(A/B)28IleCd1
1.07466	22.35342	(B/A)32ValCg1
1.16875	25.74875	(A/B)34?LeuCd1
0.89056	22.43115	(B/A)37?LeuCd2
1.04995	22.48007	(B/A)43?ValCg2
0.25734	13.29965	(A/B)54IleCd1
0.99314	25.48788	(B/A)139LeuCd2
1.11493	20.96941	(A/B)144ValCg1
0.71579	24.72911	(B/A)203LeuCd1
1.04568	22.66821	(B/A)213ValCg1
1.26256	24.63934	(B/A)213ValCg2
0.88733	19.94058	(B/A)264ValCg2
0.60690	25.27302	A18LeuCd1
1.09801	21.61963	A25ValCg1
1.12841	23.48400	A25?ValCg2
0.55629	14.80446	A51IleCd1
1.00146	20.86706	A61ValCg1
0.84516	20.46630	A66ValCg1
0.22931	13.23995	A86IleCd1
0.05469	21.92776	A88ValCd1
0.36984	25.17554	A91LeuCd1
0.86694	23.63753	A93LeuCd2
0.33835	15.49443	A105IleCd1
0.86657	14.05696	A112IleCd1
0.60593	9.16099	A127IleCd1
1.03431	24.09131	A139LeuCd1
1.14008	20.57025	A144ValCg2
0.04974	14.83515	A146IleCd1
0.26747	13.20621	A155?IleCd1
0.75798	14.07606	A160IleCd1
0.43908	12.86914	A179IleCd1
0.15756	22.38777	A200?LeuCd1
−0.6902	24.25797	A200LeuCd2
0.60946	24.55174	A203?LeuCd2
−0.3276	12.01661	A208IleCd1
0.65622	25.45629	B18LeuCd1
1.11499	21.67880	B25ValCg1
1.17568	23.76685	B25?ValCg2

^{1}H [ppm]	^{13}C [ppm]	Assign F2
1.07188	23.12258	B32?ValCg2
0.54462	14.85447	B51IleCd1
0.87900	12.96913	B60IleCd1
0.98173	20.68705	B61ValCg1
0.80899	20.62492	B66ValCg1
0.96938	21.93043	B66ValCg2
0.72549	11.29902	B80IleCd1
0.87075	13.80027	B83IleCd1
−0.0853	13.42130	B86IleCd1
0.02742	21.95458	B88ValCd1
0.37111	24.96615	B91LeuCd1
0.87336	23.69502	B93LeuCd2
0.18105	15.31761	B105IleCd1
0.90254	14.15549	B112IleCd1
0.57757	9.04679	B127IleCd1
1.04278	24.03391	B139LeuCd1
1.12700	20.57557	B144ValCg2
0.02024	14.76122	B146IleCd1
0.26962	13.02733	B155?IleCd1
0.76818	14.12425	B160IleCd1
0.42289	12.91804	B179IleCd1
0.12064	22.42099	B200LeuCd1
−0.6754	24.26260	B200LeuCd2
0.62212	24.20834	B203?LeuCd2
−0.4709	11.88199	B208IleCd1

Assignment of ILV methyl groups of Vps75$^{1\text{-}225}$ in isolation. "?" mark is a tentative assignment.

^{1}H [ppm]	^{13}C [ppm]	Assign F2	^{1}H [ppm]	^{13}C [ppm]	Assign F2
0.92842	24.58618	16Leu	0.05118	14.76039	146Ile
0.89998	23.38926	16Leu	0.21259	12.92549	155Ile
0.86937	24.85122	18Leu	0.82796	25.07487	159Leu
0.59483	25.13647	18Leu	0.67989	21.92252	159Leu
1.06122	23.19897	25Val	0.73577	14.10635	160Ile
1.04076	21.39152	25Val	0.40012	12.80682	179Ile
0.75434	14.83594	28Ile	0.85233	25.03366	188Leu
1.05111	23.09161	32Val	0.75173	22.32119	188Leu
1.05512	22.28279	32Val	0.13915	22.45782	200Leu
0.87725	22.52973	34?Leu	−0.7266	24.24102	200Leu
1.12814	25.68637	34?Leu	0.78380	24.14534	203Leu
0.78279	24.99963	37?Leu	0.59299	24.96654	203Leu
0.85621	21.62714	43Val	−0.3732	11.75181	208Ile
0.99770	22.46987	43Val	1.19431	24.60222	213Val
0.50017	14.63216	51Ile	1.02232	22.05602	213Val
0.31555	13.49786	54Ile	0.86741	24.84929	223Leu
0.87510	12.78369	60Ile	0.82178	23.03831	223Leu
0.41567	23.58521	61Val			
0.96431	20.71454	61Val			
0.80997	26.00676	62Leu			
0.11890	21.70937	62Leu			
0.94149	21.07054	66Val			
0.85975	20.56127	66Val			
0.09642	12.79849	72Ile			
0.59894	10.75062	80Ile			
0.82495	13.80162	83Ile			
0.08969	13.03123	86Ile			
0.03780	21.85495	88Val			
0.32013	25.36116	91Leu			
0.42055	21.39843	91Leu			
0.85447	23.68318	93?Leu			
0.92876	24.98051	93?Leu			
0.40461	15.21084	105Ile			
0.87660	14.11835	112Ile			
0.70881	20.29633	121Val			
0.72568	22.22405	121Val			
0.67021	21.38500	124Val			
0.70125	20.83037	124Val			
0.61775	9.54750	127Ile			
0.85418	13.47993	138Ile			
1.02445	24.19484	139Leu			
0.97674	25.28369	139Leu			
1.07033	20.91671	144Val			
1.10272	20.68396	144Val			

Assignment of ILV methyl groups of H3:H4 in the Asf1-H3:H4 complex. "?" mark is a tentative assignment. "A" marks the methyl groups of H3, "B" - of H4.

¹H[ppm]	¹³C[ppm]	Assign F1
0.82863	11.30804	B29Ile[45]
1.04607	13.32840	B34Ile[89]
0.46452	13.65557	B46Ile[55]
0.84690	12.82277	A51?Ile[121]
0.98044	15.33118	A119Ile[20]
0.50085	9.49224	B50Ile[1]
0.83126	15.06161	B26Ile[73]
1.06364	22.38630	B57Val[117]
1.07553	22.71801	B57Val[24]
0.91417	26.41810	B58Leu[228]
0.88004	13.24425	A62?Ile[71]
0.93803	13.83382	A66Ile[113]
0.93057	25.39729	A70Leu[107]
0.78508	22.88351	A71Val[75]
0.74726	15.15428	A74Ile[129]
0.59670	24.03360	A103Leu[27]
0.41416	25.72395	A103Leu[29]
0.81938	13.32056	A112Ile[99]
-0.1073	11.86058	A130Ile[79]
0.77195	24.40702	A82Leu[147]
0.83721	25.92663	A82Leu[137]
1.03759	24.26377	A109Leu[159]
0.71799	13.67572	A124Ile[18]
0.89656	22.28807	B43Val[22]
0.72001	22.47330	B43Val[242]
0.95401	21.59490	B70Val[236]
1.03053	22.88484	B70Val[238]
0.44932	20.47260	B81Val[67]

Assignment of ILV methyl groups of H3:H4 in the RVAH complex. "?" mark is a tentative assignment.

¹H[ppm]	¹³C[ppm]	Assign F1
0.78685	14.76406	B26Ile[73]
0.44475	13.22723	B46Ile[55]
0.75180	12.91389	A51?Ile[230]
1.07056	22.35116	B57Val[117]
1.09938	22.89425	B57Val[24]
0.90704	26.58572	B58Leu[228]
0.85487	13.48125	A66Ile[113]
0.81917	22.81807	A71Val[75]
0.74971	15.14938	A74Ile[129]
0.62812	23.71233	A103Leu[27]
0.44070	25.92479	A103Leu[29]
0.82001	13.28678	A112Ile[99]
-0.2186	12.04984	A130Ile[231]
0.84697	25.82752	A82Leu[137]
0.78348	24.48627	A82Leu[147]
1.03121	23.98510	A109Leu[159]

0.99001	15.61656	A119Ile[20]
0.72544	13.84888	A124Ile[18]
0.91411	22.52512	B43Val[22]
0.71505	22.44941	B43Val[242]
0.50591	10.34802	B50Ile[1]
0.46643	20.36271	B81Val[67]

Assignment of Vps75[1-225] in isolation.

