Functional characterization of N-terminal acetyltransferase 10 (*NAA10*) variants potentially causing disease

By
Christine Darbakk

Thesis submitted in partial fulfilment of the requirements for the degree of Master of Science

Department of Molecular Biology
Faculty of Mathematics and Natural Sciences
University of Bergen

November 2019

Takk til

Arbeidet med dette masterprosjektet ble utført i NAT-gruppen ved Institutt for Biovitenskap, Universitetet i Bergen. Laboratoriearbeid har vært utført på Institutt for Biomedisin ved Haukeland Universitetssykehus.

Først og fremst vil jeg takke mine fantastiske medveiledere, Nina McTiernan og Rasmus Ree. Deres hjelp har vært uvurderlig i denne prosessen. Dere har stilt på lab hver dag med strålende optimisme og godt humør. Døren deres har alltid vært åpen og dere har gledelig svart på både store og små spørsmål. Nina, takk for at du har hjulpet meg med store og små oppgaver, selv utenfor arbeidstid. Jeg setter virkelig pris på det! Og selvfølgelige, takk til dere begge for hjelpen under skriveprosessen. Dere har virkelig gjort en fantastisk jobb som mine veiledere.

Så vil jeg takke min hovedveileder Thomas Arnesen for at jeg fikk mulighet til å jobbe med dette spennende prosjektet. Du er en fantastisk, sympatisk leder som har en evne til å motivere de rundt deg. Jeg vil også takke deg for all hjelp i skriveprosessen. Uansett morgen, kveld eller helg har du tatt imot oppgaven min for retting. Å jobbe for deg har virkelig gitt verdifull erfaring og kompetanse. Jeg vil også rette en stor takk til resten av NAT-gruppen. Hver og en av dere stilte alltid opp for å hjelpe eller gi råd. Jeg har hatt mange gode opplevelser både faglig og sosialt med dere.

Jeg må også benytte anledningen til å takke min kjære medstudent, labpartner og venninne Karoline Krogstad. Takk for mange festlige lab dager og utallige lange lunsj/kaffepauser. Du har virkelig en evne til å spre glede og positivitet. Året hadde ikke vært det samme uten deg.

Til slutt vil jeg takke familien min for all støtte og motivasjon.

Bergen, November 2019
Christine Darbakk

Table of Contents

Takk til ... **3**

Selected abbreviations ... **6**

1 Summary ... **7**

2 Introduction ... **8**

2.1 Protein acetylation ... *8*
2.1.1 Lysine acetylation ... 9
2.1.2 N-terminal acetylation ... 9

2. 2 N-terminal acetyltransferases (NATs) ... *10*
2.2.1 The human NAT machinery ... 11
2.2.2 Structure of NATs ... 12

2.3 The NatA complex ... *13*
2.3.1 Other NatA components ... 14
2.3.2 The structure of human NatA ... 15
2.3.3 NatA in human diseases ... 16
2.3.4 Pathological NAA10 variants ... 17

2.4 Aim of study ... *22*

3 Materials ... **23**

3.1 Antibodies ... *23*

3.2 Apparatus and instruments ... *23*

3.3 Bacterial strain ... *23*

3.4 Buffers, solutions and media ... *23*

3.5 Chemicals and reagents ... *25*

3.6 Commercial kits ... *26*

3.7 Human cell line ... *26*

3.8 Molecular size marker ... *26*

3.9 Peptides ... *26*

3.10 Plasmids ... *26*

3.11 Primers ... *27*

3.12 Software ... *27*

3.13 Web resources ... *27*

4 Methods ... **28**

4.1 Bioinformatics analysis ... *28*
4.1.1 Multiple sequence alignment (MSA) ... 28
4.1.2 *In silico* missense mutation prediction ... 29
4.1.3 *In silico* structural illustration and predictions ... 30

4.2 Mutagenesis, plasmid preparation and DNA sequencing ... *30*
4.2.1 Introducing a missense mutation in *NAA10* using site directed mutagenesis ... 30
4.2.2 Plasmid preparation and plasmid DNA isolation ... 31
4.2.3 Verifying mutation by DNA sequencing ... 31

4.3 Maintenance and transfection of HeLa cells and NAA10 stability analysis *32*
4.3.1 Maintenance of HeLa cells 32
4.3.2 Transfection of HeLa cells 32
4.3.3 Determine half-life of NAA10 variants using cycloheximide chase assay 33
4.3.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) 34
4.3.5 Western blot (WB) analysis 34
4.4 Complex formation and intrinsic catalytic activity of NatA *36*
4.4.1 Immunoprecipitation (IP) 36
4.4.2 [^{14}C]-Ac-CoA-based *in vitro* Nt-acetylation assay 37
5 Results **39**
5.1 Genetic variants and clinical data *39*
5.2 Bioinformatic analysis *40*
5.2.1 Multiple sequence alignment (MSA) 40
5.2.2 *In silico* missense mutation predictions 42
5.2.3 *In silico* structural analysis 43
5.3 In vitro stability analysis of NAA10 variants by cycloheximide chase assay *47*
5.4 NatA complex formation and in vitro activity of NatA *49*
6 Discussion **54**
6.1 Genetic variants and clinical data *54*
6.2 Functional analysis of NAA10 variants *55*
6.2.1 NAA10 L11R variant 55
6.2.2 NAA10 H16P variant 58
6.3 Conclusion *61*
6.4 Future perspective *62*
7 References **64**
8 Supplementary **70**

Selected abbreviations

ab	Antibody	KDAC	Lysine deacetylase
Ac-CoA	Acetyl Coenzyme A	LMS	Lenz microphthalmia syndrome
CoA	Coenzyme A	MSA	Multiple sequence alignment
ddNTP	Dideoxynucleotide	NAA #	N-alpha acetyltransferases protein #
DTT	Dithiothreitol	*NAA #*	N-terminal acetyltransferase gene #
E. coli	*Escherichia coli*	NAT	N-terminal acetyltransferase
FBS	Fetal bovine serum	NDAC	N-terminal deacetylases
GNAT	GCN5-related N-acetyltransferase	Nt	N-terminal
GCN5	General control nonderepressible 5	PAGE	Polyacrylamide gel electrophoresis
His	Histidine	PenStrep	Penicillin-Streptomycin
HRP	Horseradish peroxidase	*S. pombe*	*Schizosaccharomyces pombe*
HYPK	Huntingtin-interacting protein K	SDS	Sodium dodecylsulphate
iMet	Initiator methionine	v/v	Volume/volume
IP	Immunoprecipitation	w/v	Weight/volume
KAT	Lysine acetyltransferase	WB	Western blotting
kDa	Kilodalton		

1 Summary

Approximately 80-90 % of all eukaryotic proteins are co- or post-translationally acetylated on their N-terminus by a group of enzymes called N-terminal acetyltransferases (NATs) (Arnesen et al., 2009). To date, eight NATs have been identified in eukaryotes, seven of which (NatA-NatF and NatH) are found in humans. Each of the NATs differ in subunit composition and have a distinct substrate specificity (Aksnes et al., 2019).

The NatA complex is conserved from yeast to humans, acetylating approximately 40 % of the human proteome (Arnesen et al., 2009). NatA is composed of the catalytic subunit NAA10 and the auxiliary subunit NAA15 and has the broadest substrate specificity among the NATs (Arnesen et al., 2005a, Liszczak et al., 2013). In 2011, Rope et al., reported a *NAA10* S37P missense mutation to be the cause of the lethal X-linked disorder named Ogden syndrome (Rope et al., 2011). Some years later, Esmailpour and colleagues reported that the genetically heterogeneous X-linked disorder Lenz microphtalmia syndrome (LMS) was caused by a splice mutation in *NAA10* (Esmailpour et al., 2014). The last decade, several other *NAA10* mutations have been reported to have pathological effects in the harboring patient. Intellectual disability, development delay, growth deficiency and cardiac and skeletal anomalies are among the most common phenotypes coupled to NAA10 deficiency.

The focus of this thesis has been to functionally characterize two missense mutations in *NAA10* suspected to cause disease in humans. These mutations are NAA10 L11R and NAA10 H16P, which were identified in female patients presenting with some symptoms typical of NAA10 deficiency. NatA complex formation and *in vitro* intrinsic catalytic activity, and cellular stability have been characterized, and bioinformatic analyses have been performed. The work presented in this thesis demonstrates that the NAA10 L11R variant and H16P variants have a reduced NatA complex formation and their NatA activity is functionally impaired. The L11R variant affects NatA activity to a smaller extent than the H16P mutation. In the cellular stability assay, the NAA10 L11R had a destabilizing effect, whereas NAA10 H16P appears stable and is unlikely to affect neither monomeric NAA10 nor NatA stability. The study presented in this thesis support that these variants are pathological, yet further studies are needed to define the detailed underlying mechanisms.

2 Introduction

2.1 Protein acetylation

There are approximately 19 000 – 22 000 protein-coding genes in the human genome, which in turn give rise to the extremely complex human proteome (International Human Genome Sequencing Consortium, 2004, Kim et al., 2014). Each protein-coding gene can give rise to multiple proteins mainly through two mechanisms, alternative splicing and protein modifications (Ponomarenko et al., 2016; Pertea et al., 2012; Verdin et al., 2015). There are several protein modifications taking place during or after synthesis, which are essential for regulation of protein stability, activity, localization or interactions. Protein modifications include proteolysis, glycosylation, methylation, sumoylation, phosphorylation, ubiquitination and acetylation (Boutureira and Bernardes, 2015), where the latter is the modification in focus of this thesis.

There are predominantly two types of protein acetylation; lysine acetylation (Nε-acetylation) and N-terminal acetylation (Nt-acetylation, Nα-acetylation) (Figure 2.1). Lysine acetylation is the transfer of an acetyl group donated by acetyl coenzyme A (Ac-CoA) to the ε-amino group of lysine side chains. N-terminal acetylation is the transfer of an acetyl group from Ac-CoA to the α-amino group of the amine group on the N-terminus of proteins. Addition of an acetyl to the α-amino group on the N-terminus or on the ε-amino group of lysine side chains neutralizes the positive charge and alters the electrostatic properties of proteins (Aksnes et al., 2016).

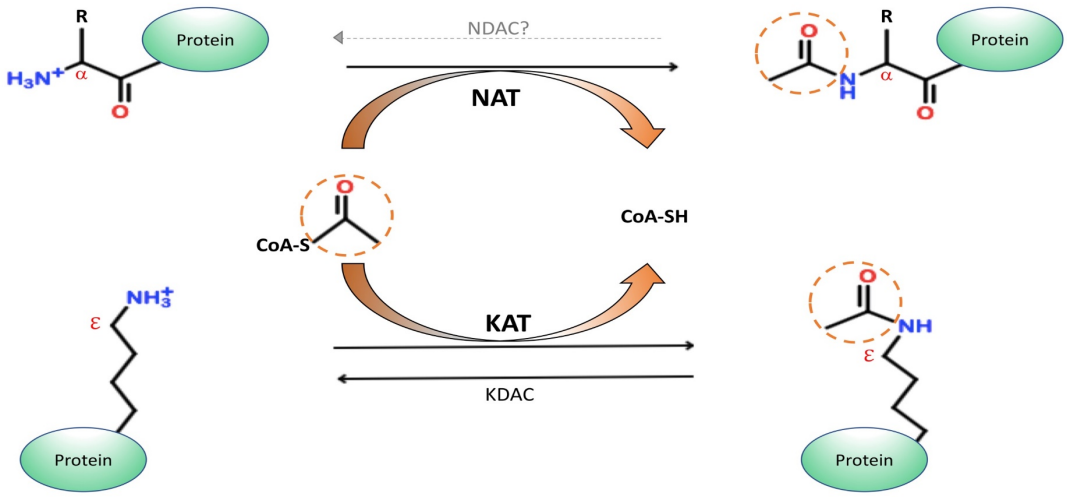

Figure 2.1: Nε-acetylation and Nt-acetylation catalyzed by KATs and NATs. Protein N-acetyltransferases catalyze the transfer of an acetyl group (orange circle) from acetyl coenzyme A (Ac-CoA) onto a substrate protein. Lysine acetyltransferases (KATs, bottom reaction) catalyze the transfer of the acetyl group to the ε-amino group of lysine side chains, whereas N-terminal acetyltransferases (NATs, top reaction) catalyze the transfer to the α-amino group of the amine group on the N-terminus. Lysines can be deacetylated by lysine

deacetylases (KDACs), whereas Nt-acetylation is considered irreversible since no N-terminal deacetylases (NDACs) are known. Figure modified from Aksnes et al., 2019.

2.1.1 Lysine acetylation

Lysine acetylation is catalyzed by a group of enzymes, termed lysine acetyltransferases (KATs). Lysine acetylation plays important roles in gene regulation, cell signaling and metabolism and is by far the most studied acetylation event. (Drazic et al., 2016). Transcription factors, nuclear receptors, chaperones, metabolic enzymes and cytoskeleton proteins are known to be subjected to lysine acetylation (Glozak et al., 2005, Kim et al., 2006, Drazic et al., 2016). Lysine acetylation is a reversible process by the action of lysine deacetylases (KDACs). By the combined activity of KATs and KDACs, lysine acetylation contributes to a dynamic regulation of several biological processes both on gene and protein levels (Drazic et al., 2016).

2.1.2 N-terminal acetylation

N-terminal acetyltransferases (NATs) are a group of enzymes responsible for catalyzing N-terminal acetylation. Nt-acetylation is one of the most common protein modifications and occurs on the majority (~80-90 %) of both cytosolic and transmembrane eukaryotic proteins (Arnesen et al., 2009; Aksnes et. al., 2016). Most proteins are Nt-acetylated co-translationally as the newly synthesized protein emerges from the ribosome, but post-translational modifications also take place. Proteins subjected to Nt-acetylation can be either fully or partially acetylated. This means that one specific protein molecule can dually exist in either acetylated or unacetylated form in the same cell with any stoichometry ranging from 0-100 % acetylated (Arnesen et al., 2009). Unlike lysine acetylation, Nt-acetylation is considered an irreversible process since no N-terminal deacetylases (NDACs) have been identified yet (Drazic et al., 2016).

Nt-acetylation influences protein subcellular targeting, protein interactions and complex formation, folding and aggregation and protein turnover through proteasomal degradation pathways such as the Ac/N-end degron pathway (Behnia et al., 2004, Dikiy et al., 2014, Scott et al., 2011, Arnesen et al., 2010, Holmes et al., 2014, Hwang et. al., 2010, Varshavsky et al., 2019). Ac/N-end degrons are specific degradation signals created by Nt-acetylation of proteins containing M, S, A, T, V, G, P or C at the N-terminus. These degradation signals are recognized by specific E3 ubiquitin ligases, submitted for degradation through the Ac/N-end degron pathway via ubiquitination of a downstream lysine residue (Hwang et al., 2010). However, a recent global reporter assay using yeast suggested that Nt-acetylation in yeast is not a major determinant for protein degradation. In fact, in some rare cases, Nt-acetylation potentially protected proteins from N-degron pathway mediated degradation (Kats et al., 2018). Shemorry et al., reported the first examples of protein regulation through the physiologically relevant transitions that shield and unshield natural Ac/N-degrons. They showed that some Ac-N degrons are conditional on the acetylated protein not having a

binding partner that protects the Nt-acetyl group from being recognized (Shemorry et al., 2013).

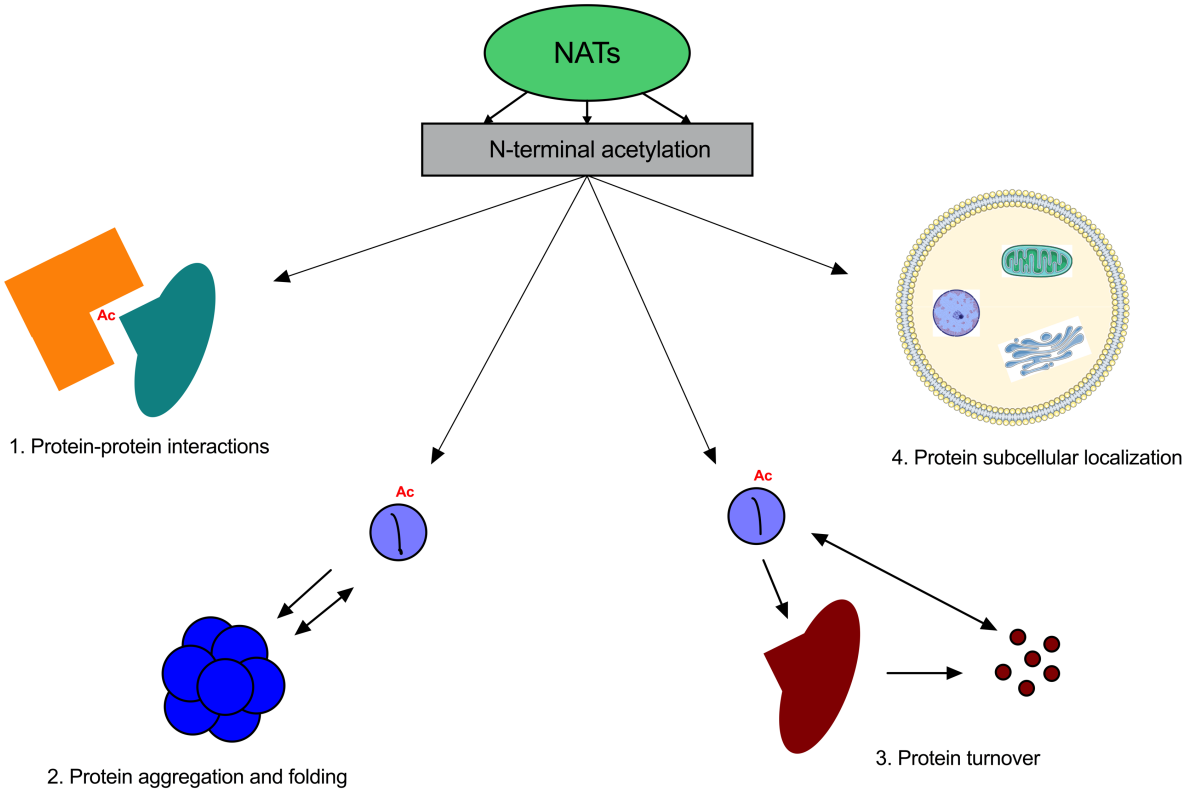

Figure 2.2: Effect of protein N-terminal acetylation. Nt-acetylation can affect a wide spectrum of protein functions including protein-protein interactions (1), facilitate protein folding and prevent protein aggregation (2), regulating protein lifetime or protein turnover (3), and target and regulate correct protein subcellular localization (4). Figure modified from Aksnes et al., 2019.

2. 2 N-terminal acetyltransferases (NATs)

Nt-acetylation is a protein modification that is conserved from bacteria to eukaryotes (Arnesen et al., 2009; Drazic et al., 2016). The group of enzymes facilitating this modification is called N-terminal acetyltransferases (NATs). To date, eukaryotes have eight defined NATs, termed NatA-NatH (Aksnes et al., 2019). NatA-NatE are common in all eukaryotes, whereas NatF is mainly found in multicellular eukaryotes like animals and plants (Van Damme et al., 2011b). NatG is exclusively found in the plant kingdom (Dinh et al., 2015), whereas NatH is found in the animal kindom (Drazic et al., 2018). In contrast to eukaryotes, only three NATs named RimI, RimJ and RimL have been identified in prokaryotes (Soppa et al., 2010).

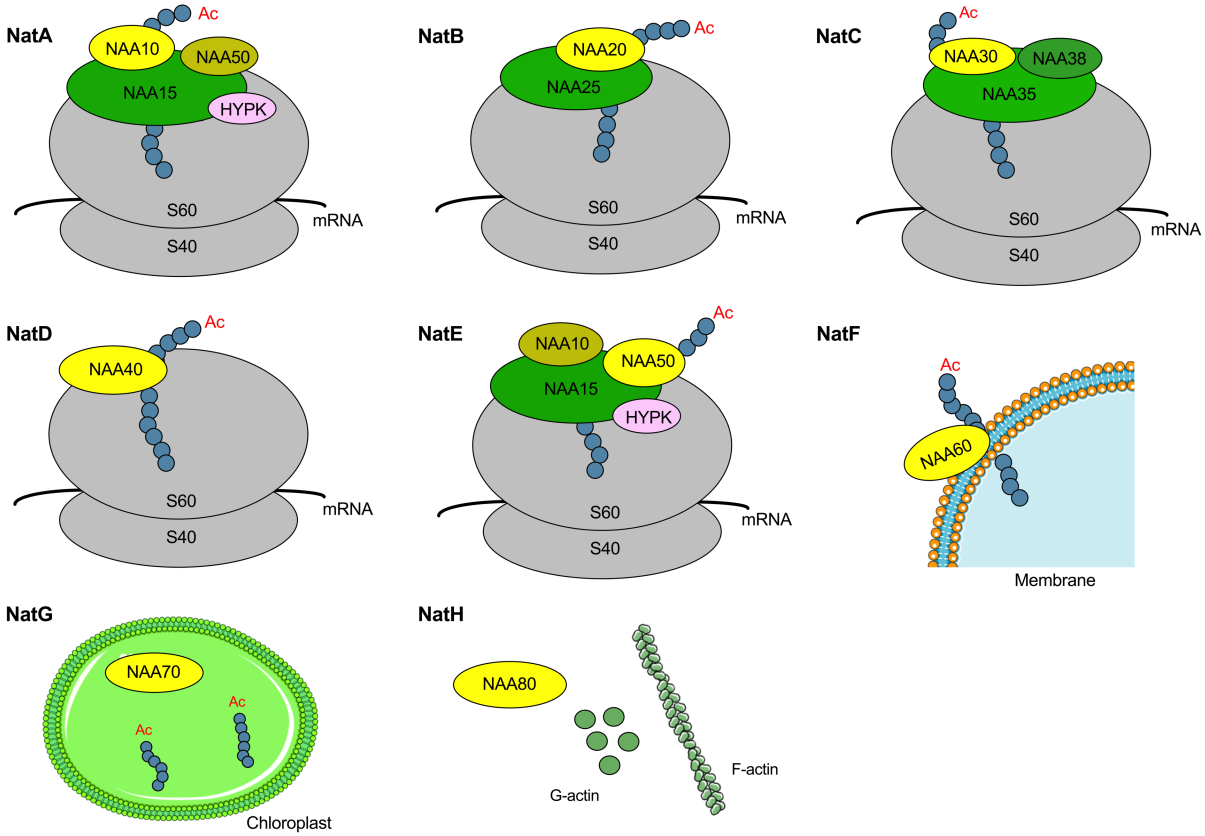

Figure 2.3: The eukaryotic NAT machinery. The eight eukaryotic NATs identified to date, NatA-NatH, are composed of either one catalytic subunit (yellow) alone or in complex with one or two auxiliary subunits (green). The ribosomal interacting NATs, NatA-NatE mediates co-translational Nt-acetylation of substrates as nascent polypeptides emerge from the ribosome. NatF associates with the Golgi apparatus where it acetylates proteins post-translationally. NatG is found exclusively in plants and NatH is localized to the cytosol where it catalyzes Nt-acetylation of actin. 60S -large ribosomal subunit, 40S – small ribosomal subunit.

2.2.1 The human NAT machinery

All the eukaryotic NATs, except NatG, have been identified in mammalian cells to date. These enzymes differ in subunit composition and substrate specificity, but together they Nt-acetylate approximately 80 % of human proteins (Aksnes et al., 2019). NatA-NatE is conserved from yeast to human, while NatF only exists in higher eukaryotes (Van Damme et al., 2011b). Each of the human NATs consist of a monomeric catalytic subunit with one or two auxiliary subunits (Fig.2.3). NatD, NatF and NatH are exceptions, having no auxiliary subunits identified so far. NatA-NatE localize to the cytosol, where they interact with the ribosome performing acetylation of substrates co-translationally as nascent polypeptide chains emerge from the ribosome (Aksnes et al., 2019). The catalytic subunits NAA10, NAA40 and NAA50 appears to additionally localize to the nucleus (Aksnes et al., 2019). NatF localize to the Golgi membrane where it acetylates membrane proteins post-translationally (Aksnes et al., 2015c). Finally, the most recently discovered NAT, NatH, localize to the cytosol where it modifies Nt-acetylation of actins (Drazic et al., 2018).

The identity of the first two amino acids at the N-terminus mostly determine whether a protein is modified and by which NAT. The substrate specificity of each NAT is represented in Table 2.1. If the second amino acid after the initiator methionine (iMet) of newly synthesized proteins is small and polar, iMet is in most cases cleaved by methionine aminopeptidase (MetAP), constructing potential substrates for NatA (Frottin et al., 2006). There are large differences among these potential NatA substrates. For instance, N-termini with serines are almost always Nt-acetylated, while N-termini with glycine and proline are rarely Nt-acetylated (Arnesen et al., 2009; Goetze et al., 2009; Van Damme et al., 2011b). The substrate specificity of the eukaryotic NATs is shown in Table 2.1.

Table 2.1 Substrate specificity of eukaryotic NATs

NAT	Substrate specificity	References
A	Ala-, Ser-, Thr-, Val-, Cys-, Gly-.	Arnesen et al., 2005a; Arnesen et al., 2009
B	Met- (Asp/ Glu/ Asn/ Gln)	Starheim et al., 2008; Van Damme et al., 2012
C	Met- (Leu/ Ile/ Phe/ Tyr/ Lys)	Starheim et al., 2009; Van Damme et al., 2016
D	Histones H2A and H4 (SGRGK...)	Hole et al., 2011; Song et al., 2003
E	Met- (Ser/ Thr/ Ala/ Val/ Leu/ Ile/ Phe/ Tyr/ Lys)	Evjenth et al., 2009; Van Damme et al., 2011a; Van Damme et al., 2015
F	Met- (Leu/ Ile/ Phe/ Tyr/ Lys)	Van Damme et al., 2011b; Aksnes et al., 2015
G	Met-, Ala-, Ser-, Thr-.	Dinh et al., 2015
H	β- and γ-actin (DDD...; EEE...)	Drazic et al., 2018; Wiame et al., 2018

2.2.2 Structure of NATs

All NATs belong to the General control nonderepressible 5 (GCN5)-related N-acetyltransferase (GNAT) superfamily. This superfamily of acetyltransferases includes serotonin N-acetyltransferase, histone acetyltransferases, aminoglycoside N-acetyltransferases and NATs. Despite a low overall sequence homology, all members of the GNAT superfamily have a structurally characteristic conserved fold made up of six to seven β-strands and four α-helices (Fig. 2.4) (Lisczak et al., 2013; Ud-Din et al., 2016). Enzymes of the GNAT superfamily share the highly conserved Ac-CoA binding motif with the consensus sequence Q/RxxGxGA/A. The GNAT fold is made up of four conserved motifs (A-D) that are arranged in order C-D-A-B (Ud-Din et al., 2016). Residues in motifs C and D are important for the stability of the protein, whereas residues in motif A and B are important for acyl-CoA and acceptor substrate binding. The variation in the GNAT structure occurs mainly on the C-terminus, but some variation in secondary structures in the immediate N-terminus occurs as well (Lisczak et al., 2013; Ud-Din et al., 2016).

The catalysis performed by NATs is believed to follow an ordered sequential Bi-Bi mechanism where binding of Ac-CoA induces a conformational change required for efficient peptide substrate binding (Evjenth et al., 2012). The active site is shielded by an elongated loop region in the NAT family exluding NatD and NatH (Magin et al., 2016, Goris et al., 2018). As a result, the entrance to the substrate binding pocket is narrower compared to KATs (Goris et al., 2018). NatH has a more open substrate binding region as the shift upon substrate binding is larger compared with the other NATs (Magin et al., 2016).

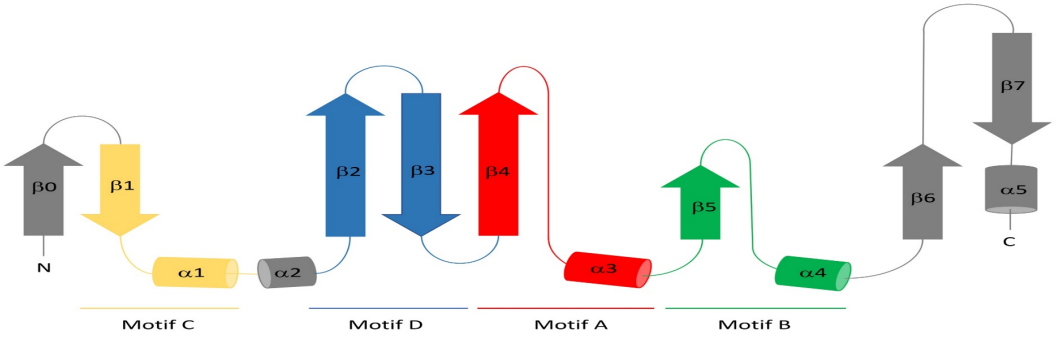

Figure 2.4: Topology of the conserved GNAT fold adapted by NATs. The motifs A-D colored in different colors are conserved among GNAT members. Residues in motifs C (yellow) and D (blue) are important for the stability of the protein, whereas residues in motif A (red) and B (green) are important for acyl-CoA and acceptor substrate binding. Figure modified from Ud-Din et. al., 2016.

2.3 The NatA complex

The NatA complex was first discovered in *Saccharomyces cerevisiae* and is by far the most studied NAT to date (Mullen et al., 1989; Park et al., 1992). NatA is conserved from yeast to humans, acetylating approximately 40 % of the human proteome (Arnesen et al., 2009). NatA is considered the major NAT due to its broad substrate specificity. The auxiliary subunit NAA15 and the catalytic subunit NAA10 forms the NatA complex together with two additional interacting partners, NAA50/NatE and HYPK (Arnesen et al., 2005a, Arnesen et al., 2010, Gautschi et.al., 2003). The human NAA15 subunit is composed of 866 amino acid residues. It functions as a ribosomal anchor and modulates the substrate specificity of NAA10 (Lisczak et al., 2013; Gautschi et. al., 2003). The human NAA10 subunit is composed of 235 residues, carrying out the acetylation. NAA15 localizes to the cytoplasm, while NAA10 localizes to both the cytoplasm and the nucleus (Arnesen et al., 2005a). Monomeric NAA10 has an altered substrate specificity, preferring acidic substrates when not bound in complex. This suggest that NAA10 might exert individual roles in the cell (Arnesen et al., 2005a; Van Damme et al., 2011a).

There are several studies demonstrating that NAA10 exerts classic roles as the catalytic subunit of the NatA complex. Approximately 40 % of the human proteome is co-translationally acetylated when NAA10 is part of the NatA complex. This contributes to diverse cellular functions such as proliferation and normal organismal development (Arnesen et al., 2009). Studies of malfunctioning NAA10 in complex with NatA have been linked to

abnormal development in humans and several organisms such as *Trypanosoma brucei* and *Drosophila melanogaster* (Rope et al., 2015; Ingram et al., 2000; Wang et al., 2010). Another study done on *Danio rerio* showed that loss of NAA10 function was associated with short growth, body axis malformations, reduced pigmentation and eye underdevelopment during embryonic development (Ree et al., 2015). Loss of NatA activity has also been linked to diminished cell proliferation or increased cell death (Arnesen et al., 2006c; Fisher et al., 2005; Gromyko et al., 2010; Myklebust et al., 2015). In *Arabidopsis,* loss of NatA activity has been linked to lethality (Linster et al., 2015).

Several studies have also suggested NAA10 to work as a KAT, where it post-translationally catalyzes lysine acetylation of internal lysine residues of several substrates such as Runx2, MSRA, PGK1, Hsp70, Cdcd25A and hypoxia-inducible factor 1α (HIF1α) among others. (Yoon et al., 2014, Shin et al., 2014, Seo et al., 2016, Lozada et al., 2016, Jeong et al., 2002). There is an ongoing debate concerning the ability of NAA10 to function as a KAT. Contradicting studies have disputed the claim that NAA10 work as a KAT. Arnesen and co-workers showed that NAA10 was not able to acetylate HIF1α *in vitro* (Arnesen et al., 2005a). The more recent report by Magin et al., in 2016 showed that purified NAA10 exhibits undetectable KAT activity (Magin et al., 2016). They report assays where there is no difference in lysine acetylation of Runx2 and MSRA with or without NAA10. They suggest that these substrates may be chemically, rather than enzymatically acetylated, and that lysine residues will not fit the substrate binding site due to an extended loop occluding lysine side chains within a polypeptide from lying across the active site (Magin et al., 2016). A study conducted by Kang and colleagues in 2018 proposes that factor inhibiting HIF (FIH) catalyzes hydroxylation at W38 of NAA10 in an oxygen-dependent manner. This is suggested to widen the substrate gate, permitting lysine acetylation (Kang et al., 2018).

Monomeric NAA10 has also been linked to non-catalytic roles independent of the NatA complex. One particular role is the direct binding of NAA10 to non-methylated DNA motifs. This binding recruits the DNA methyltransferase 1 (DNMT1), resulting in maintenance of the genomic imprinting and silencing of genes. Tumor suppressor genes are among the genes being silenced, hence contributing to cancer development (Lee et al., 2017).

2.3.1 Other NatA components

HYPK was first found to be a stably associated component of the human NatA complex and essential for optimal NatA mediated Nt-acetylation of substrates (Arnesen et al., 2010). However, recent studies involving the solved structure of the NAA10-NAA15-HYPK complex in human and *C. thermophilum* found that HYPK acts as a negative regulator of NatA activity by blocking the active site (Weyer et al., 2017; Gottlieb et al., 2018). In addition, the NatA-HYPK complex has been shown to be important in preventing the protein Huntingtin (Htt) from polyglutamine aggregation (Raychaudhuri et al., 2008).

NatA can also form a complex with the catalytic subunit NAA50 forming the ternary NatA/NAA50 complex, also known as NatE (Gautschi et.al., 2003). NAA50 is a NAT that associates with NatA, but has its own distinct substrate specificity and catalytic activity (Evjenth et al., 2009). Knockout and deletion studies of NAA50 in higher eukaryotes such as human and *Drosophila* revealed a decrease of *in vivo* NatA acetylation activity and impairment of sister chromatid cohesion. (Hou et al., 2007; Chu et al., 2011; Ribeiro et al., 2016; Williams et al., 2003; Rathore et al., 2016). Another study by Knorr et al., 2018 revealed that both NAA50 and NAA15 of NatA contributes to ribosome interaction in yeast (Knorr et. al., 2018). The active site of NAA10 was found to be 50 Å from the ribosome exit tunnel, whereas the active site of NAA50 was found to be two times that distance (Knorr et. al., 2018). The functional coupling and interplay between these two catalytic subunits are still poorly understood. A recent study from Deng and colleagues demonstrated that the catalytic activity of the two catalytic subunits within the human NatA/NAA50 complex is significantly promoted by the NatA-NAA50 interaction relative to NatA or NAA50 alone (Deng et al., 2019).

2.3.2 The structure of human NatA

The human NatA structure was solved by Gottlieb and colleagues in 2018 (Gottlieb et. al., 2018). The reported crystal structure contained the catalytic subunit NAA10 with a truncated C-terminus (1-160 out of 235 residues) and the full-length auxiliary subunit NAA15 with a 6xHis-tagged N-terminus (1-866 residues). NAA15 is composed of 45 α-helices that form a ring-like structure with a total of 13 tetratricopeptide repeat (TRP) motifs and a C-terminal Sel1-like repeat (SLR) (Zimmermann et. al. 2017). These segments participate in protein-protein interactions between NAA10, the 60S ribosomal subunit, NAA50 and HYPK (Blatch and Lassle 1999; Mittl and Schneider 2007; Liszczak et al., 2013; Neubauer 2012; Weyer et al., 2017;). NAA10 is docked to the center of the NAA15 ring-structure and forms the characteristic GNAT fold consisting of five α-helices and seven β-sheets. A study in *S. pombe* revealed that binding of NAA15 induces a conformational change in NAA10, resulting in an altered substrate specificity (Liszcak et al., 2013).

Structural studies of NatA revealed an electron density for inositol hexakisphosphate (IP_6) between NAA10 and NAA15. IP_6 appears to have a stabilizing effect on the NAA10-NAA15 interaction (Gottlieb et al., 2018), and is found to be bound by a series of hydrophilic and electrostatic interactions within α19, α24 and α25 helices of human NAA15 and the β2-loop-β3 and β4-loop-α3 segments of human NAA10 (Gottlieb et. al., 2018). Almost the same binding mode was found in the ternary NatA/NAA50 complex in *S. cerevisiae* (Deng et al., 2019). Notably, the residues of NAA15 that binds IP_6 are conserved from yeast to human, suggesting that this binding mode is structurally conserved and that it might be of biological relevance (Deng et al., 2019).

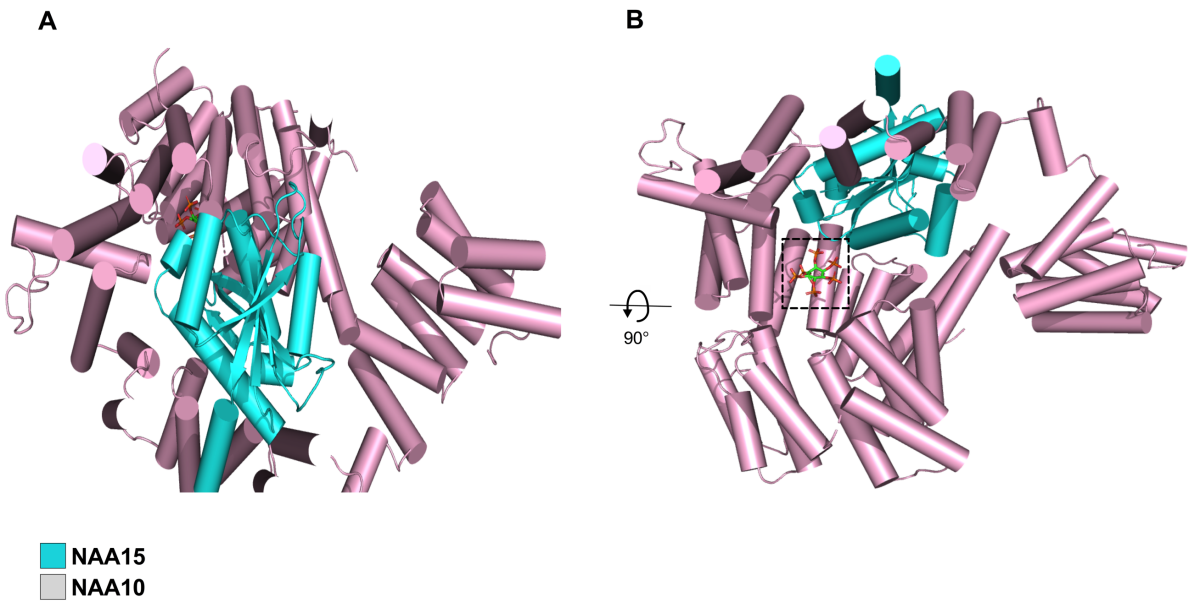

Figure 2.5: Structure of the human NatA complex. (A) The ribosomal anchorage NAA15 (light pink) binds the catalytic subunit NAA10 (cyan) in a tight binding pocket. (B) 90° rotation of the hNatA structure. The inositol hexakisphosphate (IP_6) molecule (boxed in) is believed to be structurally conserved and of biological relevance. Figure modified with PyMoL. PDB: 6C9M (Gottlieb et al., 2018).

2.3.3 NatA in human diseases

The biological importance of Nt-acetylation has over the past decade gotten more attention as an increasing number of studies report correlations between NAT expression or function and human disease. Studies have linked both NAA10 and NAA15 to tumor development, where elevated levels of these subunits are linked to tumor aggressiveness and poor prognosis (Kalvik et al., 2013, Drazic et al., 2016). Overexpressed levels of the *NAA15* gene have been found in papillary thyroid carcinoma, neuroblastoma and gastric cancer (Fluge et. al., 2002, Martin et al., 2007, Line et. al., 2002). Elevated levels of *NAA10* have also been implicated in several types of cancer such as osteosarcoma, prostate cancer, colorectal cancer, breast cancer, lung cancer and hepatocellular carcinoma (Chien et al., 2018, Midorikawa et al., 2002, Ren et al., 2008, Yu et al., 2009, Lee et al., 2010, Wang et al., 2012). Elevated levels of *NAA10* have also been found in adipose tissue in obese individuals. A study of conventional and adipocyte-specific NAA10 KO mice suggests that NAA10 promotes diet-induced obesity by suppressing beige adipocyte differentiation and thermogenic gene expression. (Lee et al., 2019).

Studies have also suggested that aberrant NatA function may be linked to neurodegenerative disorders such as Huntington's- and Alzheimer´s disease (Arnesen et al., 2010, Asaumi et al., 2005). Huntington´s disease is believed to arise from the aggregation of the protein Huntingtin (Htt) forming macrostructures leading to puncture and degenerated neural cells (Raychaudhuri et al., 2008, Chiti et al., 2006). The NatA in complex with HYPK has been suggested to prevent aggregation of Htt and is suggested to play a role involved in preventing

this disease. (Arnesen et al., 2010). NatA has also been indicated to stabilize the β-amyloid precursor protein (APP), preventing it from forming plaque amyloid β-protein that potentially contributes to development of Alzheimer´s disease (Asaumi et al., 2005).