¹H[ppm]	¹⁵N[ppm]	¹³C[ppm]	Assign F1	Assign F2	Assign F3
8.35395	122.60356	174.49827	4AspH	4AspN	3SerC
8.35370	122.63640	54.53532	4AspH	4AspN	4AspCa
8.35358	122.60997	40.28025	4AspH	4AspN	4AspCb
8.18700	119.98388	176.47474	5GlnH	5GlnN	4AspC
8.19427	120.00191	55.82693	5GlnH	5GlnN	5GlnCa
8.19428	120.00150	28.62891	5GlnH	5GlnN	5GlnCb
8.32203	121.66094	176.36226	6GluH	6GluN	5GlnC
8.32256	121.67002	56.72429	6GluH	6GluN	6GluCa
8.32256	121.67075	29.23566	6GluH	6GluN	6GluCb
8.32229	119.30827	176.57390	7AsnH	7AsnN	6GluC
8.32856	119.32724	53.24708	7AsnH	7AsnN	7AsnCa
8.32853	119.32331	38.54847	7AsnH	7AsnN	7AsnCb
8.35904	121.69037	175.45103	8GluH	8GluN	7AsnC
8.35939	121.69750	56.74014	8GluH	8GluN	8GluCa
8.35906	121.69748	29.25769	8GluH	8GluN	8GluCb
8.36651	119.73006	176.54159	9AsnH	9AsnN	8GluC
8.36705	119.73310	53.32088	9AsnH	9AsnN	9AsnCa
8.36744	119.73099	38.40116	9AsnH	9AsnN	9AsnCb
8.59928	122.49364	176.05809	10GluH	10GluN	9AsnC
8.60854	122.51886	57.98023	10GluH	10GluN	10GluCa
8.60882	122.52035	28.73528	10GluH	10GluN	10GluCb
8.30554	119.46380	177.79480	11HisH	11HisN	10GluC
8.30851	119.47079	57.02090	11HisH	11HisN	11HisCa
8.30695	119.45688	29.32778	11HisH	11HisN	11HisCb
7.95448	122.53173	176.13787	12AlaH	12AlaN	11HisC
7.96507	122.55996	54.39055	12AlaH	12AlaN	12AlaCa
7.96541	122.56449	17.77703	12AlaH	12AlaN	12AlaCb
8.13562	118.54001	180.02809	13LysH	13LysN	12AlaC
8.13678	118.54426	58.92916	13LysH	13LysN	13LysCa
8.13756	118.54763	31.25152	13LysH	13LysN	13LysCb
8.09279	124.13121	179.22546	14AlaH	14AlaN	13LysC
8.09172	124.13038	54.53786	14AlaH	14AlaN	14AlaCa
8.09301	124.13178	17.32423	14AlaH	14AlaN	14AlaCb
8.11903	117.02769	179.16501	15PheH	15PheN	14AlaC
8.12674	117.03899	60.71253	15PheH	15PheN	15PheCa
8.12609	117.03709	37.52888	15PheH	15PheN	15PheCb
7.90029	121.43513	178.63294	16LeuH	16LeuN	15PheC
7.90057	121.43306	57.41823	16LeuH	16LeuN	16LeuCa
7.89954	121.44601	40.36715	16LeuH	16LeuN	16LeuCb
7.52673	108.44402	179.88640	17GlyH	17GlyN	16LeuC
7.52780	108.44952	45.74329	17GlyH	17GlyN	17GlyCa
7.88663	122.13916	176.71675	18LeuH	18LeuN	17GlyC

```
7.88832        122.14259 57.58239   18LeuH  18LeuN  18LeuCa
7.88615        122.15998 40.05304   18LeuH  18LeuN  18LeuCb
7.84940        122.58587 179.44397  19AlaH  19AlaN  18LeuC
7.84966        122.60592 54.71173   19AlaH  19AlaN  19AlaCa
7.84938        122.59621 16.84821   19AlaH  19AlaN  19AlaCb
7.75977        119.60322 181.11642  20LysH  20LysN  19AlaC
7.76209        119.59933 58.21282   20LysH  20LysN  20LysCa
7.76069        119.60496 30.48526   20LysH  20LysN  20LysCb
7.99429        118.60056 180.04133  21CysH  21CysN  20LysC
8.00924        118.59220 63.88797   21CysH  21CysN  21CysCa
8.00536        118.59725 25.91751   21CysH  21CysN  21CysCb
7.92899        119.74346 176.74942  22GluH  22GluN  21CysC
7.93767        119.75781 58.96136   22GluH  22GluN  22GluCa
7.93621        119.76848 28.69951   22GluH  22GluN  22GluCb
7.52816        119.63904 179.22918  23GluH  23GluN  22GluC
7.54080        119.63854 28.71228   23GluH  23GluN  23GluCb
7.54018        119.65780 58.78389   23GluH  23GluN  23GluCa
7.87037        121.78010 179.95169  24GluH  24GluN  23GluC
7.87735        121.80159 59.39022   24GluH  24GluN  24GluCa
7.87644        121.77114 28.69042   24GluH  24GluN  24GluCb
8.11696        120.86833 179.43359  25ValH  25ValN  24GluC
8.12834        120.88494 66.98732   25ValH  25ValN  25ValCa
7.67166        119.73603 177.79420  26AspH  26AspN  25ValC
7.67112        119.75520 57.33229   26AspH  26AspN  26AspCa
7.67423        119.76199 40.00098   26AspH  26AspN  26AspCb
7.63776        122.12124 179.09192  27AlaH  27AlaN  26AspC
7.64575        122.15890 54.75571   27AlaH  27AlaN  27AlaCa
7.64406        122.16374 17.13439   27AlaH  27AlaN  27AlaCb
8.69933        120.10104 59.46819   29GluH  29GluN  29GluCa
8.69918        120.12587 27.97331   29GluH  29GluN  29GluCb
7.46909        119.35586 180.02562  30ArgH  30ArgN  29GluC
7.47119        119.38174 59.18256   30ArgH  30ArgN  30ArgCa
7.47502        119.35606 28.72845   30ArgH  30ArgN  30ArgCb
7.93019        120.39838 178.00595  31GluH  31GluN  30ArgC
7.93780        120.39955 59.54647   31GluH  31GluN  31GluCa
7.93993        120.41759 28.87115   31GluH  31GluN  31GluCb
9.16279        121.87584 180.44903  32ValH  32ValN  31GluC
9.16279        121.87715 66.27237   32ValH  32ValN  32ValCa
9.16421        121.83937 30.54130   32ValH  32ValN  32ValCb
7.98732        120.81154 177.15588  33GluH  33GluN  32ValC
7.97939        120.77975 66.23371   33GluH  33GluN  33GluCa
7.98771        120.82967 28.52139   33GluH  33GluN  33GluCb
9.13516        119.27950 179.08484  48AspH  48AspN  47ArgC
9.13639        119.26357 56.07009   48AspH  48AspN  48AspCa
9.14294        119.24992 38.13894   48AspH  48AspN  48AspCb
7.15420        122.01146 177.62461  49AlaH  49AlaN  48AspC
7.16215        122.03063 54.27314   49AlaH  49AlaN  49AlaCa
7.15450        122.06588 16.77960   49AlaH  49AlaN  49AlaCb
7.24473        117.94621 181.06085  50TyrH  50TyrN  49AlaC
7.24636        117.95440 60.59295   50TyrH  50TyrN  50TyrCa
7.24626        117.97245 36.69622   50TyrH  50TyrN  50TyrCb
```

```
7.76398        120.66163 178.95904  51IleH  51IleN  50TyrC
7.77429        120.62945 65.60977   51IleH  51IleN  51IleCa
7.46788        117.05473 56.07646   52AspH  52AspN  52AspCa
7.46851        117.05534 40.18396   52AspH  52AspN  52AspCb
7.41018        116.50492 177.34971  53GluH  53GluN  52AspC
7.41027        116.53782 56.05751   53GluH  53GluN  53GluCa
7.41137        116.51965 29.45080   53GluH  53GluN  53GluCb
7.87599        122.92091 176.35626  54IleH  54IleN  53GluC
7.87642        122.92503 60.09820   54IleH  54IleN  54IleCa
7.87611        122.91005 36.77654   54IleH  54IleN  54IleCb
8.39986        131.81070 53.90712   55AlaH  55AlaN  55AlaCa
8.39896        131.80418 17.79159   55AlaH  55AlaN  55AlaCb
9.03160        114.28375 178.78751  56GluH  56GluN  55AlaC
9.02850        114.24436 57.03967   56GluH  56GluN  56GluCa
9.02984        114.25316 27.17864   56GluH  56GluN  56GluCb
7.43870        119.26275 175.08883  57PheH  57PheN  56GluC
7.43908        119.27574 62.63886   57PheH  57PheN  57PheCa
7.44074        119.27763 39.24428   57PheH  57PheN  57PheCb
8.96807        117.28183 176.43329  58TrpH  58TrpN  57PheC
8.97093        117.29165 59.97826   58TrpH  58TrpN  58TrpCa
8.96971        117.30671 27.84035   58TrpH  58TrpN  58TrpCb
7.96864        121.32231 178.03452  59LysH  59LysN  58TrpC
7.96865        121.33476 59.23945   59LysH  59LysN  59LysCa
7.96716        121.32765 31.15416   59LysH  59LysN  59LysCb
8.14267        121.11021 179.83563  60IleH  60IleN  59LysC
8.14923        121.16569 63.96170   60IleH  60IleN  60IleCa
8.14674        121.17360 36.22540   60IleH  60IleN  60IleCb
7.48793        120.03720 174.78658  64GlnH  64GlnN  63SerC
7.49179        120.01344 54.80670   64GlnH  64GlnN  64GlnCa
7.49148        120.00207 28.47833   64GlnH  64GlnN  64GlnCb
7.25989        120.94059 175.94673  65HisH  65HisN  64GlnC
7.26040        120.94434 58.79756   65HisH  65HisN  65HisCa
7.26090        120.94853 29.86138   65HisH  65HisN  65HisCb
7.67074        126.90381 176.76962  66ValH  66ValN  65HisC
7.67158        126.90423 65.25836   66ValH  66ValN  66ValCa
7.67168        126.89720 31.02804   66ValH  66ValN  66ValCb
8.29147        115.74991 177.36565  67SerH  67SerN  66ValC
8.29379        115.79056 63.48213   67SerH  67SerN  67SerCb
8.29452        115.79215 59.89988   67SerH  67SerN  67SerCa
9.59884        125.86020 176.57064  68PheH  68PheN  67SerC
9.59510        125.87167 61.94084   68PheH  68PheN  68PheCa
9.59302        125.89680 38.28202   68PheH  68PheN  68PheCb
7.69855        116.12300 175.96932  69AlaH  69AlaN  68PheC
7.69966        116.13561 53.82696   69AlaH  69AlaN  69AlaCa
7.70065        116.13316 17.52478   69AlaH  69AlaN  69AlaCb
8.85053        125.86902 174.07751  96GluH  96GluN  95SerC
8.85492        125.87992 57.20720   96GluH  96GluN  96GluCa
8.85635        125.89380 28.35314   96GluH  96GluN  96GluCb
7.92883        117.32253 176.09301  97MetH  97MetN  96GluC
7.92844        117.32351 56.23611   97MetH  97MetN  97MetCa
7.93397        117.34154 31.27464   97MetH  97MetN  97MetCb
```

```
7.14433      118.17948 176.58800  98TyrH  98TyrN  97MetC
7.14550      118.18666 58.38443   98TyrH  98TyrN  98TyrCa
7.14484      118.18357 37.84289   98TyrH  98TyrN  98TyrCb
8.84284      125.25164 175.22666  99AspH  99AspN  98TyrC
8.84386      125.25180 53.34691   99AspH  99AspN  99AspCa
8.84431      125.24913 43.60727   99AspH  99AspN  99AspCb
8.66235      124.55474 70.35814   106ThrH 106ThrN 106ThrCb
8.66204      124.55491 60.87055   106ThrH 106ThrN 106ThrCa
9.20975      123.77610 172.77916  107PheH 107PheN 106ThrC
9.21703      123.78148 55.54259   107PheH 107PheN 107PheCa
9.21539      123.77917 42.85162   107PheH 107PheN 107PheCb
8.24888      119.86026 175.25917  108HisH 108HisN 107PheC
8.25572      119.86891 55.47979   108HisH 108HisN 108HisCa
8.25714      119.86193 30.65000   108HisH 108HisN 108HisCb
8.74666      123.57319 173.33042  109PheH 109PheN 108HisC
8.75296      123.58553 55.75949   109PheH 109PheN 109PheCa
8.75231      123.58376 40.33321   109PheH 109PheN 109PheCb
9.50531      120.43360 176.89504  110HisH 110HisN 109PheC
9.51084      120.41178 56.21597   110HisH 110HisN 110HisCa
9.50954      120.41667 28.25164   110HisH 110HisN 110HisCb
8.10505      117.21270 173.20647  115AspH 115AspN 114GlyC
8.10880      117.22139 55.54532   115AspH 115AspN 115AspCa
9.05407      120.15514 174.03693  116PheH 116PheN 115AspC
9.05832      120.15468 54.04817   116PheH 116PheN 116PheCa
9.05714      120.15984 42.84742   116PheH 116PheN 116PheCb
8.71418      132.62956 173.89751  117LysH 117LysN 116PheC
8.72189      132.64245 56.12526   117LysH 117LysN 117LysCa
8.72162      132.64166 32.29410   117LysH 117LysN 117LysCb
7.91779      122.58563 174.77695  118GluH 118GluN 117LysC
7.91742      122.59312 55.87364   118GluH 118GluN 118GluCa
7.91776      122.59827 28.25307   118GluH 118GluN 118GluCb
7.96285      122.40506 176.64043  119GlnH 119GlnN 118GluC
7.97334      122.39256 55.02207   119GlnH 119GlnN 119GlnCa
7.96629      122.38073 33.13341   119GlnH 119GlnN 119GlnCb
8.60864      122.56813 172.90953  120GlnH 120GlnN 119GlnC
8.61582      122.57239 54.03726   120GlnH 120GlnN 120GlnCa
8.61582      122.57016 30.42738   120GlnH 120GlnN 120GlnCb
8.79848      121.98682 174.90585  121ValH 121ValN 120GlnC
8.80485      121.99785 59.89070   121ValH 121ValN 121ValCa
8.80585      121.99768 35.09991   121ValH 121ValN 121ValCb
8.66141      123.07873 173.94045  122ThrH 122ThrN 121ValC
8.67072      123.11225 71.03217   122ThrH 122ThrN 122ThrCb
8.66606      123.04763 61.30968   122ThrH 122ThrN 122ThrCa
9.27275      128.97590 172.31093  123LysH 123LysN 122ThrC
9.27875      128.97965 54.33897   123LysH 123LysN 123LysCa
9.27897      128.97831 36.34013   123LysH 123LysN 123LysCb
8.69356      126.95478 173.11699  124ValH 124ValN 123LysC
8.70114      126.94931 61.88147   124ValH 124ValN 124ValCa
8.70094      126.91375 32.27312   124ValH 124ValN 124ValCb
9.42410      131.92730 175.32266  125PheH 125PheN 124ValC
9.43266      131.97250 56.41692   125PheH 125PheN 125PheCa
```