2.3.4 Pathological NAA10 variants

The first human genetic disorder associated with aberrant Nt-acetylation activity was reported in 2011. Rope and colleagues reported a missense mutation in exon 2 in the *NAA10* gene on the Xq28 chromosome to be the cause of the previously undescribed genetic disorder Ogden syndrome (Rope et al., 2011). Ogden syndrome is an X-linked disorder characterized by global developmental delay, hypotonia, cryptorchidism, craniofacial anomalies, cardiac malformation or arrhythmia and aged appearance (Figure 2.6).

The syndrome was originally described in 8 boys from two unrelated families. The missense mutation reported as the cause of Ogden syndrome was a S37P substitution variant in the *NAA10* gene (Rope et al., 2011). This study by Rope et al., demonstrated the *in vitro* catalytic activity of NAA10 S37P to be disrupted. A study following up on the effects caused by the S37P mutation revealed that NAA10 had a 20-80 % reduction of *in vitro* enzymatic activity dependent on the substrate polypeptide as well as an impaired NatA complex formation (Myklebust et al., 2015). In the same study, proteomic studies revealed reduced Nt-acetylation of NatA substrates *in vivo*. The proteomic studies were performed on cells derived from affected boys and healthy family members, suggesting that the cause of Ogden syndrome is at least in part caused by aberrant NatA function. The females harboring the S37P mutation in the affected families displayed skewed X-chromosome inactivation toward the wild-type allele, explaining the lack of severe symptoms in carrier females (Myklebust et al., 2015).

Figure2.6: Affected boys identified with Ogden syndrome. The lethal X-linked genetic disorder Ogden syndrome is caused by a S37P missense mutation in the *NAA10* gene. The syndrome is characterized by global developmental delay, craniofacial anomalies, cryptorchidism, hypotonia, cardiac malformation or arrhythmia and aged appearance. All of the affected boys died within 8-16 months of age. Figure from Rope et al., 2011.

In 2014, Esmailpour and co-workers identified a splice mutation in the intron 7 donor site of the *NAA10* gene in three brothers from the same family affected by the Lenz microphthalmia syndrome (LMS) (Esmailpour et al., 2014). LMS is a genetically heterogeneous X-linked disorder characterized by microphthalmia/anophthalmia, eye malformation and abnormalities of the hands, ears, teeth, skeleton and the urinary system. Approximately 60 % of affected males exhibit delayed development and/or intellectual disability and seizure disorder. Unlike

Ogden syndrome, LMS is not lethal during infancy (Esmailpour et al., 2014). The *NAA10* splice mutation c.471+2T→A, found in the three affected brothers results in a C-terminal truncated NAA10 protein, whereas the N-acetyltransferase catalytic domain is intact. The catalytic activity of NAA10 was never tested during this study, but it was speculated that the catalytic activity was still intact. Since the amounts of truncated NAA10 expressed in cells were found to be low, they proposed that the overall activity would be low as well. They concluded that the truncated C-terminus affects NAA10´s ability to interact with other proteins, resulting in dysregulation of important pathways such as WNT and the retinoic acid signaling pathway (Esmailpour et al., 2014).

Until 2015, researchers hypothesized that the reduction in catalytic activity could explain the severity of phenotypes in affected individuals. In 2015, a study conducted by Popp and colleagues identified two *de novo* missense mutations in exon 5 and 6 of the *NAA10* gene to be the cause of severe non-syndromic developmental delay. A male harboring the missense mutation R116W and an unrelated female harboring a F107V mutation had severe global delay. This was the first time a female showed a fully penetrant and severe phenotype. Functional analyses of the NAA10 R116W variant showed only a 15 % reduction in catalytic activity, whereas the NAA10 F107V variant showed a reduction of approximately 95 %. Both the male and female phenotypical features were characterized as severe, implicating that the phenotype and genotype of affected individuals could not be explained only by *in vitro* catalytic activity analysis (Popp et al., 2015).

A study by Casey et al. in 2015 supported this statement. Two brothers with intellectual disability and long QT and their mildly affected mother were all identified with the novel missense mutation Y43S in the *NAA10* gene. Functional tests revealed an 85 % reduction in catalytic activity. Even though these brothers had a NAA10 variant that exerted less catalytic activity compared to the NAA10 variant causing Ogden syndrome (20-80 % reduction) (Rope et al., 2011), the phenotypes of these two brothers were milder (Casey et al., 2015).

A study by Saunier and colleagues in 2016 investigated whether seven girls identified with the same *de novo* NAA10 R83C variant were accompanied by specific phenotypic aspects. In addition, four other girls harboring either V107F, F128I, F128L or R116W were studied in the investigation of potential shared phenotypic aspects. However, they could not deduce any obvious differences in phenotypical characteristics between the girls harboring the R83C variant or one of the other missense mutations. Hence, Saunier and colleagues concluded that it was the loss of NAT activity which is shared by all missense mutations (V107F, F128I, F128L, R116W and R83C) and not the impaired structural stability found for all mutations except R83C, that were the underlying reason for the similar phenotypes among the girls (Saunier et al., 2016). A study by Ree et al., 2019 of the novel NAA10 R83H variant detected in two unrelated boys share some of the phenotypes observed in the girls with the R83C mutation. The phenotypes observed in the girls carrying the R83C mutation include developmental delay, post-natal growth failure, intellectual disability and cardiac and skeletal abnormalities. The two boys harboring the R83H mutation were identified with intellectual disability, developmental delay cardiac abnormalities, ADHD-like behavior and limited

speech. Functional studies revealed reduced enzymatic activity of monomeric NAA10. *In silico* studies revealed that the reduction in monomeric activity was likely due to altered charge density in the Ac-CoA binding region of NAA10, resulting in an impaired enzyme-Ac-CoA binding.

In 2018, two distinct missense mutations with reduced stability and enzymatic activity of monomeric NAA10 and unaltered NatA activity were reported. One of the reported mutations was a novel *de novo* NAA10 V111G variant identified in a girl with mild/moderate non-syndromic intellectual disability (ID) and delayed motor and language development (McTiernan et al., 2018). The other undescribed NAA10 variant, I72T, was identified in three boys from two unrelated families with developmental delay, intellectual disability and cardiac abnormalities as overlapping phenotypes (Støve et al., 2018). The V111G variant was the first example of isolated NAA10 dysfunction in a case of ID, implicating that a degree of compromised NatA function might also be required in cases of syndromic ID. Both of these missense mutations broaden the phenotypic spectrum and adds to the evidence that genotype–phenotype correlations for NAA10 variants are much more complex than initially anticipated.

Cheng and Gottlieb published a comprehensive phenotypic and biochemical investigation of individuals with NAA10 and NAA15 mutations in 2019. This study contains 30 individuals from unrelated families with 17 different *de novo* or inherited *NAA10* or *NAA15* missense mutations. 23 of these were previously undescribed individuals with *NAA10*-related syndrome. The study investigated the variants´ impact on NatA catalytic activity in presence and absence of HYPK as well as the NatA thermostability in presence and absence of acetyl-CoA (Cheng et al., 2019). They found that when HYPK is present the enzymatic activities relative to their activity when HYPK is absent were potentiated, decreased or nullified depending on the specific NAA10 variant. The study comparing the thermostability of WT NatA complexes with mutant NatA complexes in the absence and presence of acetyl-CoA showed that all of the NAA10 missense variants except I72T, R83C and M147T were destabilized in the absence of acetyl-CoA compared to WT. By contrast, all variants, except I72T, were destabilized in the presence of acetyl-CoA relative to WT. Cheng et al. propose that these findings indicate that NAA10 variants are overall destabilizing for the NatA complex, wherein NAA10 variants that affect the GNAT fold reduce stability of the complex in the absence as well as the presence of acetyl-CoA (Cheng et al., 2019).

The field is expanding as new NAA10 variants are being reported more frequently. The missense mutations reported to date are listed in Table 2.2. We still do not know much about the underlying mechanisms causing NAA10-related disorder. It can not be said for sure if it is the loss of Nt-acetylation, KAT acetylation or perhaps non-catalytic functions that cause the pathological effects found in patients. The various mutations seem to affect NAA10 differently. Some variants have functional NatA activity but reduced monomeric activity, while other variants have an intact monomeric activity wherein the NatA activity is affected.

Table 2.2: Effects on protein function and phenotypic features of NAA10 mutations in patients

DNA nucleotide change	Predicted protein change	Gender	Function	Phenotype	References
c.29 A>G	p.Asp10Gly	1 male	Reduced catalytic activity and protein stability	Severe DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, delayed motor development, muscular hypotonia, ptosis, hearing loss, hypertension, agenesis and hypoplasia of corpus callosum, aged appearance, lethal cardiomyopathy.	Cheng et al., 2019
c.32 T>G	p.Leu11Arg	1 female	Reduced catalytic activity and protein stability	Hydrocephalus due to a large arachnoid cyst between the cerebellum and the occipital area, ventriculomegaly, hypoplasia, hyperplasia, partial agenesis, delayed motor and speech development, strabismus, feeding difficulties.	Cheng et al., 2019
c.109 T>C	p.Ser37Pro	8 males, 5 female carriers	Reduced catalytic activity, impaired complex formation and substrate binding	Severe DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, truncal hypotonia, aged appearance, prominent eyes, scoliosis, lethal cardiac arrhythmia and cardiomyopathy.	Rope et al., 2011, Myklebust et al., 2015
c.128 A>C	p.Tyr43Ser	2 males	Reduced catalytic activity and protein stability. Protein aggregation.	Mild DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, scoliosis, hypotonia and prolonged QT.	Casey et al., 2015
c.215 T>C	p.Ile72Thr	3 males	Reduced monomeric catalytic activity and unstable.	Coarse facial features, tachycardia, prolonged QT, lethal cardiomyopathy.	Støve et al., 2018, Cheng et al., 2019
c.247 C>T	p.Arg83Cys	18 females, 3 males	Reduced catalytic activity and impaired binding of Ac-CoA	DD, growth deficiency, post-natal growth failure, ID, cardiac and skeletal abnormalities.	Saunier et al., 2016, Cheng et al., 2019
c.248 G>A	p.Arg83His	2 males	Reduced monomeric activity and possible impaired substrate binding.	ID, DD, limited speech, ADHD-like behavior and cardiac abnormalities	Ree et al., 2019
c.259 G>T	p.Ala87Ser	3 females	Only clinically characterized	DD, growth deficiency, post-natal growth failure, epilepsy, seizures, ADHD, ASD, speech and motor delay, bilateral hearing impairment, muscle hypotonia, optic nerve hypoplasia, moderate facial features, elevated alkaline phosphatase levels, scoliosis, abnormal behavior.	Cheng et al., 2019
c.311 C>A	p.Ala104Asp	1 female	Reduced catalytic activity and protein stability.	DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, ADHD, sensory processing disorder, language disorder, hypotonia, motor delay, short stature, anxiety, astigmatism, anisometropia conjunctivae, sleep disorder, ID.	Cheng et al., 2019

c.319 G>T	p.Val107Phe	1 female	Reduced catalytic activity and reduced protein stability.	DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, minor facial features, self-stimulatory behavior and hypotonia.	Popp et al., 2015, Saunier et al., 2016
c.332 T>G	p.Val111Gly	1 female	Reduced catalytic activity and protein stability.	Mild DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, delayed language- and motor development.	McTiernan et al., 2018
c.346 C>T	p.Arg116Trp	1 male, 1 female	Minor reduced catalytic activity and impaired binding to imprinting control regions.	DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, hypotonia, minor facial features, behavioral anomalies.	Popp et al., 2015, Saunier et al., 2016
c.361 C>G	p.Leu121Val	2 females	Only clinically characterized	DD,, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, hip dysplasia, moderate/ severe ID, mild/severe language development, motor delay, severe ASD, lacks eye contact, self-stimulatory behavior, sensory issues.	Cheng et al., 2019
c.382 T>A	p.Phe128Ile	2 female	Reduced catalytic activity and reduced protein stability.	DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, limited mobility- and speech.	Saunier et al., 2016
c.384 T>A/G	p.Phe128Leu	3 females	Reduced catalytic activity and protein stability. Protein aggregation.	Severe DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, early feeding problems, cortical vision impairment, delayed motor development, ventriculomegaly, chorea, overriding toes and single palmar crease, behavior issues.	Saunier et al., 2016, Cheng et al., 2019
c.440 T>C	p.Met147Thr	1 female	Reduced catalytic activity in presence of HYPK. Only decreased stability in the presence of acetyl-CoA.	DD, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, sensory processing disorder, self-stimulatory behavior, thinning corpus callosum, microcephaly, stigmatism, acne, body odor, cortical visual impairment.	Cheng et al., 2019
c.471+2 T>A	Truncated protein	8 males	C-terminal truncation, impaired PPI, dysregulation of pathways	Retinol uptake deficiencies, microphthalmia/anophthalmia, skeletal abnormalities, genitourinary malformations, anomalies of the digits, ears, and teeth. ID, seizure disorder.	Esmailpour et al., 2014
c.455_458 del CTCA	p.Thr152fs	1 male	Removes the C-terminus, but leaves the acetyltransferase domain intact, substantially lower expression levels.	DD,, growth deficiency, post-natal growth failure, cardiac and skeletal abnormalities, agenesis of corpus collosum, craniosynostosis, severe ID, hypotonia, microcornea, microphthalmia, scoliosis, pectus excavatum, equinovarus, atrial septal defect, severely underweight, abnormal behavior and no speech development.	Cheng et al., 2019

DD = developmental delay, QT = Q wave – T wave, ID = intellectual disability, ADHD = attention deficit hyperactivity disorder, ASD = Atrial Septal Defect, PPI = protein-protein interaction.

2.4 Aim of study

In recent years, research has offered new insight and a better understanding of NAA10 structure and function, along with the pathological consequences of NAA10 deficiency in humans. After the discovery of Ogden syndrome, which was the first genetic disorder coupled to Nt-acetylation, several distinct *NAA10* mutations have been identified as the cause of disease in patients (Esmailpour et al., 2014; Popp et al., 2014; Casey et al., 2015; Saunier et al., 2016; Mc Tiernan et al., 2018; Støve et al., 2018; Ree et al., 2019; Cheng et al., 2019). Patients identified with NAA10 mutations share some phenotypic characteristics, among which are developmental delay, intellectual disability (ID), hypotonia, delayed speech and growth development, cardiac arrhythmia and various skeletal, organ and morphological anomalies. The interest of studying NAA10 abnormalities has increased the last eight years, however there are many aspects of the physiological and pathological consequences of *NAA10* mutations that still remain elusive. In the field of genetic disorders coupled to Nt-acetylation, functional studies will play an important role in the attempt to answer the questions concerning them.

Recently, two novel *NAA10* mutations were identified in two female patients with varying degree of delayed motor and language development and syndromic intellectual disability. One female was found to harbor the NAA10 L11R variant, and the other female was identified with the NAA10 H16P variant. The aim of this thesis was to functionally characterize these two novel missense mutations and investigate whether these particular mutations could be the cause of the pathological effects detected in the affected patients. In order to investigate whether these mutations have an impact on NAA10 function in terms of complex formation, catalytic activity and stability the following approaches were made: constructs encoding NAA10 with the desired mutations were generated before being transfected into human cells to undergo cellular stability analysis, and to measure NatA complex formation and catalytic activity. Disease predictions, conservation assessments and structural examinations were performed to investigate the mutations *in silico* throughout the study.

3 Materials

3.1 Antibodies

Table 3.2: Primary antibodies used for Western blotting and immunoprecipitation

Name	Description	Application	Concentration (µg/µl)	Dilution used	Supplier	Cat. #
anti-β-tubulin	Mouse, monoclonal	WB analysis of stability assay	4.5	1:3000	Sigma	T293
anti-NAA10	Rabbit, polyclonal	WB analysis of IP	0.73	1:3000	Abcam	ab155687
anti-NAA15	Rabbit, polyclonal	WB analysis of IP	1.0	1:2000	Biogenes *	
anti-V5	Mouse, monoclonal	WB analysis stability assay and IP	1.1	1:5000	Invitrogen	46-0705

* Arnesen et al., 2005a.

Table 3.2: Secondary antibodies used for Western blotting

Name (antigen)	Description	Concentration (µg/µl)	Dilution used	Supplier	Cat. #
ECL Anti-Rabbit IgG, HRP linked ab	Donkey	1.0	1:5000	GE Healthcare	NA934
ECL Anti-Mouse IgG, HRP linked ab	Sheep	1.0	1:5000	GE Healthcare	NA931

3.2 Apparatus and instruments

Table 3.3: Apparatus and instruments

Name	Use	Supplier
ChemiDoc™ XRS+	Visualizing Western blots	BioRad
Heating block, AccuBlock™ Digital Dry Bath	Heating of media and solution for cell work	Labnet International Inc.
Nano drop ND-1000	Measure DNA and protein absorbance	Saveen Werner
Thermal cycler	Site-directed mutagenesis and sequencing PCR	Eppendorf
Thermo shaker	[^{14}C]-Ac-CoA-based *in vitro* Nt-acetylation assay	Eppendorf
Tri-Carb 3100TR Liquid Scintillation Analyzer	Measure ^{14}C-signal	Packard Bioscience company

Table 3.4: Centrifuges

Name	Supplier
Eppendorf centrifuge 5810 R	Eppendorf
Thermo scientific Heraeus fresco 17	Thermo Fisher Scientific

3.3 Bacterial strain

Table 3.5: Bacterial strain used in site-directed mutagenesis and protein expression

Strain	Use	Supplier	Cat. #
NEB® 5-alpha Competent *E. Coli*	Site-directed mutagenesis	New England Biolabs	C2987H

3.4 Buffers, solutions and media

3.4.1 Buffers and solutions for bacterial work

LB-media

- 10 g/l Tryptone
- 5 g/l Yeast extract
- 5 g/l Sodium chloride
- 100 µg/µl Ampicillin

S.O.C. medium

- 2 % (w/v) Tryptone
- 0.5 % /w/v) Yeast extract
- 10 mM NaCl
- 2.5 mM $MgCl_2$
- 10 mM KCl
- 10 mM $MgCl_2$
- 10 mM $MgSO^4$
- 20 mM glucose

LB-Agar

- 15 g/l granulated agar
- 5 g/l Tryptone
- 5 g/l Yeast extract
- 10 g/l Sodium chloride
- 100 µg/µl Ampicillin

3.4.2 Buffers and solutions used for cycloheximide chase experiments, immunoprecipitation and ^{14}C-acetylation assay

Dulbecco´s Modified Eagle´s Medium (DMEM)

- Supplier: Sigma-Aldrich (D6546)
- 4500 mg/l glucose, sodium pyruvate and sodium bicarbonat
- 10 % FBS
- 3 % L-glutamine
- 1 % Penicillin Streptomycin (Pen Strep)

IPH lysis buffer

- 50 mM Tris-HCl
- 150 mM NaCl
- 5 mM EDTA
- 0.5 % NP-40
- 1x c0mplete EDTA free protease inhibitor
- pH 8.0

IPH wash buffer

- 50 mM Tris-HCl
- 300 mM NaCl
- 5 mM EDTA
- 0.5 % NP-40
- 1x c0mplete EDTA free protease inhibitor
- pH 8.0

2x Acetylation buffer

- 100 mM Tris-HCl
- 2 mM EDTA
- 20 % Glycerol
- pH 8.0

HEPES

- 10 mM HEPES
- pH 7.4

3.4.3 Buffers and solutions used in SDS-PAGE and Western blotting

1x SDS electrophoresis buffer

- 10 % BioRad 10x TGS

1x Towbin buffer

- 10 % BioRad 10x TG
- 20 % Methanol

Ponceau S solution

- 0.1 % (w/v) in 5 % acetic acid

10x Phosphate buffered saline (PBS)

- 80 g/l NaCl
- 2 g/l KCl
- 20 g/l Na_2HPO^4 x $2H_2O$
- 4 g/l KH_2PO_2
- pH 7.4

4x Sample buffer

- 62.5 mM Tris-HCl, pH 6.8
- 10 % glycerol
- 1% LDS
- 0.005% Bromophenol Blue
- dithiothreitol (DTT) to a final concentration of 50 mM

1x PBS-Tween (PBST)

- 1x PBS
- 0.1 % (v/v) Tween-20

Dry milk

- 5 % (w/v) dry milk
- 1x PBST

3.5 Chemicals and reagents

Table 3.6: Various chemicals and reagents

Supplier	Name	Cat. #
AppliChem GmbH	Dry milk powder	A0830
Bio-Rad	Laemmli™ SDS sample buffer	1610747
Lonze	L-glutamine	BE17-605E
	Penicillin Streptomycin (Pen Strep)	DE17-602E
Merck	Ethanol	
	Ethylenediaminetetraacetic acid (EDTA)	324503
	Isopropanol	33539-1L-M
	Methanol	32213
	Ponceau S solution	P7170
	Potassium dihydrogen phosphate (KH_2PO_4)	529568
	Tris	108382
	Tryptone	107213
New England Biolabs	S.O.C medium	B9020S
Roche	C0mplete EDTA free protease inhibitor cocktail tablet	11 873 580 001
	X-tremeGENE 9 DNA Transfection Reagent	06 365 809 001
Sigma-Aldrich	Agar	A1296
	Acetyl-CoA	10101893001
	Cycloheximide	C4859
	Diethiothreitol (DTT)	43816-250ML
	Dulbecco´s Modified Eagle´s medium (DMEM)	D6546-500ML
	Fetal bovine serum (FBS)	F752
	Glycerol	G5516-500ML
	HEPES	18896-100ML
	ICEPAL® CA-630 (NP-40)	18896-100ML
	Imidazole	
	Potassium chloride (KCl)	P9541-1KG
	Sodium chloride (NaCl)	31434-1KG
	Sodium dodecyl sulfate (SDS)	74255-250G
	Sodium phosphate dibasic ($Na_2HPO_4x2H_2O$)	
	Tween®-20	9005-64-5
	Yeast extract	89526
Thermo Scientific	Gibco™ Opti-MEM®	31985-070
	Gibco™ Trypsin	15400-054
	Dynabeads Protein G magnetics	10003D
VWR Chemicals	Ampicillin sodium salt	0339-EU 100g

3.6 Commercial kits

Table 3.7: Commercial kits used in various methods

Name	Use	Supplier	Cat. #
Big Dye Terminator v3.1 cycle sequencing kit	DNA sequencing	Applied Biosystems	4337456
GeneJET Plasmid Miniprep Kit	Miniprep	Thermo Scientific	K0503
Mini-PROTEAN TGX Precast Gel	SDS-PAGE	Bio-Rad	4568104/4568106
Nucleobond Xtra midi	Midiprep	Macherey-Nagel	740410.100
Q5 Site Directed Mutagenesis Kit	Mutagenesis	New England Biolabs	E0554S
SuperSignal West Pico Chemiluminescent Substrate	Protein detection (Western blot)	Thermo Scientific	34080
10x Tris/Glycine (TG)	Protein transfer (Western blot)	Bio-Rad	161-0734
10x Tris/Glycine/SDS (TGS)	SDS-PAGE	Bio-Rad	161-0772

3.7 Human cell line

Table 3.8: Human cell line

Cell line	Species	Tissue	Cell type	Supplier
HeLa	Human	Cervix	Adenocarcinoma	ATCC

3.8 Molecular size marker

Table 3.9: Molecular size marker used in SDS-PAGE

Name	Description	Range	Supplier	Cat. #
PageRuler^{TM} Plus Prestained Protein Ladder	Molecular size marker for proteins	10-250 kDa	Thermo Scientific	26619

3.9 Peptides

Table 3.10: Peptides used in ^{14}C-Nt-acetylation assay

Name	[H]-Sequence-[OH]	Derived from	Supplier
EEEI	[H] EEEIAAL RWGRPVGRRRRPVRVYP [OH]	γ-actin	Biogenes
SESS	[H] SESSSKS RWGRPVGRRRRPVRVYP [OH]	High mobility group protein A1 (HMGA1)	Biogenes

3.10 Plasmids

Table 3.11: NAA10-V5 plasmid constructs

Name	Protein	Properties	Expression	Supplier
pcDNA 3.1/V5 His TOPO	hNAA10-V5-His	Ampicillin resistance, 6x His-tag, C-terminal V5	Mammalian	Invitrogen/ T.Arnesen
pcDNA 3.1/V5 His NEB-NAA10-H16P	hNAA10-His16Pro-V5-His	Ampicillin resistance, 6x His-tag, C-terminal V5	Mammalian	Made as part of this thesis
pcDNA 3.1/V5 His NEB-NAA10-L11R	hNAA10-Leu11Arg-V5-His	Ampicillin resistance, 6x His-tag, C-terminal V5	Mammalian	Made as part of this thesis

3.11 Primers

Table 3.12: Primers used in site-directed mutagenesis

Name	Sequence (5´-3´)	Tm (°C)	Supplier
NAA10 A47C p.H16P fwd.	TAATACGACTCACTATAGGG	50.9	Sigma-Aldrich
NAA10 A47C p.H16P rev.	TCACGTAGAATCGAGACCGAGG	67.5	Sigma-Aldrich
NAA10 T32G p.L11R fwd.	GATGAAGCCCTGAAAGACGCGCAG	75.2	Sigma-Aldrich
NAA10 T32G p.L11R rev.	GCTAGTTATTGCTCAGCGG	60.2	Sigma-Aldrich

3.12 Software

Table 3.13: List of software used

Name	Use	Supplier
Image Lab™ 6.0.1	Imaging of SDS-PA gels and membranes and quantification of protein bands	Bio-Rad Laboratories
PyMOL v2.3.2	Structural analysis and creating structural figures	Schrödinger
ND-1000 v3.5.2	Measure plasmid DNA and protein concentration	Nanodrop Tchnologies Inc.
SnapGene v.4.0	Visualization of vector maps	GSL Biotech

3.13 Web resources

Table 3.14: Online resources used for bioinformatic analyses

Name	Web link	Use
ChemSpider	http://www.chemspider.com	Illustrations of chemical structures
ClustalOmega	http://www.ebi.ac.uk/Tools/msa/clustalo/	Multiple sequence alignment
CUPSAT	http://cupsat.tu-bs.de	Missense mutation prediction
DUET	http://biosig.unimelb.edu.au/duet/	Missense mutation prediction
DynaMut	http://biosig.unimelb.edu.au/dynamut/prediction	Missense mutation prediction
ESPript ENDscript	http://espript.ibcp.fr/ESPript/ESPript/ http://endscript.ibcp.fr/ESPript/ENDscript/	Annotating multiple sequence alignment with similarities and secondary structure elements
HOPE	https://www3.cmbi.umcn.nl/hope/input/	Missense mutation prediction
PMut	http://mmb.irbbarcelona.org/PMut.	Missense mutation prediction
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/	Missense mutation prediction
PROVEAN		Missense mutation prediction
SIFT	http://sift.bii.a-star.edu.sg/www/SIFT_seq_submit2.html	Missense mutation prediction
Smart Servier medical art	http://smart.servier.com/	Method illustrations
SNPs&GO^{3D}	http://snps.biofold.org/snps-and-go/snps-and-go-3d.html	Missense mutation prediction
UniProt	https://www.uniprot.org	Sequence information

4 Methods

A schematic outline of the workflow in this thesis is shown in Figure 4.1.

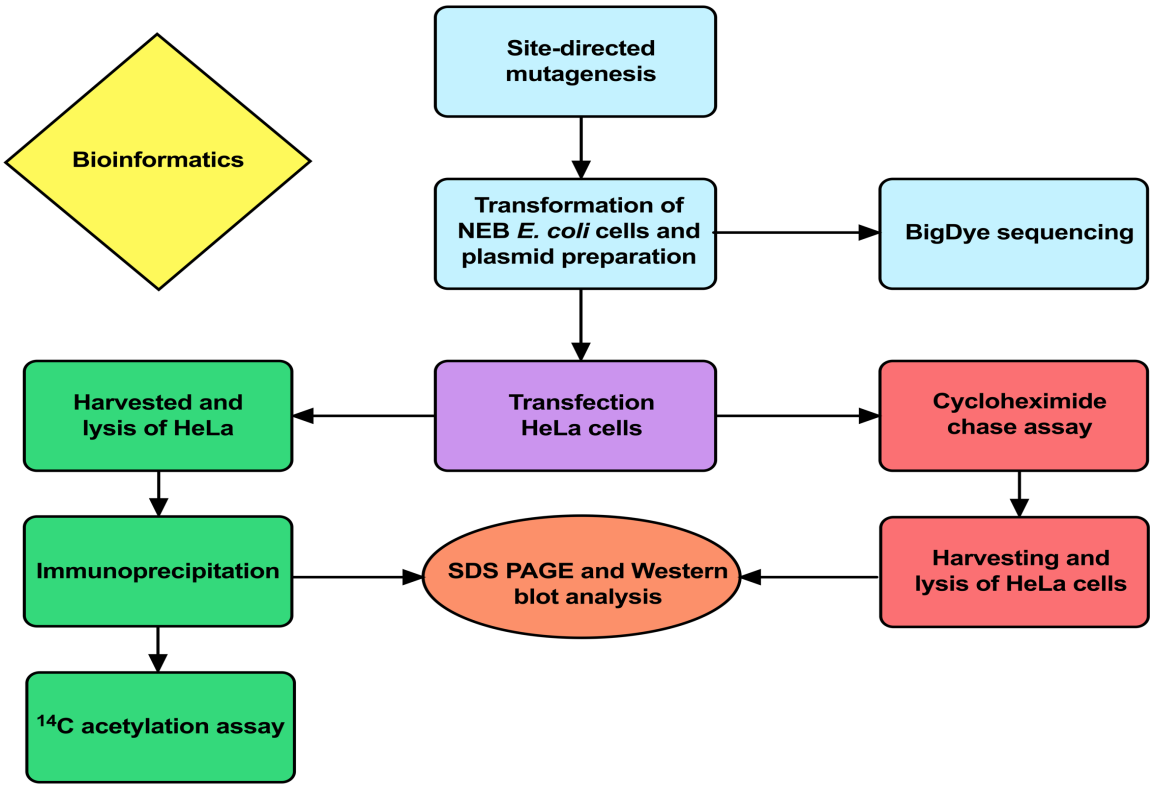

Figure 4.1: Schematic outline of the main steps of methods applied in this thesis. The initial methods applied in DNA work (blue) were used to generate constructs encoding NAA10 with the desired mutations and introduce these constructs into bacteria for plasmid preparation**.** After transfecting HeLa cells with the desired constructs, cellular stability of the expressed NAA10-V5 (red) and NatA complex formation and intrinsic catalytic activity (green) was investigated. Samples were analyzed by SDS-PAGE and Western blot (orange). Bioinformatic analyses were performed throughout the study.

4.1 Bioinformatics analysis

4.1.1 Multiple sequence alignment (MSA)

Multiple sequence alignment (MSA) is an *in silico* tool used for studying closely related genes and proteins. MSA compares homologous sequences and enables phylogenetic reconstruction, protein secondary and tertiary structure prediction and protein function prediction (Kemena et al., 2009)**.** In order to create a multiple sequence alignment for studying the amino acid conservation in orthologous NAA10 sequences, the alignment tool Clustal Omega (Sievers et al., 2011) with default settings was used. NAA10 protein sequences from *Homo Sapiens, Mus musculus, Rattus norvegicus, Xenopus laevis, Caenorhabditis elegans, Danio rerio, Arabidopsis thaliana* and *Saccharomyces pombe* was retrieved from UniProt (UniProt Consortium, 2018). Furthermore, ESPript 3.0 (Robert et al.,

2014) was used to annotate the MSA with conservation and secondary structure illustrations derived from a solved human NatA structure (PDB: 6C9M) (Gottlieb et al., 2018).

4.1.2 *In silico* missense mutation prediction

In order to retrieve consistent results of the potential impact of each missense mutation, L11R and H16P in NAA10, several *in silico* online prediction tools were used in this thesis. The NAA10 sequence and NatA structure (PDB: 6C9M, Gottlieb et al., 2018) and default settings were used in all prediction tools to assess the impact of each mutation.

The pathological impact of the missense mutations was predicted with SNPs&GO^{3D} (Capriotti et al., 2013) and PMut (López-Ferrando et al., 2017). SNPs&GO^{3D} uses protein tertiary structure, sequence profile and functional information to predict disease-associated variations. Each mutation is given a probability score ranging from 0 to 1. If the probability score is above 0.5, the mutation is predicted to be disease-associated. PMut uses sequence conservation and predicted physicochemical properties to predict the pathological impact. Each mutation is given a probability score ranging from 0 to 1. Mutations are classified as pathological if the probability is 0.5 or higher.

PolyPhen-2 (Adzhubei et al., 2010), SIFT (Sim et al., 2012) and PROVEAN (Choi et al., 2012) were used to predict the possible impact of missense mutations on protein function. PolyPhen-2 uses structure- and sequence-based features to predict the outcome of a missense mutation by comparing how well the two human alleles fit into the pattern of amino acid replacements within an MSA of homologous proteins. Each mutation is given a probability score ranging from 0 to 1. A mutation with a score close to 0 is predicted to be benign, whereas a score close to 1 is predicted to be deleterious and classified as “probably damaging”. SIFT (Sorting Intolerant From Tolerant) employs sequence homology and the physical properties of amino acids in their predictions. The probability score is presented such as PolyPhen-2 (scores ranging from 0.0-1.0), but proteins are considered deleterious with a probability score below 0.05 and tolerated if the probability score is above 0.5. PROVEAN (Protein Variation Effect Analyzer) uses homology and clustering to filter sequence variants to identify nonsynonymous or indel variants that are predicted to be functionally important. The closest related sequences get scores that are averaged within and across clusters to generate the final PROVEAN score. If the PROVEAN score is equal to or below a predefined threshold of -2.5, the protein variant is predicted to be deleterious, while a score above the threshold is characterized as neutral.

HOPE (Venselaar et al., 2010), DUET (Pires et al., 2014), CUPSAT (Gromiha et al., 2007) and DynaMut (Rodrigues et al., 2018) were used to assess insight into structural effects and the possible impact of mutations on protein stability. HOPE provides insight into structural effects of a mutation by collecting and combining structural information and sequence annotations from a series of web services and present the predictions in a report, complete with results, figures and animations. DUET predicts the effect of missense mutations based on

structural signatures and amino acids properties from homologous proteins. The stability change of a protein is presented in Gibbs free energy (ΔΔG). A positive value will indicate a stabilizing missense mutation, whereas a negative value will indicate a destabilizing missense mutation. CUPSAT predicts the changes in a protein upon point mutations by using the amino acid-atom potentials and torsion angle distribution to assess the amino acid environment of the mutation site. The stability change of a protein is presented as in DUET. The last tool used to assess the impact of mutations on protein stability and dynamics was DynaMut. DynaMut predicts the effects of a mutation by gathering structural information from the 3D space of a natively folded protein and investigate protein motions upon mutation resulting from vibrational entropy changes. The prediction outcomes are presented as ΔΔG, the vibrational entropy difference (ΔΔS) between wild-type and mutant structures and 3D visual representation of interatomic interactions, deformation and fluctuation analysis.

4.1.3 *In silico* structural illustration and predictions

The structures of human NatA (PDB: 6C9M, Gottlieb et. al., 2018) and *S. pombe* NAA10 (PBD: 4KVM, Liszczak et al., 2013) were retrieved from the RCSB protein data bank (Berman et al., 2000) and uploaded to the molecular visualization program, PyMOL. In order to investigate the structural relationships between the amino acids around the mutated sites, NAA10 of structure 4KVM was aligned to the human structure 6C9M. This alignment allowed an approximate placement of the bisubstrate inhibitor within the human NAA10 binding site, imitating the Ac-CoA- peptide substrate of NAA10. In order to get a visual impression of how the mutations might affect overall structure and substrate binding, an *in silico* mutagenesis of L11 to arginine and H16 to proline was also conducted by using PyMOL.

4.2 Mutagenesis, plasmid preparation and DNA sequencing

NEB® 5-alpha Competent *E. Coli* cells were chosen for plasmid propagation because they are ideal for subcloning efficiency transformation and high-quality plasmid preparations. NEB® 5-alpha Competent *E. Coli* is a derivate of DH5 cells. They are T1 phage resistant and the activity of nonspecific endonuclease 1 (*endA1*) is eliminated for highest quality plasmid preparations. *NAA10* was expressed in the pcDNA3.1 plasmid under the control of the cytomegalovirus (CMV) promotor. NAA10 was expressed as a fusion protein with a V5- and 6xHis-tag on the C-terminus. The V5 tag was used for immunoprecipitation and detection of NAA10-V5 variants by Western blot analysis.

4.2.1 Introducing a missense mutation in *NAA10* using site directed mutagenesis

To investigate the different point mutations in NAA10, site directed mutagenesis was used to introduce the desired point mutations into the mammalian vector pcDNA3.1/*NAA10*-V5-His.

Site directed mutagenesis is an *in vitro* technique used to make specific DNA alterations such as insertions, deletions and substitutions. The method takes advantage of explicit designed oligonucleotides to introduce a desired mutation in a double-stranded DNA plasmid. In this thesis, back-to-back primers were designed in order to create the desired missense mutations. Overlapping primers generate products that re-circulate and forms doubly nicked plasmids. By using back-to-back primers this is avoided. Back-to-back primers create non-nicked plasmids which enables exponential amplification and thereby generates significantly more of the desired product. The desired mutations were incorporated into plasmid DNA by substitution through the use of specifically designed forward and reverse primers. The mutagenesis and transformation of NEB® 5-alpha Competent *E. Coli* cells were performed by the use of Q5® Site Directed Mutagenesis Kit (New England Biolabs) according to the manufacturers´ protocol.

4.2.2 Plasmid preparation and plasmid DNA isolation

Colonies from the transformed NEB® 5-alpha Competent *E. coli* cells transformed with pcDNA3.1/*NAA10*-V5-His plasmids were inoculated in 5 ml Luria-Bertani (LB) media containing 100 µg/µl ampicillin. The culture was incubated at 37 °C with shaking at 250 rpm overnight. The following day, the pre-culture was cultivated in 100 ml LB-media containing 100 µg/µl ampicillin. The culture was incubated at 37 °C with shaking at 250 rpm overnight. The cells were harvested by centrifugation at 4 °C at 3220 x *g* for 20 min and the pcDNA3.1 plasmids were isolated using the Nucleobond Xtra Midi Kit (Macherey-Nagel) as described by the manufacturers´ protocol. The DNA concentrations of the resulting eluates were determined by measuring absorbance at 280 nm using a Nanodrop spectrophotometer, and the sequences of the mutated *NAA10* plasmids were verified by DNA sequencing.

4.2.3 Verifying mutation by DNA sequencing

To verify if the site directed mutagenesis of the plasmids had been successful, the BigDye Terminator v3.1 sequencing Kit was used to determine the plasmid DNA sequences. This kit is based on the Sanger sequencing technique which is based on the selective incorporation of chain-terminating dideoxy nucleotides (ddNTPs) by DNA polymerase during replication. These ddNTPs lacks the hydroxyl group on the 3´carbon required for the formation of a phosphodiester bond between two nucleotides. Once the ddNTP is incorporated, the extension of DNA is ceased. This results in DNA strands of different lengths ending in a ddNTP. The four ddNTP (ddATP, ddGTP, ddCTP and ddTTP) are labeled with an individual color of dye depending on the base that it carries. The DNA strands are separated by capillary electrophoresis based on their size, and the position and terminating base are determined by the labels (Sanger, 2004).

The BigDye version 3.1 sequencing kit was used in order to verify if the newly generated plasmid had obtained the correct mutation. All sequencing reactions were made with the

components listed in Table 4.1. The sequencing was performed using a thermal cycler with the sequencing program described in Table 4.2. Before delivering the sequencing reactions to the DNA sequencing facility at the University of Bergen, 10 µl ddH_2O was added to the sequencing reactions.

Table 4.1: Sequencing reaction mix

Reagent	Amount
BigDye version 3.1	1 µl
Sequencing buffer	1 µl
Template	200 ng
Primer	3.2 pmol
ddH_2O	quantum satis
Total	10 µl

Table 4.2: Sequencing program

Step	Cycles	Time
1	1	96 °C, 5 min
2	25 cycles	96 °C0 sec 50 °C, 5 sec 60 °C, 4 min
3	1	4 °C, ∞

4.3 Maintenance and transfection of HeLa cells and NAA10 stability analysis

4.3.1 Maintenance of HeLa cells

HeLa cells were maintained in a HERA cell incubator with constant 5 % CO_2 and 37 °C in Dulbecco´s modified Eagle´s medium ((DMEM), Sigma) supplemented with 10 % fetal bovine serum (FBS), 3 % L-Glutamine and 1 % PenStrep. When the cells reached confluency, they were washed with 1x PBS and trypsinized at 37 °C for 5-8 min to detach the adherent cells from the surface. Trypsin was quenched by adding fresh growth medium. Cells were typically split 1:5, 1:8 or 1:10 in a 10 cm dish for further subculturing. 0.3 x 10^6 cells per well or 2 x 10^6 cells per 10 cm dish were seeded one day prior to transfection for cycloheximide chase experiments and immunoprecipitation respectively. In order to avoid contamination, all cell cultures were handled under sterile conditions.