9.43264	131.97112	40.22809	125PheH 125PheN 125PheCb
8.65658	117.70706	173.44618	126GlnH 126GlnN 125PheC
8.66345	117.70825	52.72943	126GlnH 126GlnN 126GlnCa
8.66286	117.71064	33.49310	126GlnH 126GlnN 126GlnCb
10.3644	129.07512	176.37131	127IleH 127IleN 126GlnC
10.3644	129.07512	58.56358	127IleH 127IleN 127IleCa
10.3644	129.07512	35.59607	127IleH 127IleN 127IleCb
9.16724	129.40948	177.50527	128LysH 128LysN 127IleC
9.15486	129.35323	54.77044	128LysH 128LysN 128LysCa
9.15497	129.34791	34.23701	128LysH 128LysN 128LysCb
8.53419	125.01918	175.71179	129LysH 129LysN 128LysC
8.53460	125.02515	55.73051	129LysH 129LysN 129LysCa
8.53477	125.02342	32.66546	129LysH 129LysN 129LysCb
8.51936	111.25903	176.90814	130GlyH 130GlyN 129LysC
8.51996	111.26543	44.47558	130GlyH 130GlyN 130GlyCa
8.38513	121.17810	174.12998	131LysH 131LysN 130GlyC
8.38536	121.17978	56.50625	131LysH 131LysN 131LysCa
8.38700	121.18642	32.07670	131LysH 131LysN 131LysCb
8.34563	119.72344	176.77523	132AspH 132AspN 131LysC
8.34549	119.73292	54.29567	132AspH 132AspN 132AspCa
8.34555	119.73175	40.46366	132AspH 132AspN 132AspCb
8.13798	119.40923	175.99462	133AspH 133AspN 132AspC
8.14566	119.44174	54.59711	133AspH 133AspN 133AspCa
8.14516	119.44105	40.03627	133AspH 133AspN 133AspCb
8.21757	118.90389	176.34267	134GlnH 134GlnN 133AspC
8.21824	118.91077	55.60126	134GlnH 134GlnN 134GlnCa
8.21825	118.91074	28.40227	134GlnH 134GlnN 134GlnCb
8.12675	121.11834	176.40322	135GluH 135GluN 134GlnC
8.12737	121.12561	56.41248	135GluH 135GluN 135GluCa
8.12730	121.12459	29.64130	135GluH 135GluN 135GluCb
8.30950	120.27123	176.15174	136AspH 136AspN 135GluC
8.31017	120.27973	54.25578	136AspH 136AspN 136AspCa
8.31018	120.27784	41.01096	136AspH 136AspN 136AspCb
8.25203	108.44265	176.82222	137GlyH 137GlyN 136AspC
8.25254	108.44992	44.90162	137GlyH 137GlyN 137GlyCa
8.12374	116.87201	172.97055	138IleH 138IleN 137GlyC
8.12131	116.87059	59.11825	138IleH 138IleN 138IleCa
8.12165	116.87068	40.37399	138IleH 138IleN 138IleCb
8.71317	125.42676	174.92612	139LeuH 139LeuN 138IleC
8.71509	125.47916	55.56671	139LeuH 139LeuN 139LeuCa
8.71363	125.48921	42.67233	139LeuH 139LeuN 139LeuCb
8.42174	113.55472	175.99132	140ThrH 140ThrN 139LeuC
8.42954	113.56175	59.66690	140ThrH 140ThrN 140ThrCa
8.43100	113.56421	73.84115	140ThrH 140ThrN 140ThrCb
8.46206	112.44140	173.38499	141SerH 141SerN 140ThrC
8.46983	112.44647	63.12052	141SerH 141SerN 141SerCb
8.47003	112.44549	55.07172	141SerH 141SerN 141SerCa
7.63850	120.73718	173.35317	142GluH 142GluN 141SerC
7.64714	120.73471	52.70593	142GluH 142GluN 142GluCa
7.64623	120.74004	30.08161	142GluH 142GluN 142GluCb
9.54017	128.75143	178.99412	150GlnH 150GlnN 149ProC

9.54020	128.75674	58.27313	150GlnH	150GlnN	150GlnCa
9.54073	128.76163	27.74661	150GlnH	150GlnN	150GlnCb
8.78783	114.22606	179.70284	151SerH	151SerN	150GlnC
8.78924	114.23804	62.42783	151SerH	151SerN	151SerCb
8.78623	114.22511	60.88331	151SerH	151SerN	151SerCa
7.67237	120.32192	175.39783	152TyrH	152TyrN	151SerC
7.67425	120.33994	53.78447	152TyrH	152TyrN	152TyrCa
7.67375	120.32204	36.04635	152TyrH	152TyrN	152TyrCb
7.87530	122.56308	176.05800	153AspH	153AspN	152TyrC
7.87984	122.56259	58.33186	153AspH	153AspN	153AspCa
7.88125	122.56632	39.31892	153AspH	153AspN	153AspCb
8.40988	111.03175	179.00313	154SerH	154SerN	153AspC
8.41163	111.05365	62.70594	154SerH	154SerN	154SerCb
8.40597	111.01515	60.22721	154SerH	154SerN	154SerCa
7.26850	111.58944	174.32174	155IleH	155IleN	154SerC
7.27104	111.58792	37.10056	155IleH	155IleN	155IleCb
7.58857	118.60053	176.49142	159LeuH	159LeuN	158AspC
7.59795	118.60678	53.82930	159LeuH	159LeuN	159LeuCa
7.60205	118.62764	42.04266	159LeuH	159LeuN	159LeuCb
6.91266	121.52420	177.38729	160IleH	160IleN	159LeuC
6.91414	121.49844	62.02435	160IleH	160IleN	160IleCa
6.91318	121.49393	38.45729	160IleH	160IleN	160IleCb
8.15977	116.56796	175.06007	164ArgH	164ArgN	163LysC
8.16771	116.58575	54.27715	164ArgH	164ArgN	164ArgCa
8.16649	116.57789	41.60907	164ArgH	164ArgN	164ArgCb
7.56759	117.87969	176.53364	165SerH	165SerN	164ArgC
7.56886	117.88743	56.41550	165SerH	165SerN	165SerCa
7.56825	117.88303	62.64962	165SerH	165SerN	165SerCb
8.72587	116.91337	179.95754	167GluH	167GluN	166ProC
8.72681	116.91465	58.90349	167GluH	167GluN	167GluCa
8.72673	116.91039	28.53358	167GluH	167GluN	167GluCb
8.13380	111.53851	179.08770	168GlyH	168GlyN	167GluC
8.13768	111.53059	46.71921	168GlyH	168GlyN	168GlyCa
8.74174	122.28006	176.34039	169LysH	169LysN	168GlyC
8.74200	122.28470	59.40066	169LysH	169LysN	169LysCa
8.74199	122.27456	31.69649	169LysH	169LysN	169LysCb
7.39809	119.89238	179.79391	170LysH	170LysN	169LysC
7.39935	119.89280	58.30785	170LysH	170LysN	170LysCa
7.39932	119.90133	28.25845	170LysH	170LysN	170LysCb
7.91207	120.35369	179.04717	171LysH	171LysN	170LysC
7.91373	120.36574	59.25665	171LysH	171LysN	171LysCa
7.91253	120.36378	31.31193	171LysH	171LysN	171LysCb
8.50069	120.16887	178.01244	172TyrH	172TyrN	171LysC
8.50149	120.16270	60.81107	172TyrH	172TyrN	172TyrCa
8.50269	120.17506	38.10226	172TyrH	172TyrN	172TyrCb
8.06497	118.73034	178.00155	173ArgH	173ArgN	172TyrC
8.06512	118.73055	58.99905	173ArgH	173ArgN	173ArgCa
8.06526	118.72665	29.42456	173ArgH	173ArgN	173ArgCb
7.96176	117.37595	179.15414	174GlnH	174GlnN	173ArgC
7.96176	117.38679	58.78175	174GlnH	174GlnN	174GlnCa
7.96206	117.36970	27.61207	174GlnH	174GlnN	174GlnCb