4.3.2 Transfection of HeLa cells

The process of introducing foreign nucleic acids such as DNA, RNA or oligonucleotides into eukaryotic cells is called transfection. Transfection is a powerful analytical tool for studying gene function and regulation and protein functions. In these experiments, the HeLa cell line, which is derived from cervical cancer cells, was used because it is durable and easily transfected. 0.3 x 10^6 HeLa cells were seeded in each well in four 6-well plates 16-18 h before transfection. In order to maximize the transfection efficiency, cells were transfected in the log-phase of growth at approximately 70-80 % confluency. Three wells in each 6-well plate were transfected with 1.2 µg of NAA10 WT-V5 or 1.8 µg of NAA10-V5 variants using

XtremeGene 9 transfection reagent (Roche). 0.6 µg empty V5-vector was co-transfected with NAA10 WT-V5 to ensure equal conditions for the cells. Plasmid and XtremeGene 9 transfection reagent were added in a 1:3 ratio (µg:µl) with Gibco Opti-MEM to a final volume of 150 µl per well. After incubation for 15 min, the transfection mix was added dropwise to the cells. To minimize cell death due to the cytotoxic transfection reagent, the DMEM growth medium was replaced 4 h post transfection.

4.3.3 Determine half-life of NAA10 variants using cycloheximide chase assay

In order to investigate the turnover of NAA10 WT-V5 and each NAA10-V5 variant, cycloheximide chase assay were performed. Cycloheximide chase assay is a method used to determine the half-life of a protein. Cycloheximide is an inhibitor of protein biosynthesis that prevents eukaryotic cytosolic translational elongation by ribosomes. This method enables analysis of protein half-life without the interference of translation of new proteins (Buchanan et al., 2016). Cells are exposed to cycloheximide and collected at specific time points. Since an abnormal protein may be more stable or degraded more rapidly compared to native proteins, cycloheximide chase experiments can be used to analyze protein turnover and indirectly measure the stability of a protein. In this experiment, HeLa cells overexpressing NAA10 WT-V5 and NAA10-V5 variants were treated with cycloheximide for six hours. Cells were harvested at specific time points (0, 2, 4, and 6 h) and lysed. Proteins were separated by SDS-PAGE and protein abundance at each time point was analyzed by Western blot.

48 h post transfection, the growth medium of each well was replaced with 2 ml medium containing 50 µg/ml cycloheximide. 0, 2, 4, and 6 h after adding cycloheximide, cells were washed and harvested in 1200 µl cold 1x PBS. Cells harvested at time point 0 h were not treated with cycloheximide. Cells were centrifuged at 4 °C and 1000 x *g* for 5 min. The supernatant was discarded, and the cell pellets were resuspended in 300 µl cold 1x PBS. Cells were spun down under same conditions above and the supernatant was removed. The resulting cell pellets were stored at -80 °C.

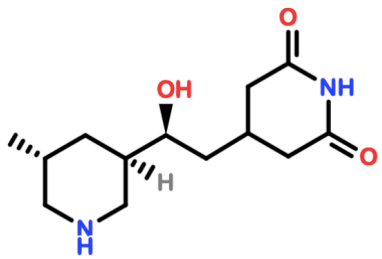

Figure 4.2: Cycloheximide (CHX) inhibits eukaryotic protein synthesis. Cycloheximide inhibits the translational elongation of peptides by associating with the 60S ribosomal unit E site in eukaryotic cells (Buchanan et al., 2016)

4.3.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is an analytical method for separating proteins according to their relative molecular mass. This method uses a polyacrylamide gel as support medium and sodium dodecyl sulphate (SDS) to denature the proteins. Protein samples are mixed with sample buffer containing dithiothreitol (DTT), SDS, glycerol and bromophenol blue. The anionic detergent SDS and reducing agent DTT disrupts the secondary and tertiary structure of proteins, resulting in linear peptides. SDS introduces an overall negative charge proportional to its relative molecular mass. Proteins are separated in a polyacrylamide gel by applying an electric field. The polyacrylamide gel is a mesh-like matrix containing pores causing larger proteins to encounter more resistance and therefore migrating slower than smaller proteins.

Samples were mixed with 4x sample buffer and denatured on a heat block at 95 °C for 5 min. Subsequently, samples were loaded onto an 8-16 % Mini-PROTEAN TGX Precast Gel (Bio-Rad). PageRuler Plus Prestained Protein Ladder (Thermo Scientific) was used as molecular weight marker. The gel was run in 1x electrophoresis running buffer (1x TGS buffer) at 100 V for 10 min and then at 200 V until the proteins were sufficiently separated. Afterwards, the proteins were transferred to a membrane for total protein detection (Western blotting, section 4.3.5).

4.3.5 Western blot (WB) analysis

With the intention of investigating the protein stability of NAA10, protein abundance at each time point was analyzed by Western blot (WB). WB is used to identify specific proteins in a complex mixture of proteins extracted from cells. Proteins are first separated by SDS-PAGE before being electrophoretically transferred onto a nitrocellulose membrane by the use of a transfer sandwich. The transfer sandwich consists of sponges at each end and filter papers to protect the gel and blotting membrane. The blotting membrane is placed between the gel and the positively charged electrode, enabling the negatively charged proteins to migrate towards the positively charged electrode and onto the membrane (Towbin et al., 1979). Finally, by the use of either fluorophore- or enzyme-conjugated antibodies, staining or radioactive labeling the protein of interest can be detected and visualized on the membrane.

Cell pellets were resuspended in 40 µl IPH lysis buffer and incubated on ice for 20 min. Afterwards, the cell debris was pelleted by centrifugation at 17000 x *g* and 4 °C for 5 min. The supernatants were transferred to new tubes and 4x sample buffer was added. SDS-PAGE (Methods, section 4.3.4) was used to separate the proteins before they were electrophoretically transferred to a membrane. The protein transfer was run in 1x Towbin buffer at 100 V for 40 min. To verify that the protein transfer was successful, the negatively charged stain Ponceau S was used as a control. In order to prevent antibodies from binding to the membrane nonspecifically, the membrane was blocked with 5 % dry milk in 1x PBS-

Tween (PBS-T) at room temperature for 1 h. Subsequently, the membrane was incubated with primary ab in 1 % dry milk at 4 °C with gentle shaking overnight. The following day, the membrane was washed 3x 10 min with 1x PBS-T on a shaker in order to minimize background and remove unbound antibodies. Afterwards, the membrane was incubated with secondary ab in 3 % dry milk for 1 h at room temperature with gentle shaking. After incubation the secondary ab was removed, and the membrane was washed 3x10 min with 1x PBS-T. An additional washing step with 1x PBS for 5 min was added before treating the membrane with SuperSignal® West Pico Chemiluminescent substrate (Thermo Scientific) for 2 min. The secondary ab is coupled to the enzyme horseradish peroxidase (HRP) that catalyzes the chemical reaction resulting in a chemiluminescent product that emits light (Figure 4.3). HRP allows visualization of the protein of interest by catalyzing the oxidation of HRP substrates by hydrogen peroxide. Detection and imaging of the membrane was done using the ChemiDoc XRS+ connected with ImageLab™ v.3.0 (Bio-Rad). The protein abundance at each time point was normalized to the quantification of the corresponding α-V5 band at time point 0 h as well as α-β-tubulin bands as a loading control.

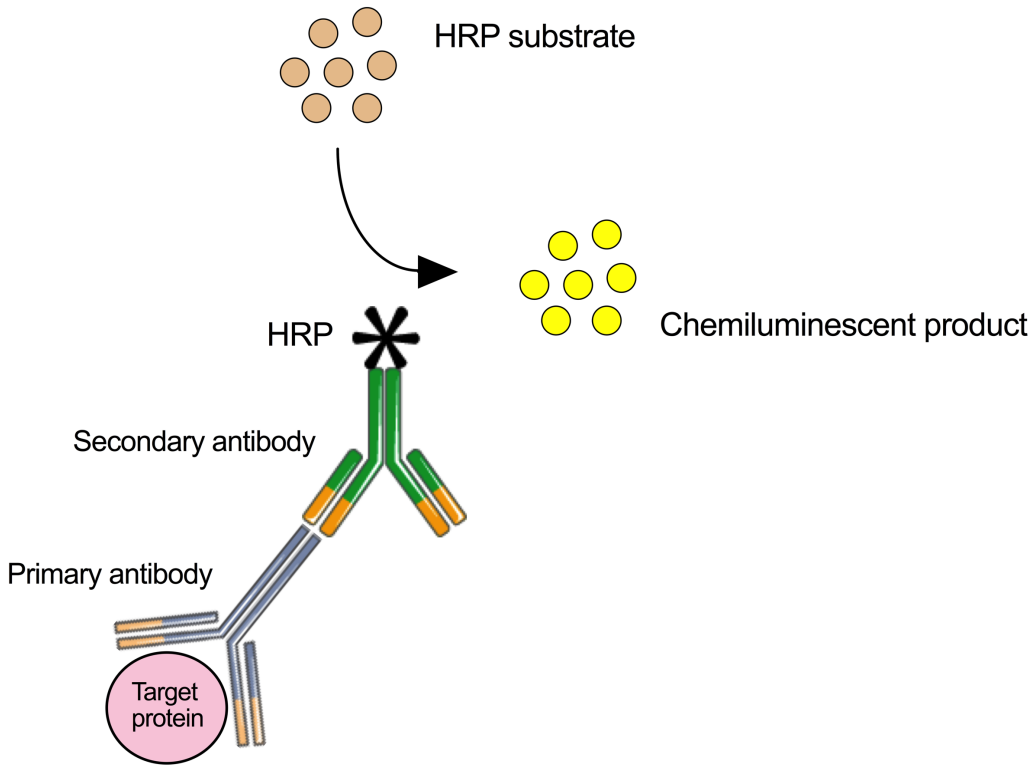

Figure 4.3: Detection of protein using luminescence. A specific primary ab (grey) is bound to the target protein (pink) by recognizing a specific binding site (epitope). This primary ab is recognized and bound by a specific secondary ab coupled with horseradish peroxidase (HRP). HRP (black) catalyzes the oxidation of HRP substrate (beige) resulting in a chemiluminescent product (yellow) that emits light, which empower an indirect visualization of the target protein. Figure was created using antibody illustrations from Servier Medical Art (http://smart.servier.com/).

4.4 Complex formation and intrinsic catalytic activity of NatA

In order to investigate how *NAA10* mutations affect the NatA complex formation and the intrinsic catalytic activity, immunoprecipitation and Carbon-14 (^{14}C) Ac-CoA-based *in vitro* Nt-acetylation assay were used. Complex formation was investigated by subjecting HeLa cells transfected with NAA10 WT-V5 and NAA10-V5 variants to immunoprecipitation (IP). V5-antibody and magnetic beads were used for pulling down of complexes and analyzing the immunoprecipitated by Western blot. The intrinsic catalytic activity was investigated by using the immunoprecipitated NatA complexes in a [^{14}C]-Ac-CoA-based *in vitro* Nt-acetylation assay.

4.4.1 Immunoprecipitation (IP)

Immunoprecipitation is a method used to separate a protein of interest from a cell lysate by the use of specific antibodies. An antibody specific for the protein of interest is incubated with cell lysate to enable the antibody to bind to the protein. The antibody-antigen immune complex is then pulled down from the solution using agarose or magnetic beads (Figure 4.4). High yields can be obtained with small volumes using both agarose and magnetic beads due to the high binding capacity of the beads. In order to not lose the protein of interest when using agarose beads, it is important to not disturb the bead-pellet during washing. Agarose beads are isolated by use of centrifugation, while magnetic beads can be isolated using a magnet. The protein of interest can be separated by SDS-PAGE and analyzed by Western blot. In these experiments, Dynabeads Protein G was used to isolate the protein of interest. Dynabeads Protein G are magnetic beads with recombinant Protein G (~17 kDa) covalently coupled to the surface. The recombinant immunoglobulin G (IgG) binding Protein G is lacking the albumin binding region, but the Fc binding domain is still present. Protein G binds antibodies preferentially through the heavy chain on the Fc region but can also in some cases bind within the Fab region (Graille et al., 2000, Idusogie et al., 2000).

For each IP experiment, five 10 cm dishes with 2 x 10^6 HeLa cells each were transfected with 4 µg NAA10 WT-V5, five dishes with 6 µg of each NAA10-V5 variant and five dishes with 6 µg LacZ-V5. 2 µg empty V5-vector was co-transfected with NAA10 WT-V5 to ensure equal conditions for the cells. The transfection was performed as described in section 4.4.2, except for 18 µl XtremeGene 9 transfection reagent and Opti-MEM to a final volume of 500 µl per cell dish. Cells were harvested on ice 48 h post-transfection using 2 ml 1x PBS. Cells were centrifuged at 1000 x *g* for 5 min at 4 °C. The resulting pellet was lysed inn 200 µl IPH buffer per cell dish and incubated on a rotating wheel at 4 °C for 15 min. The lysed cells were then centrifuged at 17000 x *g* for 5 min at 4 °C to remove cell debris. The supernatant was transferred to a new tube and the centrifugation step was repeated. The supernatant was transferred to new tubes and 30 µl cell lysate was saved for Western blot analysis.

To allow formation of immune complexes, the cell lysates were incubated with 4 µg α-V5 on a rotating wheel at 4 °C for 2-3 h. Afterwards, 40 µl washed Protein G magnetic beads

(Thermo Scientific) were added to each IP sample and incubated at 4 °C on a rotating wheel overnight in order to retrieve the immune complexes. The following day, each IP sample was placed on a magnetic holder and 30 µl of the supernatant were saved for Western blot analysis. The IP beads were washed three times with IPH wash buffer and twice with 1x acetylation buffer before being resuspended in 90 µl 2x acetylation buffer. After resuspension, the IP beads were ready to be used in ^{14}C acetylation assays.

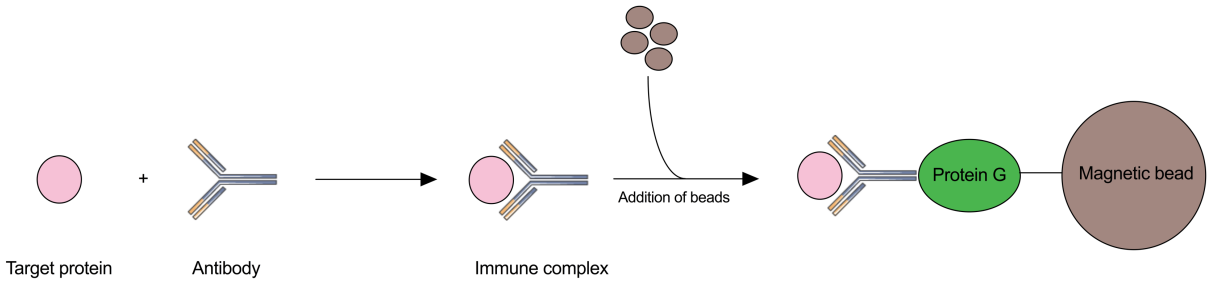

Figure 4.4: Schematic illustration of immunoprecipitation of protein using magnetic beads. A specific antibody (blue) recognizes and binds to the protein of interest (red) in the cell lysate in order to form an immune complex. These immune complexes are pulled out and retrieved by adding Protein G magnetic beads to the cell lysate. Protein G binds specifically to the heavy chain off the antibody on the immune complex, allowing untargeted proteins to be washed off when isolating the beads with a high-power magnet. Figure was created using antibody illustrations from Servier Medical Art (http://smart.servier.com/).

4.4.2 [^{14}C]-Ac-CoA-based *in vitro* Nt-acetylation assay

For the purpose of investigating the intrinsic catalytic activity of NatA for the immunoprecipitated complexes, a [^{14}C]-Ac-CoA-based acetylation assay was performed. This assay is a highly sensitive method that enables detection of very weak acetylation events (Drazic and Arnesen, 2017). The method enables radioactive detection of peptides that are Nt-acetylated by the NatA complexes using ^{14}C-Ac-CoA as a donor. The Nt-acetylation reaction is quenched by placing the magnetic beads (contains the enzyme) on a magnet followed by removing the radioactive supernatant (contains the substrate). The radioactive supernatant is transferred onto P81phosphocellulose filter squares that bind only positively charged oligopeptides, while unincorporated ^{14}C-Ac-CoA will be washed off. To determine the radioactivity of the incorporated ^{14}C-Ac, filter squares are soaked in scintillation fluid and measured by a liquid scintillation analyzer. A scintillation analyzer detects light that are indirectly generated by ionizing radiation caused by interactions between a scintillator material and emitted β-particles. The measured Nt-acetylation activity is adjusted to the amount of NatA complexes used, which is determined by quantification of the IP samples by Western blot analysis (Drazic and Arnesen, 2017).

To examine and compare the catalytic activity between NAA10 WT and NAA10 variants, ^{14}C- acetylation assays were performed. Three replicates of 200 mM specific peptide (SESS or EEEI), 50 µM isotope-labeled [^{14}C]-Ac-CoA, 10 µl IP beads and dH_2O were mixed to a final volume of 25 µl. In addition, two negative replicates without peptides were prepared.

Samples were incubated at 37 °C with shaking at 1400 rpm in a thermo shaker for 30 min. After the enzyme reaction the magnetic beads were isolated by placing the samples on a magnetic holder and 23 µl of supernatant were transferred onto P81 phosphocellulose filter squares. The filter squares were washed 3x 5 min with 10 mM HEPES buffer (pH 7.4) followed by air drying. In order to measure the incorporated [^{14}C]-Ac signal, filter squares were placed in individual tubes filled with 5 ml scintillation fluid and the signal was measured by a scintillation analyzer (Tri-Carb 3100TR Liquid Scintillation Analyzer, Packard BioScience company). The IP samples were analyzed by Western blot (as described in section 4.4.5). The measured activity was normalized by quantification of the corresponding α-V5 bands and α-NAA15 bands.

5 Results

5.1 Genetic variants and clinical data

The objective of this thesis was to functionally characterize two different NAA10 variants, NAA10 L11R and NAA10 H16P, suspected to cause pathological effects in the respective patients harboring these. The patient harboring the NAA10 L11R mutation is a girl three years of age with a mild pyramidal syndrome, strabismus, feeding difficulties, no speech development and has a large arachnoid cyst between the cerebellum and the occipital area. The patient harboring the NAA10 H16P variant is a ten-year-old girl with severe syndromic intellectual disability, disturbed behavior and severe delayed motor and language development. Her blood leukocyte X-inactivation pattern was skewed (95/5) toward the maternally inherited X-chromosome. Figure 5.1 illustrates the human *NAA10* gene with its eight exons (E1-E8) and the functional protein domains are indicated below the gene. The missense mutations identified to date (in grey boxes) and the two missense mutations L11R (orange) and H16P (green) investigated in this thesis are placed in the *NAA10* exon in which they occur. The L11R and H16P variants are in exon 2 in the NAA15 interaction domain that is a part of the conserved GNAT fold of NAA10.

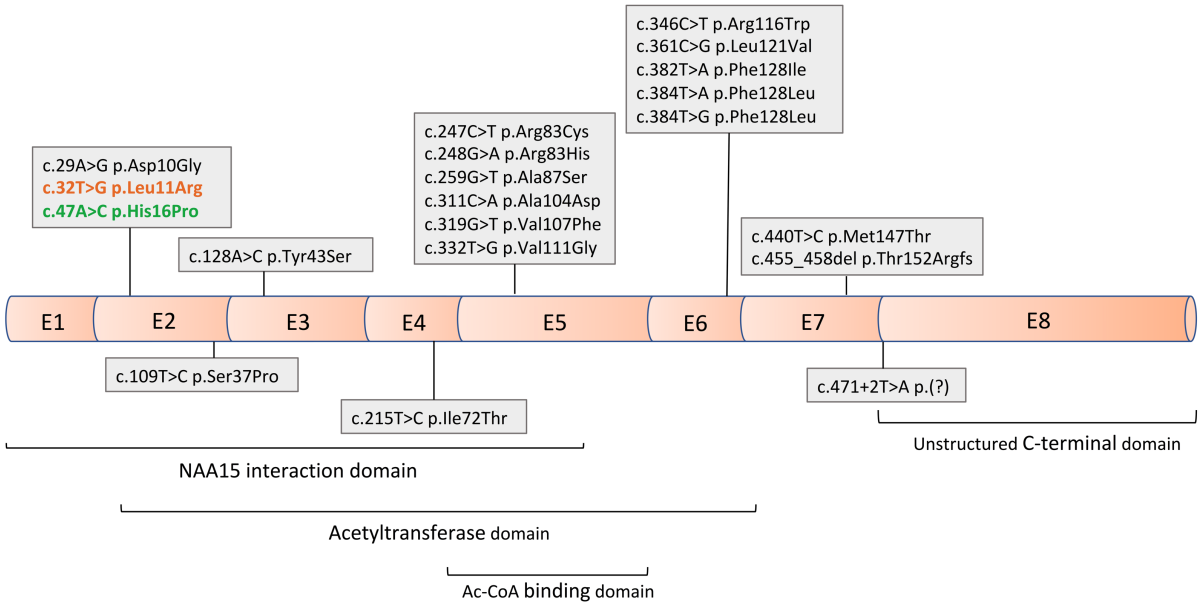

Figure 5.1: Schematic overview of the genomic structure of the human *NAA10* gene with *NAA10* missense mutations and functional domains. NAA10 variants identified to date are listed in grey boxes in the *NAA10* exon they occur. The *NAA10* variants L11R (orange) and H16P (green) are studied in this thesis. The functional domains of NAA10 are represented below the exon structure.

5.2 Bioinformatic analysis

5.2.1 Multiple sequence alignment (MSA)

In order to investigate the degree of amino acid conservation in orthologous NAA10 amino acid sequences, a MSA was assembled using ESPript v.3 (Xavier & Patrice, 2014) as described in methods section 4.1.1. As seen from the MSA (Figure 5.2A), the NAA10 N-terminus is highly conserved between the represented species, while the C-terminus varies greatly. The Ac-CoA binding motif is shown by the blue box and the amino acid residues L22, E24, Y26, R112 and Y138 that are important for substrate binding are indicated with black stars (*) (Liszczak et al., 2013). These residues are completely conserved among the represented species. Secondary structure information (displayed in black above the alignment) were obtained from the 3D structure of human NAA10 (Gottlieb & Marmorstein, 2018).

The amino acid L11 is completely conserved among all the represented species, suggesting that this amino acid is important for NAA10 function and/or stability. The amino acid H16 is less conserved through evolution. However, the particular amino acid for all species except for *A. thaliana* is a charged/polar amino acid, which is able to form hydrogen bonds. This suggests that the variation may be tolerated as long as the properties of the amino acid are conserved. Figure 5.2B shows the structure of hNAA10 with the zones of conservation in red. Both amino acid residues are found on the α1-helix where L11 (orange) face inside and H16 (green) face outside the protein.

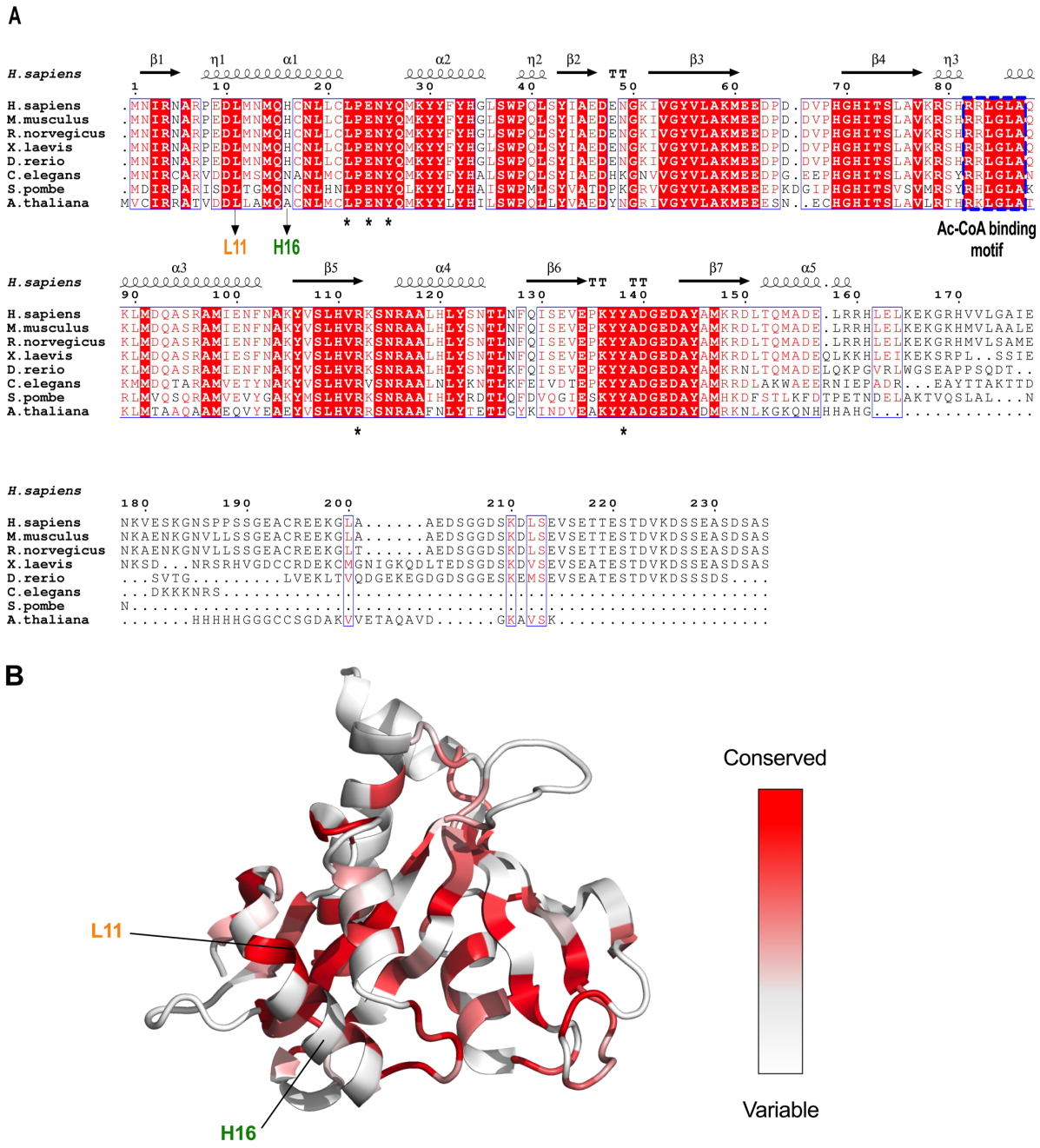

Figure 5.2: NAA10 multiple sequence alignment and structure with zones of conservation. A) An MSA containing NAA10 sequences from different species was generated using Clustal Omega and visualized using ESPript 3.0. Completely conserved residues are highlighted with white letters in red boxes. Highly conserved residues are labeled with red letters on white background (Xavier & Patrice, 2014). **B)** A structure of hNAA10 (PDB ID: 6C9M) with zones of conservation shown in red was generated using ENDscript 2.0 (Xavier & Patrice, 2014) and visualized using PyMOL. *, residues important for substrate binding. η (eta) refers to a 3_{10}-helix.

5.2.2 *In silico* missense mutation predictions

In silico missense mutation prediction was performed for each NAA10 variant in order to get a preliminary assessment of their impact on protein function and stability. Several *in silico* tools were employed as described in methods section 4.1.2. Three selected prediction tools are presented here, while the rest of the results from all the prediction tools applied are listed in Table 8 in the supplementary data.

The PMut predictor (Lòpez-Ferrando et al., 2017) uses sequence conservation and predicted physicochemical properties to predict the pathological nature of a given single amino acid variant. Table 5.1 presents the predicted pathological effect of each missense mutation based on the NAA10 protein sequence. PMut predicted both mutations to be disease causing with a probability score of 0.85 for L11R and a score of 0.70 for H16P. The percentages represent the confidence degree of the predicted probability scores.

Table 5.1: Pathological effects of missense mutations by PMut.

NAA10 mutation	Prediction	Probability*
L11R	Disease	0.85 (91%)
H16P	Disease	0.70 (86%)

* Score ranges from 0.0 (neutral) to 1.0 (pathological). Mutation are classified as pathological if probability > 0.5. Percentage represents the confidence degree of the predicted scores.

SIFT (Sim et al., 2013) was used to investigate whether the missense mutations would affect protein function. SIFT considers the position, amino acid properties and sequence homology to predict whether the type of amino acid change affects protein function. The results of the predictions are listed in Table 5.2. Substitutions with a probability score less than 0.05 is considered deleterious. The NAA10 L11R variant was predicted to affect protein function with a score of 0.01. The NAA10 H16P variant was considered tolerated with a score of 0.06 just above the threshold.

Table 5.2: Missense mutation`s impact on protein function by SIFT.

NAA10 mutation	Prediction	Probability score*
L11R	Affect protein function	0.01
H16P	Tolerated	0.06

* Scores ranges from 0.0 (deleterious) – 1.0 (tolerated). Substitutions with scores less than 0.05 is predicted as deleterious.

CUPSAT (Parthiban et al., 2006) was used to predict changes in protein stability upon missense mutations. CUPSAT predicts ΔΔG, the difference in free energy of unfolding between wild-type and mutant proteins by using structural environment specific atom potentials and torsion angle potentials. CUPSAT distinguishes the amino acid environment of the native and mutated residue using its solvent accessibility and secondary structure specificity. Table 5.3 show the predicted overall stability and the ability to adapt the torsion

angle of the two missense mutations L11R and H16P. The mutation of leucine to arginine was predicted to be favorable overall with a ΔΔG value of 1.25 kcal/mol. As for the H16P variant, the exchange of histidine to a proline was predicted to be an overall destabilizing substitution with a ΔΔG value of -0.42 kcal/mol.

Table 5.3: Missense mutation`s impact on protein stability by CUPSAT.

NAA10 mutation	Overall stability	Torsion*	Predicted ΔΔG (kcal/mol)
L11R	Stabilizing	Favorable	1.25
H16P	Destabilizing	Unfavorable	-0.42

* Overall stability is calculated from atom potential and torsion angle potentials. ΔΔG, the difference in free energy of unfolding between wild-type and mutant proteins is calculated by using structural environment specific atom potentials and torsion angle potentials.

5.2.3 *In silico* structural analysis

Single nucleotide base substitution in coding regions can result in an amino acid change and subsequent altered protein structure or function. Native protein structure is essential for biological function and structural alterations can result in loss of function. *In silico* structural analyses (methods section 4.1.3) were performed in order to evaluate how changing one single amino acid residue in NAA10 might impact its structure, function and interactions.

Figure 5.3 presents an overview of the human NatA structure (PDB: 6C9M, Gottlieb et al., 2018) with Ac-CoA and the peptide SASE implemented by superimposition with *S. pombe* NAA10 (PDB: 4KVM, Liszczak et al., 2013) using PyMOL. The structure of Ac-CoA is shown in blue, along with the peptide SASE shown in yellow. Leucine (L11) is a nonpolar amino acid with a hydrophobic side chain located at the α1-helix where it faces inside the protein toward the α2-helix of NAA10. Histidine (H16) is a positively charged polar amino acid located at the α1-helix where it faces outside the protein toward the ligand IP_6 located in the groove between NAA10 and NAA15. The phosphate groups of IP_6 have one or two negative charges at physiological pH, making it an anion. The IP_6 molecule is suggested to stabilize and strengthen the interaction between NAA10 and NAA15 and therefore being of biological relevance for the NatA complex (Gottlieb & Marmorstein, 2018; Deng et al, 2019).

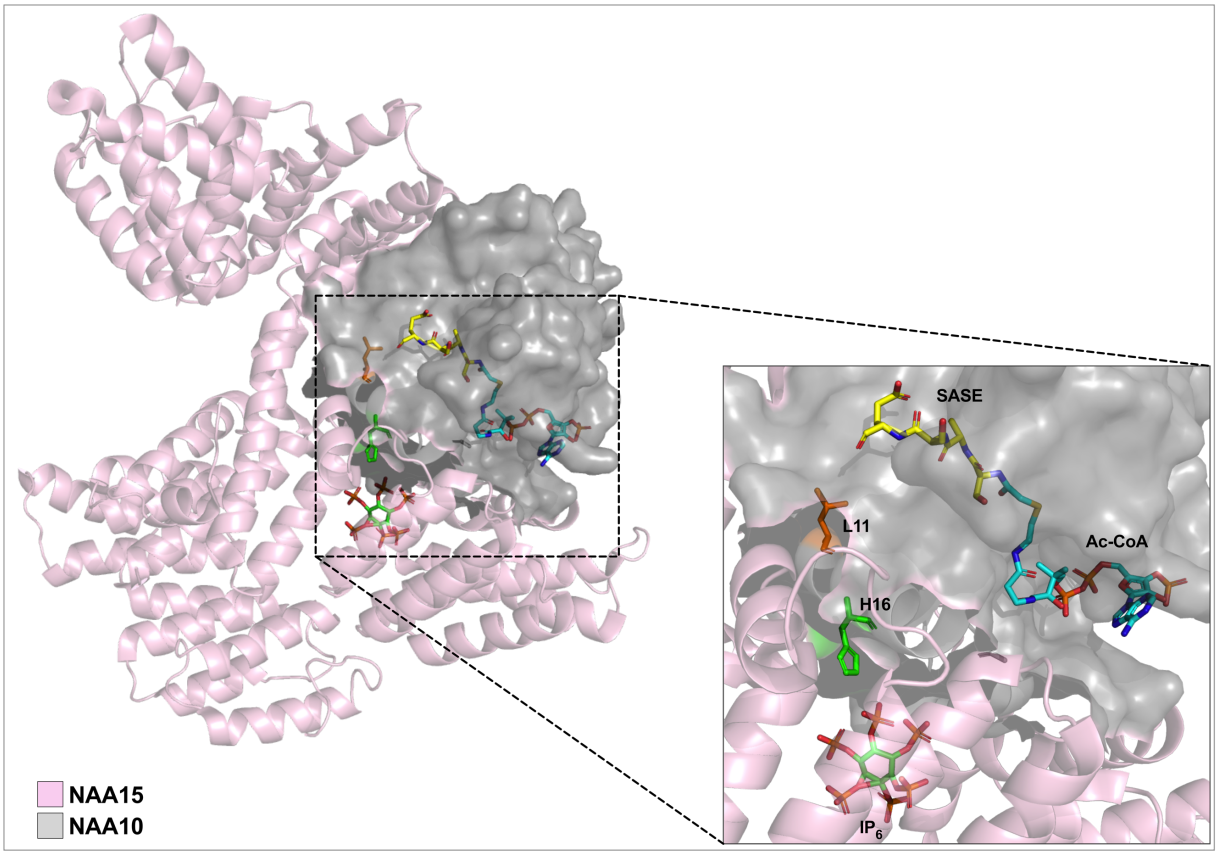

Figure 5.3: Structural overview of the human NatA complex. The human NatA structure (PDB: 6C9M) with Ac-CoA and the peptide SASE from *S. pombe* NAA10 (PDB: 4KVM) visualized in PyMOL. Amino acids of interest are colored orange and green. Ac-CoA is shown in blue, along with the peptide SASE shown in yellow.

Figure 5.4A presents a closeup of the location of L11 and its surroundings. Leucine (L11) is surrounded by polar uncharged (Q15 and N13) and hydrophobic residues (P8, M14, M28, Y31, F32). The side chain of L11 interacts with M14, M28 and F32, making up a hydrophobic pocket by hydrophobic interactions. Figure 5B illustrates the outcome of the substitution of the buried, neutral residue leucine, to the bigger, positively charged arginine in the hydrophobic pocket. The L11R mutation may cause loss of hydrophobic interactions in the core of the protein. As seen in Figure 5.4B, the hydrophobic interactions between L11 to the s-methyl thioester side chain of M14 and the polar interaction between carboxyl group of Y31 is not predicted when leucine is exchanged with arginine. Arginine is predicted by DynaMut (visualized with PyMOL) to make a new weak polar interaction with the carbonyl group of alanine (A6) located on β1 and a weak hydrogen to the hydroxyl group of serine (S42) located on β2.

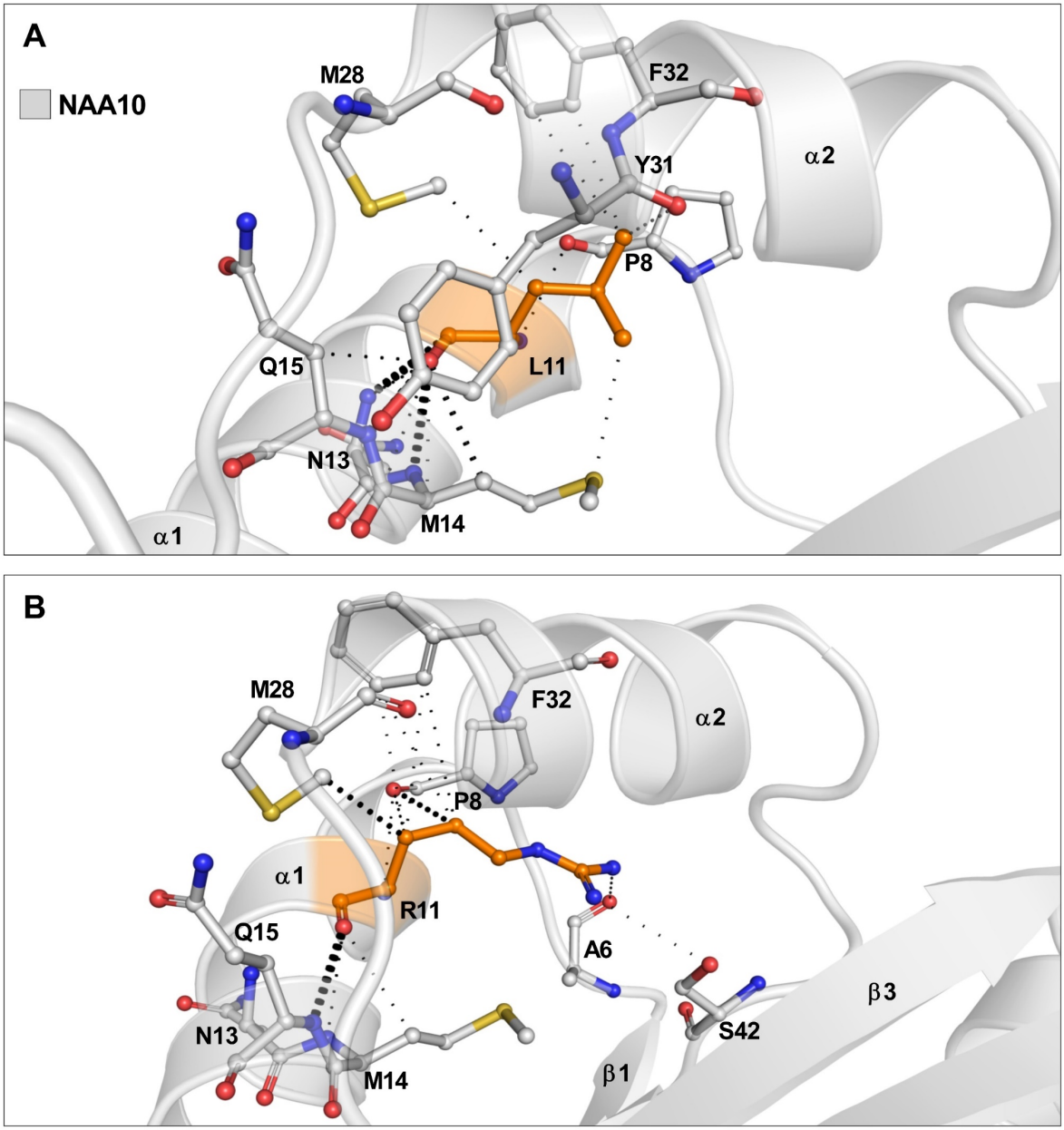

Figure 5.4: Structural visualization of L11R mutation site and surrounding residues. **A)** Overview of NAA10 (grey) with the amino acid L11 subjected to mutation (orange). Leucine is surrounded by the polar uncharged residues (Q15 and N13) and the hydrophobic residues (P8, M14, M28, Y31, F32) making up a hydrophobic pocket by hydrophobic interactions (green dotted line). **B)** Overview of the amphipathic amino acid, arginine (R11) with the predicted hydrogen bonding and hydrophobic interactions (red dotted lines and green dotted lines). Arginine is predicted by DynaMut (visualized in PyMOL) to make a new weak interaction with the backbone carbonyl group of alanine (A6) located on β1 and a weak hydrogen to the hydroxyl group of serine (S42) located on β2.

Figure 5.5 shows a closeup of the H16P variant with its surrounding residues and the ligand IP_6. The side chain of H16 points toward IP_6 in the groove between NAA10 and NAA15. The amino acids surrounding H16 are M12, N13, M14, N18, L19 and L20, all found in the α1-helix. In addition, two amino acids (W486 and F551) from NAA15 point toward the side chain of H16. The differences in properties between histidine and proline can easily cause

loss of internal and external interactions. As seen in Figure 5.5A, histidine is predicted by PyMOL to have polar interactions with the ligand IP_6 and interactions the two NAA15 residues W486 and F551. These interactions can be lost when proline is introduced. The P16 points toward IP_6 in the groove between NAA10 and NAA15 (Figure 5.5B). The amino acids surrounding P16 is the same as H16. P16 interacts with F551, however the potential polar bond with IP_6 and interactions to W486 of NAA15 are lost.