8.42057	107.23727	179.96317	175GlyH	175GlyN	174GlnC
8.42142	107.23904	47.42214	175GlyH	175GlyN	175GlyCa
8.12922	120.76916	177.64006	176MetH	176MetN	175GlyC
8.12873	120.77207	55.67549	176MetH	176MetN	176MetCa
8.12761	120.79843	30.79081	176MetH	176MetN	176MetCb
7.09478	114.12142	175.50188	177LysH	177LysN	176MetC
7.09444	114.12096	55.78481	177LysH	177LysN	177LysCa
7.09466	114.12457	31.82792	177LysH	177LysN	177LysCb
7.63178	109.47026	176.75727	178ThrH	178ThrN	177LysC
7.63676	109.49940	73.96488	178ThrH	178ThrN	178ThrCb
7.63931	109.50597	60.39343	178ThrH	178ThrN	178ThrCa
10.1599	121.59590	177.35891	179IleH	179IleN	178ThrC
10.1588	121.69450	36.64672	179IleH	179IleN	179IleCb
9.12814	117.94972	173.99483	198AspH	198AspN	197GlyC
9.13549	117.94023	54.71536	198AspH	198AspN	198AspCa
9.13339	117.94538	37.68500	198AspH	198AspN	198AspCb
7.48131	118.52078	177.34811	199SerH	199SerN	198AspC
7.47508	118.54665	61.63254	199SerH	199SerN	199SerCb
7.47413	118.55133	60.77990	199SerH	199SerN	199SerCa
7.55440	121.90286	176.96796	200LeuH	200LeuN	199SerC
7.56622	121.91559	56.79704	200LeuH	200LeuN	200LeuCa
7.56528	121.95290	39.46101	200LeuH	200LeuN	200LeuCb
7.16364	120.21842	176.79067	201AlaH	201AlaN	200LeuC
7.16487	120.22446	55.91042	201AlaH	201AlaN	201AlaCa
7.16381	120.24236	16.60208	201AlaH	201AlaN	201AlaCb
8.00061	112.91385	179.35477	202SerH	202SerN	201AlaC
8.00078	112.91927	62.03183	202SerH	202SerN	202SerCb
8.00109	112.93078	60.90765	202SerH	202SerN	202SerCa
7.14299	117.58059	176.08378	212CysH	212CysN	211PheC
7.14754	117.56894	59.93262	212CysH	212CysN	212CysCa
7.14110	117.57604	27.20628	212CysH	212CysN	212CysCb
8.23067	122.63834	175.28308	213ValH	213ValN	212CysC
8.22792	122.62153	67.20113	213ValH	213ValN	213ValCa
8.22820	122.62777	30.43913	213ValH	213ValN	213ValCb
7.43424	120.21359	177.53553	214LysH	214LysN	213ValC
7.43685	120.24013	58.87183	214LysH	214LysN	214LysCa
7.34565	117.93835	179.72729	215TyrH	215TyrN	214LysC
7.34615	117.93475	59.97160	215TyrH	215TyrN	215TyrCa
7.34311	117.92207	37.12276	215TyrH	215TyrN	215TyrCb
8.43796	121.51918	177.97859	216TyrH	216TyrN	215TyrC
8.43590	121.50976	61.79735	216TyrH	216TyrN	216TyrCa
8.43493	121.51215	38.23811	216TyrH	216TyrN	216TyrCb
8.50890	120.62347	177.27087	217AlaH	217AlaN	216TyrC
8.51459	120.62749	54.52109	217AlaH	217AlaN	217AlaCa
8.51031	120.62334	17.28465	217AlaH	217AlaN	217AlaCb
7.84133	118.40102	181.00820	218GluH	218GluN	217AlaC
7.84231	118.40558	58.26618	218GluH	218GluN	218GluCa
7.84224	118.41090	28.81598	218GluH	218GluN	218GluCb
7.44964	121.20542	178.14891	219AlaH	219AlaN	218GluC
7.45783	121.21848	53.31136	219AlaH	219AlaN	219AlaCa
7.45894	121.21482	17.93232	219AlaH	219AlaN	219AlaCb

7.51058	115.63344	178.99748	220GlnH 220GlnN 219AlaC
7.51874	115.64104	55.29031	220GlnH 220GlnN 220GlnCa
7.51843	115.64003	27.60355	220GlnH 220GlnN 220GlnCb
7.54574	120.61145	176.85826	221ArgH 221ArgN 220GlnC
7.54654	120.61753	56.78384	221ArgH 221ArgN 221ArgCa
7.54654	120.61568	29.63488	221ArgH 221ArgN 221ArgCb
8.04293	120.57688	176.57182	222AspH 222AspN 221ArgC
8.04342	120.58453	54.34379	222AspH 222AspN 222AspCa
8.04347	120.58312	40.39640	222AspH 222AspN 222AspCb
7.81917	121.81504	176.23294	223LeuH 223LeuN 222AspC
7.81982	121.81977	54.92165	223LeuH 223LeuN 223LeuCa
7.81933	121.82727	41.40303	223LeuH 223LeuN 223LeuCb
8.09520	121.67596	177.42992	224GluH 224GluN 223LeuC
8.09578	121.68587	55.95470	224GluH 224GluN 224GluCa
8.09575	121.68542	29.77901	224GluH 224GluN 224GluCb
7.87096	127.28196	175.41020	225AspH 225AspN 224GluC
7.87158	127.28878	41.66717	225AspH 225AspN 225AspCb

Backbone assignment of the H3 N-terminal tail in the free state.

^{1}H [ppm]	^{15}N [ppm]	Assign F2
8.35517	116.7957	3ThrN
8.4896	124.33355	4LysN
8.53959	122.38567	5GlnN
8.23802	116.21263	6ThrN
8.37246	126.9707	7AlaN
8.34483	121.06223	8ArgN
8.45208	123.2935	9LysN
8.48933	117.60216	10SerN
8.29587	115.56191	11ThrN
8.45544	111.15577	12GlyN
8.31997	108.99975	13GlyN
8.19697	120.94341	14LysN
8.32975	127.18873	15AlaN
8.45252	121.75446	17ArgN
8.41222	123.10462	18LysN
8.45602	122.17956	19GlnN
8.32994	124.19079	20LeuN
8.37081	125.18376	21AlaN
8.07971	113.77084	22ThrN
8.32352	123.80836	23LysN
8.2819	125.27925	24AlaN
8.23447	123.63566	25AlaN
8.27062	120.64398	26ArgN
8.32483	117.42028	28SerN
8.37129	126.93835	29AlaN
8.51173	124.79052	31AlaN
8.14674	112.88179	32ThrN
8.43083	111.09164	33GlyN
8.30326	108.96487	34GlyN
8.04177	119.61209	35ValN
8.44879	125.93765	36LysN

Backbone assignment of Vps75 CTAD in the free state.

^{1}H [ppm]	^{15}N [ppm]	Assign F2
8.24048	120.81722	225AspN
8.31828	121.27124	226GluN
8.43565	121.37159	227GluN
8.34357	109.43144	228GlyN
8.32743	120.64705	229GluN
8.47913	117.03491	230SerN
8.46077	110.67207	231GlyN
8.08678	121.5159	232LeuN
8.37335	116.83376	233SerN
8.44145	126.45876	234AlaN
8.30667	119.43977	235AspN
8.27642	109.12087	236GlyN
8.28909	120.49188	237AspN
8.33787	116.08453	238SerN
8.51191	122.62996	239GluN
8.34394	121.04188	240AspN
8.30462	120.96413	241AspN
8.3913	121.50661	242AspN
8.52193	109.66177	243GlyN
8.27504	115.94298	244SerN
8.31397	123.57117	245LeuN
8.30623	108.91082	246GlyN
8.23536	120.66655	247GluN
8.14851	120.40989	248ValN
8.40114	123.7516	249AspN
8.12447	123.84257	250LeuN
8.37913	122.46593	252LeuN
8.31201	116.14013	253SerN
8.41462	122.24296	254AspN
8.23226	120.18264	255GluN
8.3881	123.24803	256GluN
8.47038	115.20384	258SerN
8.26743	117.32857	259SerN
8.15801	122.5527	260LysN
8.159	121.86127	261LysN
8.30403	122.60495	262ArgN
8.4297	124.09551	263LysN
7.76971	125.85609	264ValN

Appendix 4. PCR primer sequences

Primers used to generate mutants for the assignment of the methyl groups of Vps75:

Vps_V25I_R: 5' CTCACGTTCAATAGCGTCGATTTCTTCTTCGCATTTAGC 3'
Vps_V25I_F: 5' GCTAAATGCGAAGAAGAAATCGACGCTATTGAACGTGAG 3'

Vps_V32I_F: 5' GACGCTATTGAACGTGAGATAGAATTATACAGATTGAACAAAATG 3'
Vps_V32I_R: CATTTTGTTCAATCTGTATAATTCTATCTCACGTTCAATAGCGTC 3'

Vps_V61I_F: 5' GAAATTGCAGAGTTCTGGAAAATTATCCTTTCACAACACGTTAGTTTTG 3'
Vps_V61I_R: 5' CAAAACTAACGTGTTGTGAAAGGATAATTTTCCAGAACTCTGCAATTTC 3'

Vps_V124I_F: 5' GAGCAGCAAGTCACCAAGATCTTCCAAATCAAAAAAGGCAAAG 3'
Vps_V124I_R: 5' CTTTGCCTTTTTTGATTTGGAAGATCTTGGTGACTTGCTGCTC 3'

Vps_L188I_F: 5' TGGTTCAGATGGACCGGCATAAAACCCGGGAAGGAGTTC 3'
Vps_L188I_R: 5' GAACTCCTTCCCGGGTTTTATGCCGGTCCATCTGAACCA 3'

Vps_V213I_F: 5' GGCCTCAGCATAGTACTTGATACAAAAAGGGTAGATCTCTTC 3'
Vps_V213I_R: 5' GAAGAGATCTACCCTTTTTGTATCAAGTACTATGCTGAGGCC 3'

Vps_I28V_F: 5' GAAGAAGTGGACGCTGTTGAACGTGAGGTAG 3'
Vps_I28V_R: 5' CTACCTCACGTTCAACAGCGTCCACTTCTTC 3'

Vps_I54V_F: 5' GCGTATATTGACGAAGTTGCAGAGTTCTGG 3'
Vps_I54V_R: 5' CCAGAACTCTGCAACTTCGTCAATATACGC 3'

Vps_I72V_F: 5' CACGTTAGTTTTGCGAATTACGTAAGGGCTTCAGATTTTAAATAC 3'
Vps_I72V_R: 5' GTATTTAAAATCTGAAGCCCTTACGTAATTCGCAAAACTAACGTG 3'

Vps_I127V_F: 5' GTCACCAAGGTGTTCCAAGTCAAAAAAGGCAAAGATG 3'
Vps_I127V_R: 5' CATCTTTGCCTTTTTTGACTTGGAACACCTTGGTGAC 3'

Primers used to generate mutants for the PRE measurements:

H3_R63C_F: 5' CCACCGAGCTGCTCATCTGCAAACTGCCTTTCCAG 3'
H3_R63C_R: 5' CTGGAAAGGCAGTTTGCAGATGAGCAGCTCGGTGG 3'

H3_Q76C_F: 5' CCTGGTCCGGGAGATCGCTTGTGACTTCAAGACCGACCTG 3'
H3_Q76C_R: 5' CAGGTCGGTCTTGAAGTCACAAGCGATCTCCCGGACCAGG 3'

H3_E94C_F: 5' GCCGTTATGGCTCTGCAGTGTGCCAGCGAGGCTTATCTG 3'
H3_E94C_R: 5' CAGATAAGCCTCGCTGGCACACTGCAGAGCCATAACGGC 3'

H3_C110A_F: 5' CTTTGAGGACACCAACCTGGCCGCCATCCACGCCAAGAG 3'
H3_C110A_R: 5' CTCTTGGCGTGGATGGCGGCCAGGTTGGTGTCCTCAAAG 3'

H4_Q93C_F: 5' GTTGTTTACGCTCTGAAACGTTGTGGTCGTACCCTGTACGGTTTC 3'
H4_Q93C_R: 5' GAAACCGTACAGGGTACGACCACAACGTTTCAGAGCGTAAACAAC 3'

H4_Y98C_F: 5' GTCAGGGTCGTACCCTGTGCGGTTTCGGTGGTTAAAG 3'
H4_Y98C_R: 5' CTTTAACCACCGAAACCGCACAGGGTACGACCCTGAC 3'

Plasmids containing the remaining histone mutants (H4 R45C, H4 T30C, H4 T82C, and H3 C110A K115C) were a kind gift from Prof. Dr. T. Owen-Hughes lab.