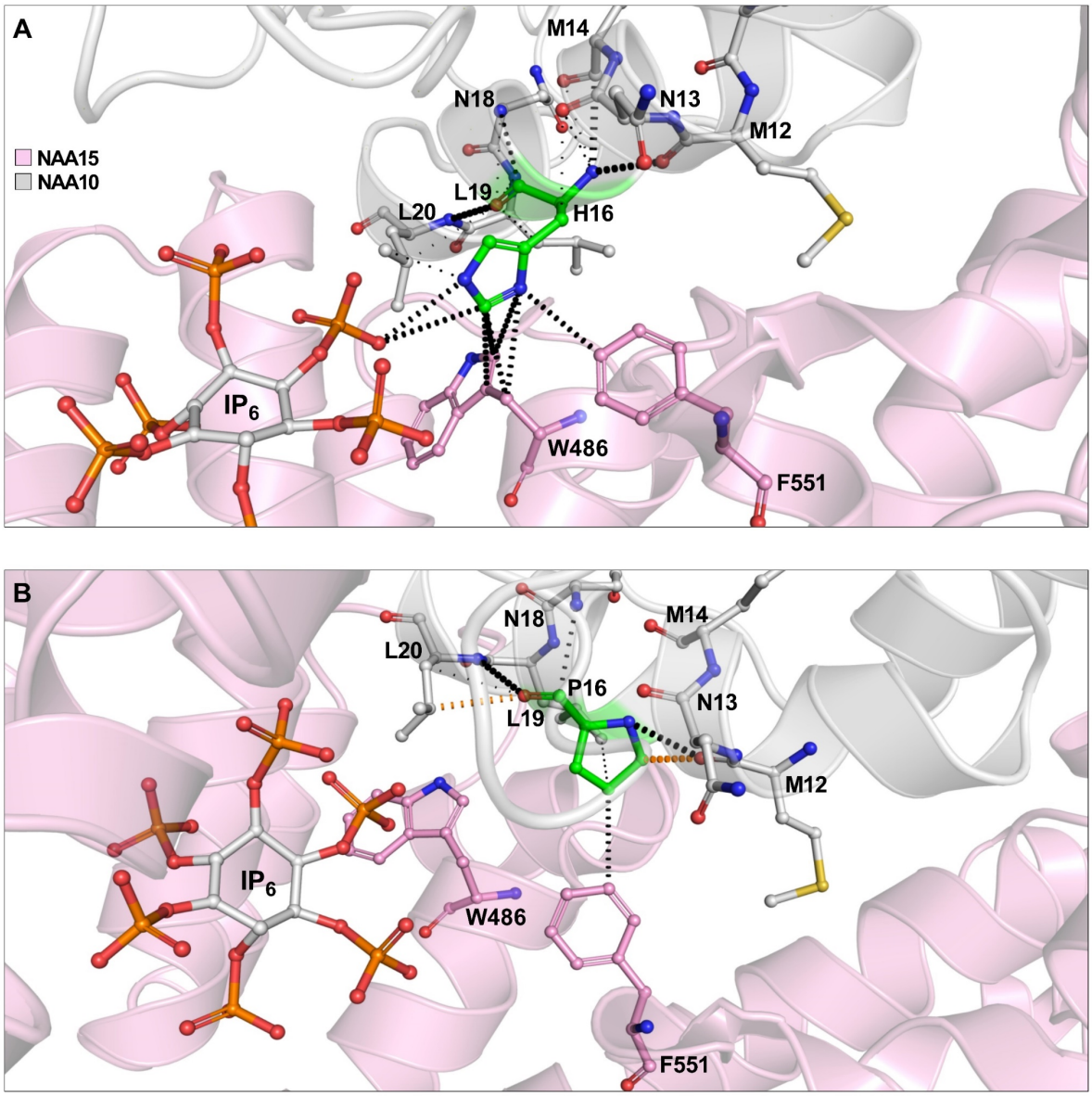

Figure 5.5: Structural visualization of H16P mutation site and surrounding residues. **A)** Overview of NAA10 (grey) in complex with NAA15 (pink) with the amino acid H16 subjected to mutation (green). The side chain of H16 points toward IP_6 in the groove between NAA10 and NAA15. The amino acids surrounding H16 are M12, N13, M14, N18, L19 and L20 all found in the α1-helix. H16 shows hydrogen bonding and/or ionic interactions with two amino acids from NAA15 (W486 and F551) (black dotted lines) **B)** Overview of NAA10 (grey) in complex with NAA15 (pink) with the mutated amino acid P16 (green).

Another interesting observation made as histidine is substituted with proline, is the loss of interactions inside the α1-helix. The hydrogen bonds made by histidine to other residues (Figure 5.6A) could potentially be lost and affected when proline is introduced. As seen from Figure 5.6B, the backbone bonds (NH-CO) between H16 and M14 is lost when proline is introduced. The weak hydrogen bonds between histidine and the two asparagine residues (N13 and N18) and leucine (L19) are also lost when proline is introduced.

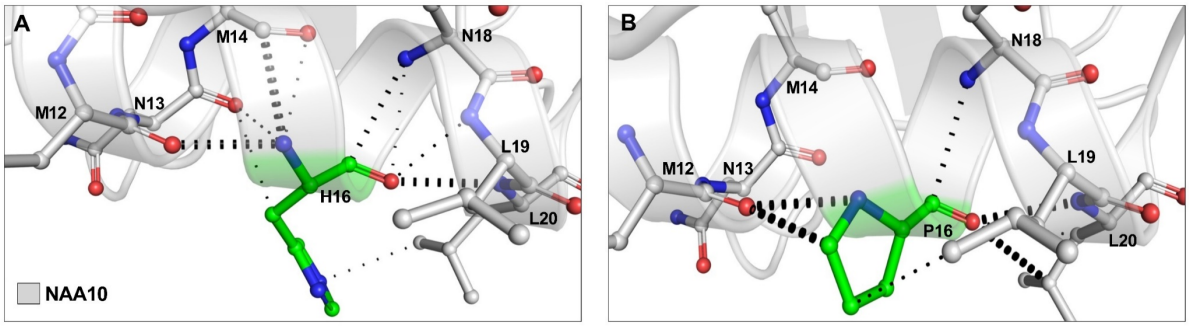

Figure 5.6: H16P mutation potensially causes loss of stabilizing hydrogen bonds in the α1-helix of NAA10. **A)** The predicted hydrogen bonding and within the α-helix with H16 subjected to mutation is indicated by black dots. **B)** The predicted hydrogen binding when proline is introduced in position 16. The interactions between H16 and asparagine (N13 and N18) and leucine (L19) are also lost when proline is introduced. The loss of stabilizing hydrogen bonds found in the α-helix may destabilize or disrupt the helix by potentially introduce a bend, which can have severe effects in the structure of the protein.

5.3 *In vitro* stability analysis of NAA10 variants by cycloheximide chase assay

The structure of a protein can be altered by a single missense mutation, resulting in a destabilized abnormal protein. In order to investigate how the NAA10 mutations affected NAA10 stability in human cells, cycloheximide chase assays were performed as described in methods section 4.3.3. Cycloheximide inhibits protein biosynthesis by preventing translational elongation, giving an indirect measure of the protein stability in cells, since abnormal proteins often are degraded more rapidly as compared to native proteins.

HeLa cells were transfected with NAA10 WT-V5, NAA10 L11R-V5 or NAA10 H16P-V5. Two days after transfection, cells were treated with 50 µg/µl cycloheximide for up to six hours. Cells were harvested at specific timepoints (0, 2, 4, and 6h) and the cell lysates were analyzed by Western blot. The amount of NAA10-V5 present at timepoint 2, 4 and 6h post treatment was normalized to the quantification of the corresponding α-V5 band at time point 0 h as well as α-β-tubulin bands as a loading control.

The results of three independent cycloheximide chase experiments performed for NAA10 L11R and NAA10 WT are shown in Figure 5.7. The L11R variant showed a reduced cellular stability compared to NAA10 WT. After two hours, the amount of NAA10 L11R molecules present in cells was 27.3 ± 4.6 % that relates to an approximately 70% reduction relative to the amount of NAA10 before treatment. In comparison, the amount of NAA10 WT molecules

was 49.7 ± 3.1 %, equal a 50 % reduction relative to the amount of NAA10 before treatment. After four hours, the amount of NAA10 L11R was 15.9 ± 3.1 %, while the amount of NAA10 WT molecules was 29.7 % ± 6.5 %. This corresponds to an approximately 84 % reduction for NAA10 L11R and a 70 % reduction for NAA10 WT relative to the amount of NAA10 before cycloheximide treatment. After six hours, the amount of NAA10 L11R present in cells was only 7.5 ± 3.1 %, whereas the amount of NAA10 WT was 22.4 ± 5.3 %. The amount of molecules corresponds to an approximately 90 % reduction for NAA10 L11R and a 80 % reduction for NAA10 WT six hours after treatment.

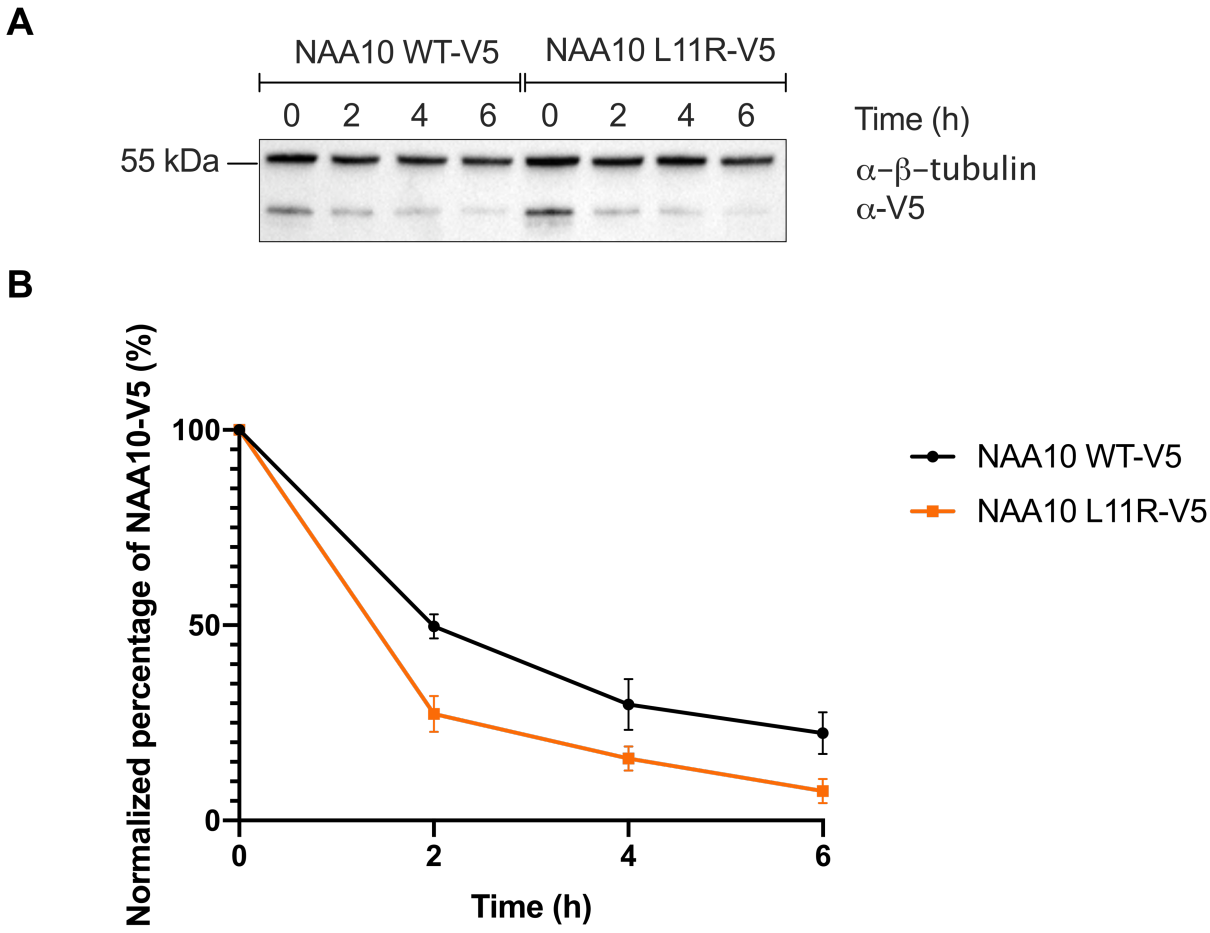

Figure 5.7: Stability analysis of the NAA10 L11R variant by cycloheximide chase assay. **(A)** HeLa cells were transfected with NAA10 WT-V5 or NAA10 L11R-V5 in order to investigate and compare their cellular stability. 48 h post transfection, the cells were treated with 50 µg/µl cycloheximide for up to 6 h. Cell lysates were analyzed by Western blot with α-V5 and α-β-tubulin ab **(B)** The α-V5 ab signal was normalized to the α-β-tubulin signal at timepoint 2, 4 and 6 h post treatment and expressed as a percentage of the amount of NAA10-V5 present at timepoint 0 (before treatment). The graph includes results from three independent experiments.

The average stability curve of three independent cycloheximide chase experiments performed for NAA10 H16P and WT is presented in Figure 5.8. There was no significant difference observed between the NAA10 H16P and WT turnover rate. After two hours, the amount of NAA10 H16P molecules was 61.4 ± 18.3 %, which corresponds to a 35 % reduction relative

to the amount of NAA10 before treatment. The amount of NAA10 WT molecules was 65.4 ± 19.2 %, which relates to a 35 % reduction as well. Four hours after treatment, the amount of NAA10 H16P was 35.9 ± 5.1 % and the amount of NAA10 WT was 39.4 ± 1.5 %, resulting in a 60 % reduction of molecules relative to the amount of NAA10 before treatment. At last, after six hours the amount of NAA10 H16P molecules was 24 ± 3.2 % and the amount of NAA10 WT was 27.3 ± 6.6 %, equivalent to an approximately 70 % reduction in molecules relative to the amount of NAA10 before treatment. There is no significant difference between the mutant and WT, indicating that NAA10 H16P most likely does not affect neither monomeric NAA10 nor NatA stability.

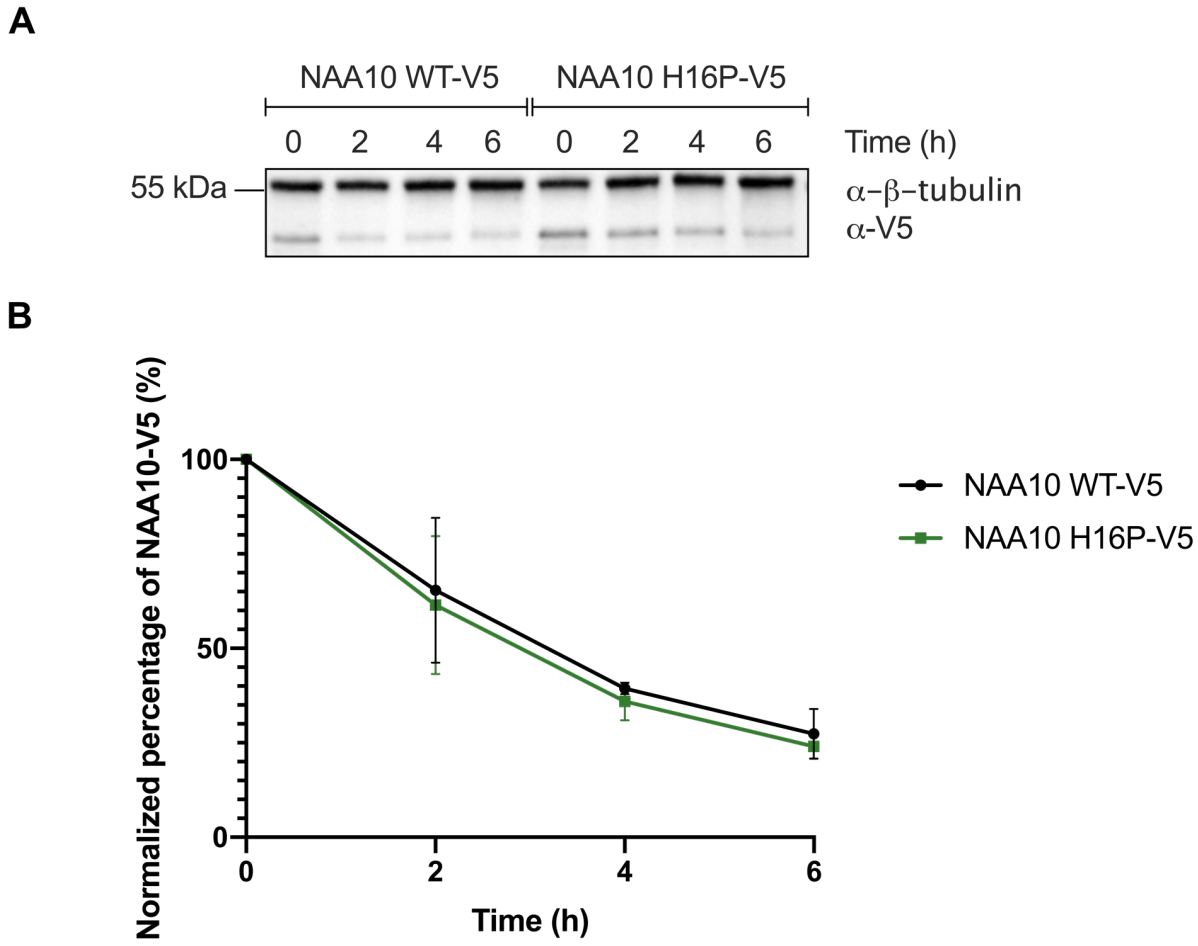

Figure 5.8: Stability analysis of the NAA10 H16P variant by cycloheximide chase assay. (A) HeLa cells were transfected with NAA10 WT-V5 and NAA10 L11R-V5 in order to investigate and compare their cellular stability. 48 h post-transfection, the cells were treated with 50 µg/µl cycloheximide for up to 6 h. Cell lysates were analyzed by Western blot with α-V5 and α-β-tubulin ab **(B)** The α-V5 ab signal normalized relative to the α-β-tubulin signal at timepoint 2, 4 and 6h post treatment is expressed as a percentage of the amount of NAA10-V5 present at timepoint 0 (before treatment). The graph includes results from three independent experiments.

5.4 NatA complex formation and *in vitro* activity of NatA

The catalytic subunit NAA10 exerts its NatA activity in complex with auxiliary subunit NAA15 with a substrate preference of N-terminal alanine, serine, threonine, glycine, valine and cysteine (Liszczak et al., 2013), while monomeric NAA10 acetylates acidic N-termini *in*

vitro (Van Damme et. al., 2011). It was of interest to investigate how NAA10 L11R and NAA10 H16P affected the *in vitro* complex formation and the intrinsic activity of NatA.

HeLa cells were transfected with V5-tagged NAA10 WT, NAA10 L11R or NAA10 H16P. LacZ-V5 was used as a control plasmid. After two days of growth, cells were harvested and lysed, and NAA10-V5 was subsequently pulled down by IP with α-V5 ab and Protein G magnetic beads as described in methods section 4.4.1. Thereafter, NAT activity in the NAA10-V5 immunoprecipitate was measured by ^{14}C-acetylation assay as described in methods section 4.4.2. The amount of NAA10 (which will be a mixture of monomeric NAA10 and NAA10 in complex with NAA15) and NAA15 (a measure of NatA complex only) present in each sample was determined by SDS-PAGE and Western blot with α-V5 and α-NAA15 ab. The measured catalytic activity was normalized to both the amount of NAA10-V5 and NAA15 present in the sample.

The Western blot analysis of NAA10-V5 precipitate of NAA10 L11R and NAA10 WT is shown in Figure 5.9A. As seen from the cell lysate, the expression of NAA10 WT and NAA10 L11R were almost equal, whereas the amount of immunoprecipitated NAA10 L11R was higher than NAA10 WT (44 % more). In contrast, approximately 56 % less NAA15 was co-immunoprecipitated with NAA10 L11R relative to NAA10 WT.

The catalytic activity of NAA10 L11R toward the NatA substrate SESS is normalized to α-NAA15 bands (NAT activity per NatA complex). As seen in Figure 5.9B, the catalytic activity was reduced almost 1.8-fold relative to WT. The activity toward the monomeric NAA10 substrate EEEI is normalized to α-V5 bands (NAT activity per NAA10-V5 molecule). The monomeric catalytic activity of NAA10 L11R was equal relative to WT. The amount of NAA15 can serve as a measurement of the amount of NatA complex present in each sample since the contribution of monomeric NAA10 on acetylation of SESS is minimal (Liszczak et al., 2013). The data indicates that NAA10 L11R impairs NatA activity, whereas the monomeric activity of NAA10 L11R toward EEEI is unaltered.