Primers used to generate mutants for validation of interaction interfaces and activity assays:

H3Δ28_F: 5' GTTTACTTTAAGAAGGAGATATACC ATG TCCGCTCCTGCTACCGGCGGA 3'
H3_ΔN _R: 5' GGGCTTTGTTAGCAGCCGGATTTAAGCCCTCTCGCCTCGGATTC 3'

H3Δ35_F: 5' GTTTACTTTAAGAAGGAGATATACC ATG GTCAAGAAACCTCACCGTTACCGG 3'
H3_ΔN_R: 5' GGGCTTTGTTAGCAGCCGGATTTAAGCCCTCTCGCCTCGGATTC 3'

Production of H3$^{29-135}$ and H3$^{35-135}$ was done in two steps. First, megaprimers (PCR products encoding a gene of interest) were generated using the above-stated pairs of primers and the pET3d plasmid containing wild-type H3 sequence. The resulting megaprimers then acted as a primer pair in a restriction-free (RF) cloning protocol.

H4Δ95_F: 5' CGCTCTGAAACGTCAGGGTTAAACCCTGTACGGTTTCGGTG 3'
H4Δ95_R: 5' CACCGAAACCGTACAGGGTTAACCCTGACGTTTCAGAGCG 3'

Rtt109Δ424_R: 5' TTGCTTTTTTACGCGGTTACAGCATGGTAATTGCCAG 3'
Rtt109Δ424_F: 5' CTGGCAATTACCATGCTGTAACCGCGTAAAAAAGCAA 3'

Vps75Δ225_F: 5' GAGGGACTTGGAAGACTAAGAAGGAGAATCCGG 3'
Vps75Δ225_R: 5' CCGGATTCTCCTTCTTAGTCTTCCAAGTCCCTC 3'

Vps_E206A_F: 5' GGCGTCATTGTTCAGTGCAGAGATCTACCCTTTTG 3'
Vps_E206A_R: 5' CAAAAGGGTAGATCTCTGCACTGAACAATGACGCC 3'

Vps_EE206/207AA_F: 5' GTCATTGTTCAGTGCAGCGATCTACCCTTTTGCTG 3'
Vps_EE206/207AA_R: 5' CAGCAAAAGGGTAGATCGCTGCACTGAACAATGAC 3'

H3_R69E_F: 5' CAAACTGCCTTTCCAGGAGCTGGTCCGGGAGATCG 3'
H3_R69E_R: 5' CGATCTCCCGGACCAGCTCCTGGAAAGGCAGTTTG 3'

H3_R69E/R72E_F: 5' CTTTCCAGGAGCTGGTCGAGGAGATCGCTCAGGAC 3'
H3_R69E/R72E_R: 5' GTCCTGAGCGATCTCCTCGACCAGCTCCTGGAAAG 3'

Acknowledgements

I am extremely grateful to my supervisor, Prof. Dr. Teresa Carlomagno, for the guidance during my PhD and thesis writing and for her sincere involvement in this research, for the insightful discussions, for her great ideas and tenacity. Thank you for giving me the opportunity to learn a lot from you and for the chance to have a taste of different techniques.

I would like to thank the members of my thesis advisory committee, Prof. Dr. Thomas Brüser, Prof. Dr. Wulf Blankenfeldt and Prof. Dr. Jens Boch, for taking an interest in this work and kindly agreeing to be a part of the thesis committee.

I had a great pleasure working with Dr. Lukas Lercher, who has initiated this project and who was a great mentor and partner in science during the first half of my PhD and a good friend throughout it. I am deeply grateful to Dr. John Kirkpatrick for a great amount of assistance with the NMR experiments, for patience, and for the development of the protocol for the distance derivation from PRE data. I am also very grateful to Dr. Luca Codutti for finding a solution for setting up MD simulations for my system and for introducing me to MD and MALS.

I would like to thank Dr. Frank Gabel for the assistance with SANS data collection and his input in the data analysis and interpretation. I want to also express my gratitude to Prof. Dr. Ezgi Karaca for the consultations on M3 protocol and for the user-friendly scripts deposited by her.

I want to thank all present and past members of the Carlomagno group for creating a nice working atmosphere and for sharing their expertise. In particular I want to thank Dr. Andrea Graziadei, whose help cannot be overestimated, for proofreading this manuscript and for providing a great deal of moral support throughout my studies. I am also thankful to Dr. Megha Karanth for her help with the proofreading, for the coffee breaks and for her friendship.

On a personal note, I would like to thank my family for their love, support and understanding during this time.

Bibliography

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindahl E (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2:** 19–25 Available at: http://dx.doi.org/10.1016/j.softx.2015.06.001

Ahmad K & Henikoff S (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9:** 1191–1200

Albaugh BN, Arnold KM, Lee S & Denu JM (2011) Autoacetylation of the histone acetyltransferase Rtt109. *J. Biol. Chem.* **286:** 24694–24701

Albaugh BN, Kolonko EM & Denu JM (2010) Kinetic mechanism of the Rtt109-Vps75 histone acetyltransferase-chaperone complex. *Biochemistry* **49:** 6375–6385

Alvarez F, Muñoz F, Schilcher P, Imhof A, Almouzni G & Loyola A (2011) Sequential establishment of marks on soluble histones H3 and H4. *J. Biol. Chem.* **286:** 17714–17721

Anbo H, Sato M, Okoshi A & Fukuchi S (2019) Functional Segments on Intrinsically Disordered Regions in Disease-Related Proteins. *Biomolecules* **9:** Available at: http://dx.doi.org/10.3390/biom9030088

Arents G & Moudrianakis EN (1995) The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proceedings of the National Academy of Sciences* **92:** 11170–11174 Available at: http://dx.doi.org/10.1073/pnas.92.24.11170

Battiste JL & Wagner G (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39:** 5355–5365

Bazan JF (2008) An old HAT in human p300/CBP and yeast Rtt109. *Cell Cycle* **7:** 1884–1886

Bennett CB, Lewis LK, Karthikeyan G, Lobachev KS, Jin YH, Sterling JF, Snipe JR & Resnick MA (2001) Genes required for ionizing radiation resistance in yeast. *Nat. Genet.* **29:** 426–434

Berndsen CE, Tsubota T, Lindner SE, Lee S, Holton JM, Kaufman PD, Keck JL & Denu JM (2008) Molecular functions of the histone acetyltransferase chaperone complex Rtt109–Vps75. *Nature Structural & Molecular Biology* **15:** 948–956 Available at: http://dx.doi.org/10.1038/nsmb.1459

Bonangelino CJ, Chavez EM & Bonifacino JS (2002) Genomic Screen for Vacuolar Protein Sorting Genes inSaccharomyces cerevisiae. *Molecular Biology of the Cell* **13:** 2486–2501 Available at: http://dx.doi.org/10.1091/mbc.02-01-0005

Borgia A, Borgia MB, Bugge K, Kissling VM, Heidarsson PO, Fernandes CB, Sottini A, Soranno A, Buholzer KJ, Nettels D, Kragelund BB, Best RB & Schuler B (2018) Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555:** 61–66

Bowman A, Hammond CM, Stirling A, Ward R, Shang W, El-Mkami H, Robinson DA, Svergun DI, Norman DG & Owen-Hughes T (2014) The histone chaperones Vps75 and Nap1 form ring-like, tetrameric structures in solution. *Nucleic Acids Res.* **42:** 6038–6051

Bowman A, Ward R, Wiechens N, Singh V, El-Mkami H, Norman DG & Owen-Hughes T (2011) The Histone Chaperones Nap1 and Vps75 Bind Histones H3 and H4 in a Tetrameric Conformation. *Molecular Cell* **41:** 398–408 Available at: http://dx.doi.org/10.1016/j.molcel.2011.01.025

Burgess RJ & Zhang Z (2013) Histone chaperones in nucleosome assembly and human disease. *Nat. Struct. Mol. Biol.* **20:** 14–22

Campos EI, Fillingham J, Li G, Zheng H, Voigt P, Kuo W-HW, Seepany H, Gao Z, Day LA, Greenblatt JF & Reinberg D (2010) The program for processing newly synthesized histones H3.1 and H4. *Nat. Struct. Mol. Biol.* **17:** 1343–1351

Chen F, Qi S, Zhang X, Wu J, Yang X & Wang R (2019a) lncRNA PLAC2 activated by H3K27 acetylation promotes cell proliferation and invasion via the activation of Wnt/β−catenin pathway in oral squamous cell carcinoma. *Int. J. Oncol.* **54:** 1183–1194

Chen S, Rufiange A, Huang H, Rajashankar KR, Nourani A & Patel DJ (2015) Structure–function studies of histone H3/H4 tetramer maintenance during transcription by chaperone Spt2. *Genes & Development* **29:** 1326–1340 Available at: http://dx.doi.org/10.1101/gad.261115.115

Chen Y, Zhang Y, Ye H, Dou Y, Lu D, Li X, Limper AH, Han J & Su D (2019b) Structural basis for the acetylation of histone H3K9 and H3K27 mediated by the histone chaperone Vps75 in Pneumocystis carinii. *Signal Transduction and Targeted Therapy* **4:** Available at: http://dx.doi.org/10.1038/s41392-019-0047-8

Clore GM & Iwahara J (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* **109:** 4108–4139

Cote JM, Kuo Y-M, Henry RA, Scherman H, Krzizike DD & Andrews AJ (2019) Two factor authentication: Asf1 mediates crosstalk between H3 K14 and K56 acetylation. *Nucleic Acids Res.* **47:** 7380–7391

Daganzo SM, Erzberger JP, Lam WM, Skordalakes E, Zhang R, Franco AA, Brill SJ, Adams PD, Berger JM & Kaufman PD (2003) Structure and function of the conserved core of histone deposition protein Asf1. *Curr. Biol.* **13:** 2148–2158

Dahlin JL, Kottom T, Han J, Zhou H, Walters MA, Zhang Z & Limper AH (2014) Pneumocystis jirovecii Rtt109, a novel drug target for Pneumocystis pneumonia in immunosuppressed humans. *Antimicrob. Agents Chemother.* **58:** 3650–3659

Dahlin JL, Nissink JWM, Strasser JM, Francis S, Higgins L, Zhou H, Zhang Z & Walters MA (2015) PAINS in the assay: chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *J. Med. Chem.* **58:** 2091–2113

Dahlin JL, Sinville R, Solberg J, Zhou H, Han J, Francis S, Strasser JM, John K, Hook DJ, Walters MA & Zhang Z (2013) A cell-free fluorometric high-throughput screen for

inhibitors of Rtt109-catalyzed histone acetylation. *PLoS One* **8:** e78877

Danilenko N, Lercher L, Kirkpatrick J, Gabel F, Codutti L & Carlomagno T (2019) Histone chaperone exploits intrinsic disorder to switch acetylation specificity. *Nat. Commun.* **10:** 3435

Das C & Tyler JK (2013) Histone exchange and histone modifications during transcription and aging. *Biochim. Biophys. Acta* **1819:** 332–342

Dennehey BK, Noone S, Liu WH, Smith L, Churchill MEA & Tyler JK (2013) The C Terminus of the Histone Chaperone Asf1 Cross-Links to Histone H3 in Yeast and Promotes Interaction with Histones H3 and H4. *Molecular and Cellular Biology* **33:** 605–621 Available at: http://dx.doi.org/10.1128/mcb.01053-12

Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N & Rando OJ (2007) Dynamics of replication-independent histone turnover in budding yeast. *Science* **315:** 1405–1408

Drogaris P, Villeneuve V, Pomiès C, Lee E-H, Bourdeau V, Bonneil E, Ferbeyre G, Verreault A & Thibault P (2012) Histone deacetylase inhibitors globally enhance h3/h4 tail acetylation without affecting h3 lysine 56 acetylation. *Sci. Rep.* **2:** 220

Dunker AK, Obradovic Z, Romero P, Garner EC & Brown CJ (2000) Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* **11:** 161–171

Dunne O, Weidenhaupt M, Callow P, Martel A, Moulin M, Perkins SJ, Haertlein M & Forsyth VT (2017) Matchout deuterium labelling of proteins for small-angle neutron scattering studies using prokaryotic and eukaryotic expression systems and high cell-density cultures. *European Biophysics Journal* **46:** 425–432 Available at: http://dx.doi.org/10.1007/s00249-016-1186-2

Eitoku M, Sato L, Senda T & Horikoshi M (2008) Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. *Cell. Mol. Life Sci.* **65:** 414–444

English CM, Adkins MW, Carson JJ, Churchill MEA & Tyler JK (2006) Structural basis for the histone chaperone activity of Asf1. *Cell* **127:** 495–508

Fan Z, Beresford PJ, Oh DY, Zhang D & Lieberman J (2003) Tumor suppressor NM23-H1 is a granzyme A-activated DNase during CTL-mediated apoptosis, and the nucleosome assembly protein SET is its inhibitor. *Cell* **112:** 659–672

Feigin LA & Svergun DI (1987) Structure Analysis by Small-Angle X-Ray and Neutron Scattering. Available at: http://dx.doi.org/10.1007/978-1-4757-6624-0

Feng H, Zhou B-R & Bai Y (2018) Binding Affinity and Function of the Extremely Disordered Protein Complex Containing Human Linker Histone H1.0 and Its Chaperone ProTα. *Biochemistry* **57:** 6645–6648

Fillingham J, Recht J, Silva AC, Suter B, Emili A, Stagljar I, Krogan NJ, Allis CD, Keogh M-C & Greenblatt JF (2008) Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. *Mol. Cell. Biol.* **28:** 4342–4353

Franco AA, Lam WM, Burgers PM & Kaufman PD (2005) Histone deposition protein Asf1 maintains DNA replisome integrity and interacts with replication factor C. *Genes Dev.*

19: 1365–1375

Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM & Svergun DI (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *Journal of Applied Crystallography* **50:** 1212–1225 Available at: http://dx.doi.org/10.1107/s1600576717007786

Franke D & Svergun DI (2009) DAMMIF, a program for rapidab-initioshape determination in small-angle scattering. *Journal of Applied Crystallography* **42:** 342–346 Available at: http://dx.doi.org/10.1107/s0021889809000338

Frielinghaus H, Feoktystov A, Berts I & Mangiapia G (2015) KWS-1: Small-angle scattering diffractometer. *Journal of large-scale research facilities JLSRF* **1:** Available at: http://dx.doi.org/10.17815/jlsrf-1-26

Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, Shabanowitz J, Mishra N, Strahl BD, David Allis C & Hunt DF (2007) Organismal Differences in Post-translational Modifications in Histones H3 and H4. *Journal of Biological Chemistry* **282:** 7641–7655 Available at: http://dx.doi.org/10.1074/jbc.m607900200

Gown AM, Jiang JJ, Matles H, Skelly M, Goodpaster T, Cass L, Reshatof M, Spaulding D & Coltrera MD (1996) Validation of the S-phase specificity of histone (H3) in situ hybridization in normal and malignant cells. *J. Histochem. Cytochem.* **44:** 221–226

Green EM, Antczak AJ, Bailey AO, Franco AA, Wu KJ, Yates JR 3rd & Kaufman PD (2005) Replication-independent histone deposition by the HIR complex and Asf1. *Curr. Biol.* **15:** 2044–2049

Grzesiek S & Bax A (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *Journal of Magnetic Resonance (1969)* **96:** 432–440 Available at: http://dx.doi.org/10.1016/0022-2364(92)90099-s

Guinier A (1939) La diffraction des rayons X aux très petits angles : application à l'étude de phénomènes ultramicroscopiques. *Annales de Physique* **11:** 161–237 Available at: http://dx.doi.org/10.1051/anphys/193911120161

Hainer SJ & Martens JA (2011) Identification of histone mutants that are defective for transcription-coupled nucleosome occupancy. *Mol. Cell. Biol.* **31:** 3557–3568

Hamiche A & Shuaib M (2013) Chaperoning the histone H3 family. *Biochim. Biophys. Acta* **1819:** 230–237

Hammond CM, Sundaramoorthy R, Larance M, Lamond A, Stevens MA, El-Mkami H, Norman DG & Owen-Hughes T (2016) The histone chaperone Vps75 forms multiple oligomeric assemblies capable of mediating exchange between histone H3–H4 tetramers and Asf1–H3–H4 complexes. *Nucleic Acids Research* **44:** 6157–6172 Available at: http://dx.doi.org/10.1093/nar/gkw209

Han J, Zhou H, Li Z, Xu R-M & Zhang Z (2007) Acetylation of lysine 56 of histone H3 catalyzed by RTT109 and regulated by ASF1 is required for replisome integrity. *J. Biol. Chem.* **282:** 28587–28596

Holliday R & Pugh J (1975) DNA modification mechanisms and gene activity during

development. *Science* **187:** 226–232 Available at: http://dx.doi.org/10.1126/science.1111098

Huang H, Strømme CB, Saredi G, Hödl M, Strandsby A, González-Aguilera C, Chen S, Groth A & Patel DJ (2015) A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. *Nat. Struct. Mol. Biol.* **22:** 618–626

Humphrey W, Dalke A & Schulten K (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14:** 33–8, 27–8

Ikura M, Kay LE & Bax A (1990a) A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* **29:** 4659–4667

Ikura M, Kay LE, Tschudin R & Bax A (1990b) Three-dimensional NOESY-HMQC spectroscopy of a 13C-labeled protein. *Journal of Magnetic Resonance (1969)* **86:** 204–209 Available at: http://dx.doi.org/10.1016/0022-2364(90)90227-z

Iwahara J, Schwieters CD & Marius Clore G (2004) Ensemble Approach for NMR Structure Refinement against1H Paramagnetic Relaxation Enhancement Data Arising from a Flexible Paramagnetic Group Attached to a Macromolecule. *Journal of the American Chemical Society* **126:** 5879–5896 Available at: http://dx.doi.org/10.1021/ja031580d

Jackson V (1988) Deposition of newly synthesized histones: hybrid nucleosomes are not tandemly arranged on daughter DNA strands. *Biochemistry* **27:** 2109–2120 Available at: http://dx.doi.org/10.1021/bi00406a044

Jacrot B (1976) The study of biological structures by neutron scattering from solution. *Reports on Progress in Physics* **39:** 911–953 Available at: http://dx.doi.org/10.1088/0034-4885/39/10/001

Jessulat M, Alamgir M, Salsali H, Greenblatt J, Xu J & Golshani A (2008) Interacting proteins Rtt109 and Vps75 affect the efficiency of non-homologous end-joining in Saccharomyces cerevisiae. *Arch. Biochem. Biophys.* **469:** 157–164

Jiao Y, Seeger K, Lautrette A, Gaubert A, Mousson F, Guerois R, Mann C & Ochsenbein F (2012) Surprising complexity of the Asf1 histone chaperone-Rad53 kinase interaction. *Proc. Natl. Acad. Sci. U. S. A.* **109:** 2866–2871

Johnson GJ (2005) Encyclopedia of Analytical Science (2nd edition)2005421Edited by Paul Worsfold, Alan Townshend and Colin Poole. Encyclopedia of Analytical Science (2nd edition). Amsterdam: Elsevier Academic Press 2005. 10 vols., ISBN: 0 12 764199 9 £2,950, $4,570. *Reference Reviews* **19:** 38–39 Available at: http://dx.doi.org/10.1108/09504120510632723

Jones PA & Liang G (2009) Rethinking how DNA methylation patterns are maintained. *Nat. Rev. Genet.* **10:** 805–811

Kang B, Pu M, Hu G, Wen W, Dong Z, Zhao K, Stillman B & Zhang Z (2011) Phosphorylation of H4 Ser 47 promotes HIRA-mediated nucleosome assembly. *Genes & Development* **25:** 1359–1364 Available at: http://dx.doi.org/10.1101/gad.2055511

Karaca E, Rodrigues JPGLM, Graziadei A, Bonvin AMJJ & Carlomagno T (2017) M3: an integrative framework for structure determination of molecular machines. *Nat. Methods*

14: 897–902

Katan-Khaykovich Y & Struhl K (2011) Splitting of H3-H4 tetramers at transcriptionally active genes undergoing dynamic histone exchange. *Proc. Natl. Acad. Sci. U. S. A.* **108:** 1296–1301

Kato H, van Ingen H, Zhou B-R, Feng H, Bustin M, Kay LE & Bai Y (2011) Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl. Acad. Sci. U. S. A.* **108:** 12283–12288

Keck KM & Pemberton LF (2011) Interaction with the histone chaperone Vps75 promotes nuclear localization and HAT activity of Rtt109 in vivo. *Traffic* **12:** 826–839

Kemble DJ, McCullough LL, Whitby FG, Formosa T & Hill CP (2015) FACT Disrupts Nucleosome Structure by Binding H2A-H2B with Conserved Peptide Motifs. *Mol. Cell* **60:** 294–306

Klinker H, Haas C, Harrer N, Becker PB & Mueller-Planitz F (2014) Rapid purification of recombinant histones. *PLoS One* **9:** e104029

Kolonko EM, Albaugh BN, Lindner SE, Chen Y, Satyshur KA, Arnold KM, Kaufman PD, Keck JL & Denu JM (2010) Catalytic activation of histone acetyltransferase Rtt109 by a histone chaperone. *Proc. Natl. Acad. Sci. U. S. A.* **107:** 20275–20280

Konarev PV, Volkov VV, Sokolova AV, Koch MHJ & Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* **36:** 1277–1282 Available at: http://dx.doi.org/10.1107/s0021889803012779

Korzhnev DM, Kloiber K, Kanelis V, Tugarinov V & Kay LE (2004) Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *J. Am. Chem. Soc.* **126:** 3964–3973

Kozak S, Lercher L, Karanth MN, Meijers R, Carlomagno T & Boivin S (2016) Optimization of protein samples for NMR using thermal shift assays. *J. Biomol. NMR* **64:** 281–289

Kristie TM (2016) Chromatin Modulation of Herpesvirus Lytic Gene Expression: Managing Nucleosome Density and Heterochromatic Histone Modifications. *MBio* **7:** e00098–16

Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrín-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, et al (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440:** 637–643

Kulaeva OI, Hsieh F-K, Chang H-W, Luse DS & Studitsky VM (2013) Mechanism of transcription through a nucleosome by RNA polymerase II. *Biochim. Biophys. Acta* **1829:** 76–83

Kuo Y-M, Henry RA, Huang L, Chen X, Stargell LA & Andrews AJ (2015) Utilizing targeted mass spectrometry to demonstrate Asf1-dependent increases in residue specificity for Rtt109-Vps75 mediated histone acetylation. *PLoS One* **10:** e0118516

Laskey RA, Mills AD, Philpott A, Leno GH, Dilworth SM & Dingwall C (1993) The role of nucleoplasmin in chromatin assembly and disassembly. *Philos. Trans. R. Soc. Lond. B

Biol. Sci. **339:** 263–9; discussion 268–9

Lercher L, Danilenko N, Kirkpatrick J & Carlomagno T (2018) Structural characterization of the Asf1–Rtt109 interaction and its role in histone acetylation. *Nucleic Acids Research* **46:** 2279–2289 Available at: http://dx.doi.org/10.1093/nar/gkx1283

Le S, Davis C, Konopka JB & Sternglanz R (1997) Two new S-phase-specific genes from Saccharomyces cerevisiae. *Yeast* **13:** 1029–1042

Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO & Shaw DE (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78:** 1950–1958

Li Q, Zhou H, Wurtele H, Davies B, Horazdovsky B, Verreault A & Zhang Z (2008) Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* **134:** 244–255

Li S & Shogren-Knaak MA (2009) The Gcn5 bromodomain of the SAGA complex facilitates cooperative and cross-tail acetylation of nucleosomes. *J. Biol. Chem.* **284:** 9411–9417

Liu WH & Churchill MEA (2012) Histone transfer among chaperones. *Biochem. Soc. Trans.* **40:** 357–363

Luger K, Mäder AW, Richmond RK, Sargent DF & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389:** 251–260

Luger K, Rechsteiner TJ & Richmond TJ (1999) Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* **304:** 3–19

Maas NL, Miller KM, DeFazio LG & Toczyski DP (2006) Cell Cycle and Checkpoint Regulation of Histone H3 K56 Acetylation by Hst3 and Hst4. *Molecular Cell* **23:** 109–119 Available at: http://dx.doi.org/10.1016/j.molcel.2006.06.006

Masumoto H, Hawke D, Kobayashi R & Verreault A (2005) A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* **436:** 294–298

Ma X-J, -J. Ma X, Wu J, Altheim BA, Schultz MC & Grunstein M (1998) Deposition-related sites K5/K12 in histone H4 are not required for nucleosome deposition in yeast. *Proceedings of the National Academy of Sciences* **95:** 6693–6698 Available at: http://dx.doi.org/10.1073/pnas.95.12.6693

McCullough L, Pham TH, Parnell TJ, Chandrasekharan MB, Stillman DJ & Formosa T (2008) FACT activity and histone H3-K56 acetylation promote optimal establishment of chromatin architecture independent of ongoing transcription in Saccharomyces cerevisiae. Available at: http://dx.doi.org/10.1101/383984

McCullough LL, Pham TH, Parnell TJ, Connell Z, Chandrasekharan MB, Stillman DJ & Formosa T (2019) Establishment and Maintenance of Chromatin Architecture Are Promoted Independently of Transcription by the Histone Chaperone FACT and H3-K56 Acetylation in Saccharomyces cerevisiae. *Genetics* **211:** 877–892 Available at: http://dx.doi.org/10.1534/genetics.118.301853

Mertens HDT & Svergun DI (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of Structural Biology* **172:** 128–141

Available at: http://dx.doi.org/10.1016/j.jsb.2010.06.012

Minard LV, Williams JS, Walker AC & Schultz MC (2011) Transcriptional Regulation by Asf1. *Journal of Biological Chemistry* **286:** 7082–7092 Available at: http://dx.doi.org/10.1074/jbc.m110.193813

Minezaki Y, Homma K, Kinjo AR & Nishikawa K (2006) Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *J. Mol. Biol.* **359:** 1137–1149

Mousson F, Ochsenbein F & Mann C (2007) The histone chaperone Asf1 at the crossroads of chromatin and DNA checkpoint pathways. *Chromosoma* **116:** 79–93

Mund M, Overbeck JH, Ullmann J & Sprangers R (2013) LEGO-NMR spectroscopy: a method to visualize individual subunits in large heteromeric complexes. *Angew. Chem. Int. Ed Engl.* **52:** 11401–11405

Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN & Dunker AK (2005) Comparing and combining predictors of mostly disordered proteins. *Biochemistry* **44:** 1989–2000

Ollerenshaw JE, Tugarinov V & Kay LE (2003) Methyl TROSY: explanation and experimental verification. *Magnetic Resonance in Chemistry* **41:** 843–852 Available at: http://dx.doi.org/10.1002/mrc.1256

Papamokos GV, Tziatzos G, Papageorgiou DG, Georgatos SD, Politou AS & Kaxiras E (2012) Structural role of RKS motifs in chromatin interactions: a molecular dynamics study of HP1 bound to a variably modified histone tail. *Biophys. J.* **102:** 1926–1933

Park Y-J & Luger K (2006) The structure of nucleosome assembly protein 1. *Proc. Natl. Acad. Sci. U. S. A.* **103:** 1248–1253

Park Y-J, Sudhoff KB, Andrews AJ, Stargell LA & Luger K (2008) Histone chaperone specificity in Rtt109 activation. *Nat. Struct. Mol. Biol.* **15:** 957–964

Parthun MR, Widom J & Gottschling DE (1996) The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell* **87:** 85–94

Pervushin K, Riek R, Wider G & Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **94:** 12366–12371

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera?A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25:** 1605–1612 Available at: http://dx.doi.org/10.1002/jcc.20084

Price DJ & Brooks CL 3rd (2004) A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.* **121:** 10096–10103

Radovani E, Cadorin M, Shams T, El-Rass S, Karsou AR, Kim H-S, Kurat CF, Keogh M-C, Greenblatt JF & Fillingham JS (2013) The Carboxyl Terminus of Rtt109 Functions in Chaperone Control of Histone Acetylation. *Eukaryotic Cell* **12:** 654–664 Available at:

http://dx.doi.org/10.1128/ec.00291-12

Rambo RP & Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496:** 477–481

Recht J, Tsubota T, Tanny JC, Diaz RL, Berger JM, Zhang X, Garcia BA, Shabanowitz J, Burlingame AL, Hunt DF, Kaufman PD & Allis CD (2006) Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proceedings of the National Academy of Sciences* **103:** 6988–6993 Available at: http://dx.doi.org/10.1073/pnas.0601676103

Reiter C, Heise F, Chung H-R & Ehrenhofer-Murray AE (2015) A link between Sas2-mediated H4 K16 acetylation, chromatin assembly in S-phase by CAF-I and Asf1, and nucleosome assembly by Spt6 during transcription. *FEMS Yeast Res.* **15:** Available at: http://dx.doi.org/10.1093/femsyr/fov073

Reverón-Gómez N, González-Aguilera C, Stewart-Morgan KR, Petryk N, Flury V, Graziano S, Johansen JV, Jakobsen JS, Alabert C & Groth A (2018) Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. *Mol. Cell* **72:** 239–249.e5

Rosa JL, Boyartchuk VL, Zhu LJ & Kaufman PD (2010) Histone acetyltransferase Rtt109 is required for Candida albicans pathogenesis. *Proceedings of the National Academy of Sciences* **107:** 1594–1599 Available at: http://dx.doi.org/10.1073/pnas.0912427107

Rosa JL, Bajaj V, Spoonamore J & Kaufman PD (2013) A small molecule inhibitor of fungal histone acetyltransferase Rtt109. *Bioorg. Med. Chem. Lett.* **23:** 2853–2859

Salzmann M, Pervushin K, Wider G, Senn H & Wuthrich K (1998) TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proceedings of the National Academy of Sciences* **95:** 13585–13590 Available at: http://dx.doi.org/10.1073/pnas.95.23.13585

Schneider J, Bajwa P, Johnson FC, Bhaumik SR & Shilatifard A (2006) Rtt109 is required for proper H3K56 acetylation: a chromatin mark associated with the elongating RNA polymerase II. *J. Biol. Chem.* **281:** 37270–37274

Schütz S & Sprangers R (2019) Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems. *Progress in Nuclear Magnetic Resonance Spectroscopy* Available at: http://dx.doi.org/10.1016/j.pnmrs.2019.09.004

Selth LA, Lorch Y, Ocampo-Hafalla MT, Mitter R, Shales M, Krogan NJ, Kornberg RD & Svejstrup JQ (2009) An rtt109-independent role for vps75 in transcription-associated nucleosome dynamics. *Mol. Cell. Biol.* **29:** 4220–4234

Selth L & Svejstrup JQ (2007) Vps75, a new yeast member of the NAP histone chaperone family. *J. Biol. Chem.* **282:** 12358–12362

Shang W-H, Hori T, Westhorpe FG, Godek KM, Toyoda A, Misu S, Monma N, Ikeo K, Carroll CW, Takami Y, Fujiyama A, Kimura H, Straight AF & Fukagawa T (2016) Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. *Nat. Commun.* **7:** 13465

Sharp JA, Fouts ET, Krawitz DC & Kaufman PD (2001) Yeast histone deposition protein

Asf1p requires Hir proteins and PCNA for heterochromatic silencing. *Curr. Biol.* **11:** 463–473

Shintomi K, Iwabuchi M, Saeki H, Ura K, Kishimoto T & Ohsumi K (2005) Nucleosome assembly protein-1 is a linker histone chaperone in Xenopus eggs. *Proc. Natl. Acad. Sci. U. S. A.* **102:** 8210–8215

Sobel RE, Cook RG, Perry CA, Annunziato AT & Allis CD (1995) Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc. Natl. Acad. Sci. U. S. A.* **92:** 1237–1241

Solomon I & Bloembergen N (1956) Nuclear Magnetic Interactions in the HF Molecule. *J. Chem. Phys.* **25:** 261–266

Stavropoulos P, Nagy V, Blobel G & Hoelz A (2008) Molecular basis for the autoregulation of the protein acetyl transferase Rtt109. *Proc. Natl. Acad. Sci. U. S. A.* **105:** 12236–12241

Su D, Hu Q, Zhou H, Thompson JR, Xu R-M, Zhang Z & Mer G (2011) Structure and Histone Binding Properties of the Vps75-Rtt109 Chaperone-Lysine Acetyltransferase Complex. *Journal of Biological Chemistry* **286:** 15625–15629 Available at: http://dx.doi.org/10.1074/jbc.c111.220715

Sugumar R, Adithavarman AP, Dakshinamoorthi A, David DC & Ragunath PK (2016) Virtual Screening of Phytochemicals to Novel Target (HAT) Rtt109 in Pneumocystis Jirovecii using Bioinformatics Tools. *J. Clin. Diagn. Res.* **10:** FC05–8

Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography* **25:** 495–503 Available at: http://dx.doi.org/10.1107/s0021889892001663

Svergun DI (1999) Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophysical Journal* **76:** 2879–2886 Available at: http://dx.doi.org/10.1016/s0006-3495(99)77443-6

Svergun DI, Richard S, Koch MH, Sayers Z, Kuprin S & Zaccai G (1998) Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. U. S. A.* **95:** 2267–2272

Tang Y, Holbert MA, Delgoshaie N, Wurtele H, Guillemette B, Meeth K, Yuan H, Drogaris P, Lee E-H, Durette C, Thibault P, Verreault A, Cole PA & Marmorstein R (2011) Structure of the Rtt109-AcCoA/Vps75 complex and implications for chaperone-mediated histone acetylation. *Structure* **19:** 221–231

Tang Y, Holbert MA, Wurtele H, Meeth K, Rocha W, Gharib M, Jiang E, Thibault P, Verreault A, Cole PA & Marmorstein R (2008a) Fungal Rtt109 histone acetyltransferase is an unexpected structural homolog of metazoan p300/CBP. *Nat. Struct. Mol. Biol.* **15:** 738–745

Tang Y, Meeth K, Jiang E, Luo C & Marmorstein R (2008b) Structure of Vps75 and implications for histone chaperone function. *Proc. Natl. Acad. Sci. U. S. A.* **105:** 12206–12211

Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B & Zhao Y (2011)

Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146:** 1016–1028

Thompson PR, Wang D, Wang L, Fulco M, Pediconi N, Zhang D, An W, Ge Q, Roeder RG, Wong J, Levrero M, Sartorelli V, Cotter RJ & Cole PA (2004) Regulation of the p300 HAT domain via a novel activation loop. *Nature Structural & Molecular Biology* **11:** 308–315 Available at: http://dx.doi.org/10.1038/nsmb740

Tompa P & Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends in Biochemical Sciences* **33:** 2–8 Available at: http://dx.doi.org/10.1016/j.tibs.2007.10.003

Trewhella J, Duff AP, Durand D, Gabel F, Guss JM, Hendrickson WA, Hura GL, Jacques DA, Kirby NM, Kwan AH, Pérez J, Pollack L, Ryan TM, Sali A, Schneidman-Duhovny D, Schwede T, Svergun DI, Sugiyama M, Tainer JA, Vachette P, et al (2017) 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update. *Acta Crystallogr D Struct Biol* **73:** 710–728

Tropberger P & Schneider R (2013) Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat. Struct. Mol. Biol.* **20:** 657–661

Tsubota T, Berndsen CE, Erkmann JA, Smith CL, Yang L, Freitas MA, Denu JM & Kaufman PD (2007) Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. *Mol. Cell* **25:** 703–712

Tugarinov V & Kay LE (2003) Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *Journal of the American Chemical Society* **125:** 13868–13878 Available at: http://dx.doi.org/10.1021/ja030345s

Tugarinov V & Kay LE (2006) Relaxation rates of degenerate 1H transitions in methyl groups of proteins as reporters of side-chain dynamics. *J. Am. Chem. Soc.* **128:** 7299–7308

Tugarinov V & Kay LE (2013) Estimating Side-Chain Order in [U-2H;13CH3]-Labeled High Molecular Weight Proteins from Analysis of HMQC/HSQC Spectra. *The Journal of Physical Chemistry B* **117:** 3571–3577 Available at: http://dx.doi.org/10.1021/jp401088c

Turner BM (1991) Histone acetylation and control of gene expression. *J. Cell Sci.* **99 (Pt 1):** 13–20

Vuister GW, Clore GM, Gronenborn AM, Powers R, Garrett DS, Tschudin R & Bax A (1993) Increased Resolution and Improved Spectral Quality in Four-Dimensional 13C/13C-Separated HMQC-NOESY-HMQC Spectra Using Pulsed Field Gradients. *Journal of Magnetic Resonance, Series B* **101:** 210–213 Available at: http://dx.doi.org/10.1006/jmrb.1993.1035

Wang L, Tang Y, Cole PA & Marmorstein R (2008) Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. *Curr. Opin. Struct. Biol.* **18:** 741–747

Warren C, Matsui T, Karp JM, Onikubo T, Cahill S, Brenowitz M, Cowburn D, Girvin M & Shechter D (2017) Dynamic intramolecular regulation of the histone chaperone nucleoplasmin controls histone binding and release. *Nat. Commun.* **8:** 2215

Warren C & Shechter D (2017) Fly Fishing for Histones: Catch and Release by Histone Chaperone Intrinsically Disordered Regions and Acidic Stretches. *J. Mol. Biol.* **429:** 2401–2426

Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73:** 1–16

Wittekind M & Mueller L (1993) HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *Journal of Magnetic Resonance, Series B* **101:** 201–205 Available at: http://dx.doi.org/10.1006/jmrb.1993.1033

Xu M, Long C, Chen X, Huang C, Chen S & Zhu B (2010) Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. *Science* **328:** 94–98

Yuan J, Pu M, Zhang Z & Lou Z (2009) Histone H3-K56 acetylation is important for genomic stability in mammals. *Cell Cycle* **8:** 1747–1753 Available at: http://dx.doi.org/10.4161/cc.8.11.8620

Yu Q, Olsen L, Zhang X, Boeke JD & Bi X (2011) Differential contributions of histone H3 and H4 residues to heterochromatin structure. *Genetics* **188:** 291–308

Zhang L, Serra-Cardona A, Zhou H, Wang M, Yang N, Zhang Z & Xu R-M (2018) Multisite Substrate Recognition in Asf1-Dependent Acetylation of Histone H3 K56 by Rtt109. *Cell* **174:** 818–830.e11

Zimm BH (1948) The Scattering of Light and the Radial Distribution Function of High Polymer Solutions. *The Journal of Chemical Physics* **16:** 1093–1099 Available at: http://dx.doi.org/10.1063/1.1746738

Zuiderweg ERP, McIntosh LP, Dahlquist FW & Fesik SW (1990) Three-dimensional 13C-resolved proton NOE spectroscopy of uniformly 13C-labeled proteins for the NMR assignment and structure determination of larger molecules. *Journal of Magnetic Resonance (1969)* **86:** 210–216 Available at: http://dx.doi.org/10.1016/0022-2364(90)90228-2

Curriculum vitae

Nataliya Danilenko

Education:

2015 – 2020   **Leibniz University of Hannover (BMWZ, Hannover, Germany)**,
Ph.D. student in Prof. Dr. T. Carlomagno group.
Thesis: Structural basis for histone H3 acetylation by Rtt109 in complex with histone chaperones Asf1 and Vps75.

2014 – 2015   **European Molecular Biology Laboratory (EMBL Heidelberg, Germany)**, Ph.D. student in Prof. Dr. T. Carlomagno group.

2009 – 2014   **Moscow State University (Moscow, Russia)**.
Diploma in Biochemistry.
Thesis: Transcription factors involved in proteasomal genes regulation in *S. cerevisiae.*

Research experience:

2014 – present   BMWZ, Leibniz University of Hannover, Hannover, Germany.
Ph.D. student; Advisor: Prof. Dr. T. Carlomagno.

2012 – 2014   Engelhardt Institute of Molecular Biology RAS, Moscow, Russia.
Diploma student; Advisor: Prof. V.L. Karpov

Conferences and courses:

2018   7th Murnau Conference: New Frontiers in Structural Biology (Murnau, Germany).
Presented poster.

2017   AILM2017: advanced isotopic labeling methods for integrated structural biology (Grenoble, France).
Presented poster, selected for promoted talk.

2016   EMBO NMR course (Joachimsthal, Germany).

2016   Biochemistry 2016. Shaping the Molecules of Life: Chemical Biology of Nucleic Acid and Protein Modifications (Frankfurt, Germany).
Presented poster.

Publications:

Danilenko N, Lercher L, Kirkpatrick J, Gabel F, Codutti L & Carlomagno T (2019) Histone chaperone exploits intrinsic disorder to switch acetylation specificity. *Nat. Commun.* 10(1).

Lercher L, **Danilenko N**, Kirkpatrick J & Carlomagno T (2017) Structural characterization of the Asf1-Rtt109 interaction and its role in histone acetylation. *Nucleic Acids Res.* 46(5).

Karpov DS, Grineva EN, Leinsoo AT, Nadolinskaia NI, **Danilenko NK**, Tutyaeva V V., Spasskaya DS, Preobrazhenskaya O V., Lysov YP & Karpov VL (2017) Functional analysis of Debaryomyces hansenii Rpn4 on a genetic background of Saccharomyces cerevisiae. *FEMS Yeast Res.* 17(1).