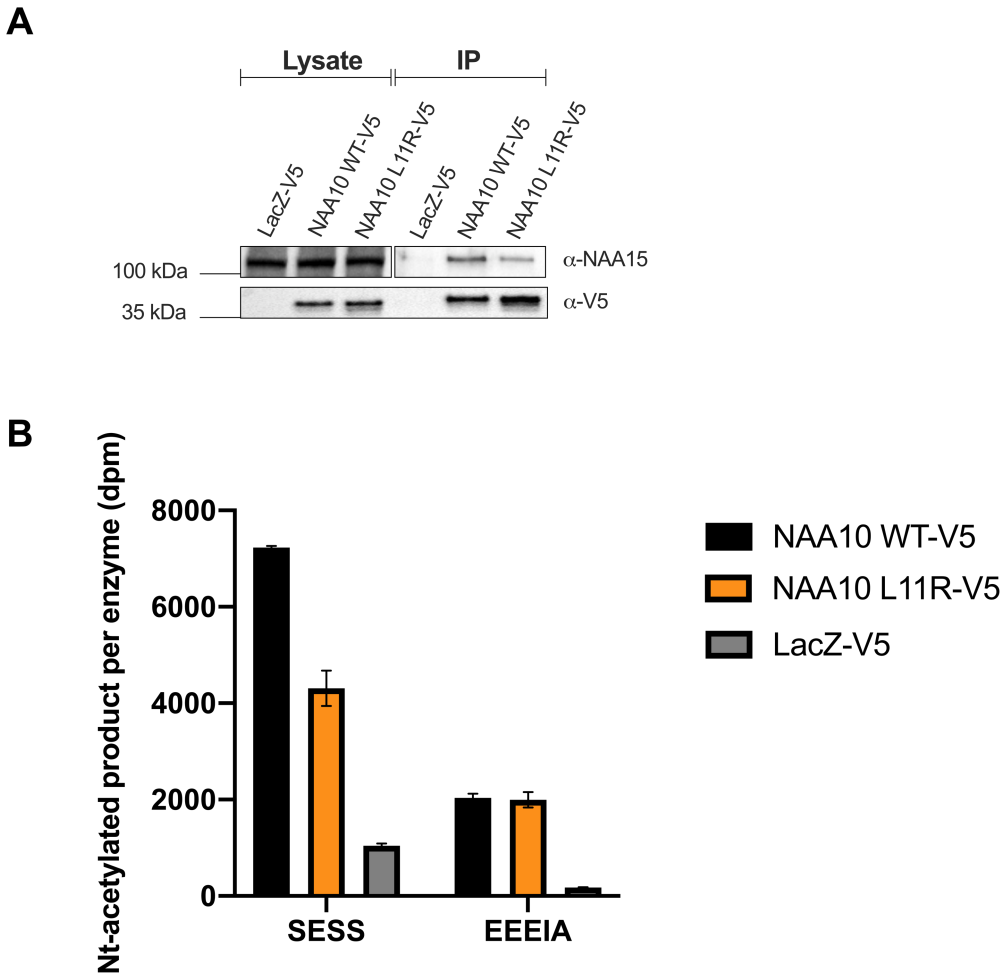

Figure 5.9: NatA complex formation and intrinsic activity of NAA10 L11R and NAA10 WT. A) Western blot analysis of V5-tagged NAA10 WT, NAA10 L11R and LacZ immunoprecipitated from transfected HeLa cells using α-V5. **B)** NAT activity of NAA10 WT-V5, NAA10 L11R-V5 and LacZ-V5 immunoprecipitate toward NatA substrate SESS and monomeric NAA10 substrate EEEI measured by ^{14}C-acetylation assay. The measured catalytic activity toward SESS was normalized to α-NAA15 bands to show relative product formation per NatA complex. The measured catalytic activity toward EEEI was normalized to α-V5 bands to show relative product formation relative to the NAA10 signal.

Figure 5.10 shows the Western blot analysis of the IP-activity experiment of NAA10 H16P and NAA10 WT. As seen from the cell lysate, the expression of WT and the H16P variant was approximately equal, while the IP of the H16P variant gave a 37 % higher yield compared to WT. In contrast, approximately 47 % less NAA15 was co-immunoprecipitated with NAA10 H16P relative to NAA10 WT.

The catalytic activity of NAA10 H16P toward the NatA substrate SESS is normalized to α-NAA15 bands (NAT activity per NatA complex). As seen in Figure 5.10B, the catalytic activity was reduced almost 4-fold relative to WT. The activity toward the monomeric NAA10 substrate EEEI is normalized to α-V5 bands (NAT activity per NAA10-V5 molecule). The monomeric catalytic activity of NAA10 H16P was increased relative to WT. Altogether, the data suggest that the H16P variant impairs NatA activity, while monomeric NAA10 activity is not impaired.

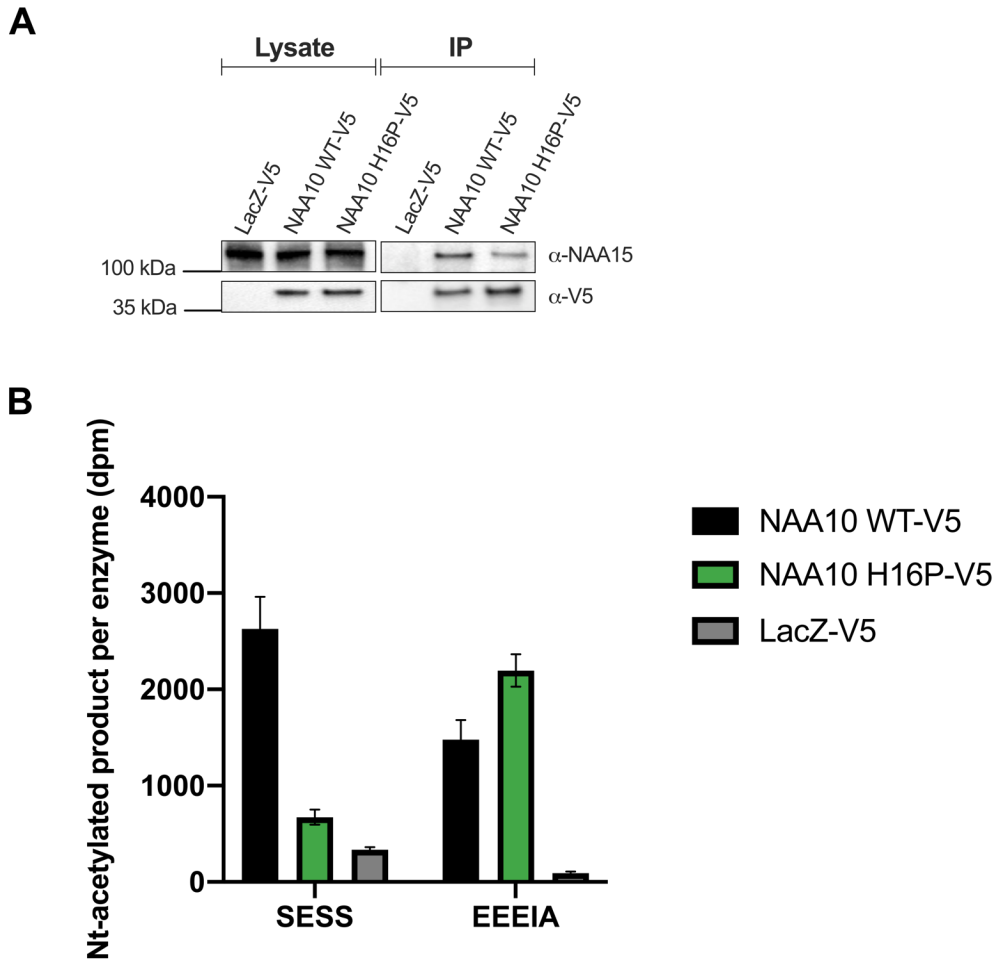

Figure 5.10: NatA complex formation and intrinsic activity of NAA10 H16P and NAA10 WT. A) Western blot analysis of V5-tagged NAA10 WT, NAA10 H16P and LacZ immunoprecipitated from transfected HeLa cells using α-V5. **B)** NAT activity of NAA10 WT-V5, NAA10 H16P-V5 and LacZ-V5 precipitate toward NatA substrate SESS and monomeric NAA10 substrate EEEI measured by ^{14}C-acetylation assay. The measured catalytic activity toward SESS was normalized to α-NAA15 bands to show relative product formation per NatA complex. The measured catalytic activity toward EEEI was normalized to α-V5 bands to show product formation relative to the NAA10 signal.

In order to confirm if the binding of NAA10 to NAA15 is altered by the mutations as suggested by the data above, a α-NAA15 immunoprecipitation was performed as well. Given the results from the α-V5 IP samples, it was expected that with an equal amount of immunoprecipitated NAA15 we would see less NAA10-V5 co-precipitated for both NAA10 variants relative to WT. The transfection of HeLa cells, harvesting and lysis of the cells were performed as in the α-V5 IP experiment, except this time NAA15 was pulled down by IP with α-NAA15 ab and Protein G magnetic beads as described in methods section 4.4.1.

The Western blot analysis of the α-NAA15 IP experiment of NAA10 L11R and NA10 WT is shown in Figure 5.11A. As seen from the cell lysate, the expression of NAA10 WT and NAA10 L11R were quite similar. The amount of NAA15 in the IP samples were equal as well. As expected, more NAA10-V5 was co-immunoprecipitated with NAA10 WT relative to NAA10 L11R. This experiment was only performed twice due to time limitations and showed

inconsistent results. Therefore, the experiment needs to be repeated several times in order to make any conclusions.

Figure 5.11B shows the Western blot analysis of the α-NAA15 IP experiment of NAA10 H16P and NAA10 WT. The expression seen in cell lysate and the amount of NAA15 in the IP samples for NAA10 WT and NAA10 H16P were relatively equal. Unexpectedly, more NAA10-V5 was co-immunoprecipitated with NAA10 H16P relative to WT. Due to time limitations, this experiment was only performed twice, with consistent results. However, the experiment needs to be repeated several times in order to make any interpretations and a certain conclusion.

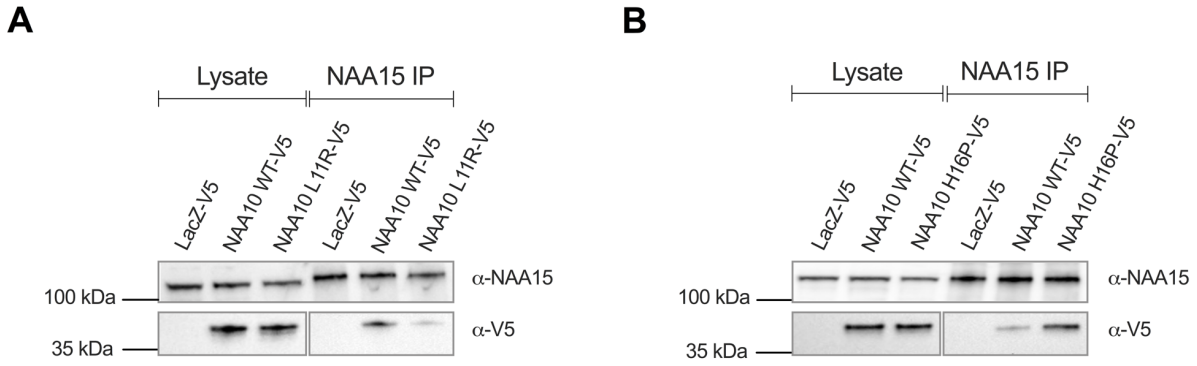

Figure 5.11: NatA complex formation analysis by α-NAA15 IP. A) Western blot analysis of V5-tagged NAA10 WT, NAA10 L11R and LacZ immunoprecipitated from transfected HeLa cells using α-NAA15. **B)** Western blot analysis of V5-tagged NAA10 WT, NAA10 H16P and LacZ immunoprecipitated from transfected HeLa cells using α-NAA15.

6 Discussion

The objective of this thesis was to functionally characterize two NAA10 variants which have been identified in two female patients with varying degrees of delayed language and motor development and syndromic intellectual disability. The NAA10 L11R variant was recently published by Cheng et al., 2019, whereas the NAA10 H16P variant has to my knowledge not been studied. In order to functionally characterize these two missense mutations, mammalian plasmids encoding the NAA10 variants were constructed before HeLa cells were transfected with them. The NatA complex formation and *in vitro* intrinsic catalytic activity were then investigated by IP and ^{14}C-Ac-CoA-based Nt-acetylation assays. The results were analyzed using SDS-PAGE and Western blot. Furthermore, the cellular stability of the two NAA10 variants was assessed by cycloheximide chase experiments.

6.1 Genetic variants and clinical data

There is a wide spectrum of phenotypes and phenotype-severity among individuals with various pathological NAA10 variants. The boys carrying the S37P mutation causing Ogden syndrome had impaired NatA complex formation and catalytic activity and died during infancy. However, mothers and sisters of the boys with the S37P mutation (S37P carriers) were healthy (Rope et al., 2011). A NAA10 splice mutation identified in males with LMS resulted in dysregulation of pathways involved in eye development. The variant was never functionally tested; still it was speculated that the catalytic activity was intact. Since the amount of truncated NAA10 expressed in cells were found to be lower, they proposed that the overall activity would be lower as well (Esmailpour et al., 2014). The males identified with the NAA10 Y43S mutation had a more reduced catalytic activity compared to the boys with Ogden syndrome, yet the phenotypes were milder (Casey et al. in 2015). For girls carrying the V107F, F128I, F128L, R116W and R83C mutations, Saunier et al. concluded that it is the loss of catalytic activity, which is a common feature of these missense mutations, that is the underlying reason for the similar phenotypes among the girls (Saunier et al., 2016). Ree et al. reported a R83H mutation identified in 2 unrelated boys that resulted in reduced monomeric NAA10 activity (Ree et al., 2019). They concluded that it was likely due to altered charge density in the Ac-CoA binding region of NAA10, resulting in an impaired enzyme-Ac-CoA binding. For the V111G and I72T mutations, both mutations showed a reduced stability and enzymatic activity of monomeric NAA10 and unaltered NatA activity (Mc Tiernan et al., 2018, Støve et al., 2018). Cheng et al., 2019 presented several NAA10 mutations that had one or more of these features: reduced NatA activity and reduced stability in presence and absence of HYPK and Ac-CoA (Cheng et al., 2019). The various mutations affect NAA10 differently. Some NAA10 variants have functional NatA activity but reduced monomeric NAA10 activity, while other variants have an unaltered monomeric NAA10 activity but the NatA activity is affected Others destabilize NAA10 and some NAA10 mutations are associated with defects in DNA methylation and genomic imprinting (Lee et al., 2017).

The individuals harboring NAA10 mutations share some general characteristics including intellectual disabilities, developmental delay, postnatal growth failure, skeletal and cardiac abnormalities (Saunier et al., 2016). The mechanism for how impaired Nt-acetylation of NatA or monomeric NAA10 substrates or impaired functions such as KAT, NAT or non-catalytic regulation events could induce the severe phenotype observed in patients carrying NAA10 mutations is still not clear. However, differences in acetylation status of multiple target proteins and/or complexes in different tissues have been speculated to contribute to the differences in phenotype and severity observed in patients.

In the last decade, boys harboring the NAA10 mutations S37P, D10G, I72T and R83C have died (Cheng et al., 2019). There are probably many more cases, but they are not registered because sequencing is not common in cases involving early death. To my knowledge, none of the girls identified with NAA10 mutations have had a lethal outcome, however, phenotypes have been characterized as severe. It has been postulated that missense mutations may generally cause more profound effects in males because they only have one *NAA10* allele and therefore inability to have X-chromosome inactivation (XCI) (Saunier et al., 2016). Skewed X-chromosome inactivation has been observed in females carrying *NAA10* variants. However, skewing is not functionally demonstrated in all cases, making it difficult to speculate about the impact of X-chromosome skewing on the phenotypic expression in those individuals (Cheng et al., 2019). The girl identified with the NAA10 L11R mutation share some of the phenotypes observed in patients identified with NAA10 deficiency. She has motor disorder (pyramidal syndrome), no speech development, strabismus and feeding difficulties. She also has a large arachnoid cyst between the cerebellum and the occipital area. Skewing was not functionally demonstrated in this case. The girl harboring the NAA10 H16P mutation shares some of the general phenotypic characteristics associated with NAA10 deficiency. She has severe syndromic intellectual disability, severe delayed motor and language development. In addition, she shows a disturb behavior. Skewing was functionally demonstrated, indicating that she inherited the H16P mutation maternally.

6.2 Functional analysis of NAA10 variants

6.2.1 NAA10 L11R variant

NAA10 L11 is completely conserved in all the surveyed eukaryotes, suggesting it is important for maintaining protein function or stability (Figure 5.2). The *in silico* predictions by PMut and SIFT of NAA10 L11R indicated that the mutation would affect protein function (Table 5.2) and has a pathological effect (Table 5.1). As for the analysis of the predicted overall stability and the ability to adapt the torsion angle, CUPSAT predicted L11R to be favorable overall (Table 5.3). The calculations performed in CUPSAT are based on atom potentials and torsion angle potentials where in case of unfavorable torsion angles; the atom potentials may have higher impact on stability that results in a stabilizing mutation. The potential energy of a molecule is described by a system of functions known as force field (Wu, 2008). The potential energy includes bond length contribution, bond angle contribution, torsion angle

contribution, electrostatic contribution and van der Waals contribution (Wu, 2008). As seen in the structural analysis of the L11R mutation (Figure 5.4), arginine interacts with more side groups relative to leucine, explaining why the potential energy is greater for arginine.

According to the structural analysis, L11 is located in α1 where its side chain creates a hydrophobic pocket together with the side chains of P8 and M14 on α1 and M28, Y31, F32 on α2 of NAA10 (Figure 5.4). L11 is found to be highly conserved in the catalytic subunit of NatA, NatB and NatC from *S. pombe* and humans, indicating an important biological function among the NATs conserved throughout evolution (Liszczak et al., 2013). The substitution of a leucine with an arginine will probably introduce changes in the hydrophobic surface leucine occupies, which might affect the highly conserved α1-loop-α2 region that make up parts of the NatA substrate binding site and forms hydrophobic interfaces upon association with NAA15 (Liszczak et al., 2013). The mutation of a leucine to arginine which has an amphipathic side chain could potentially cause loss of interaction in the hydrophobic surface between α1 and α2 in NAA10 and introduce stiffening of the structure between α1 and α2 of NAA10 (Figure 6.1). The completely conserved residues L22, E24, Y26 that are important for substrate binding are found in the α1-α2-loop. Loss of flexibility in this key region can result in affecting substrate binding or even global structural integrity, which again might affect NAA10-dependent cellular pathways.

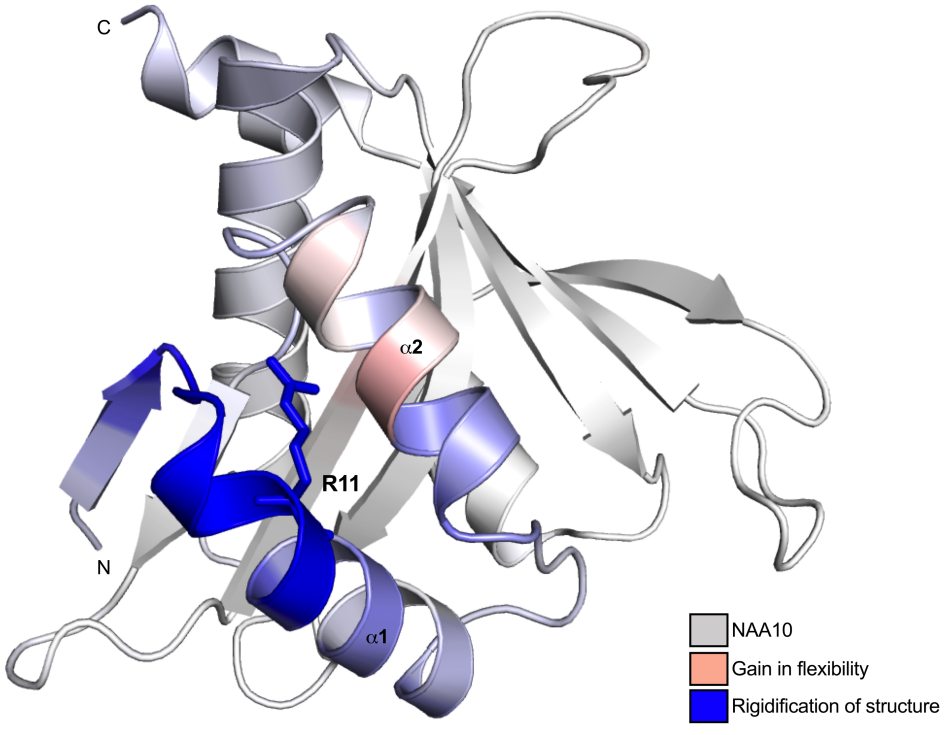

Figure 6.2: The L11R mutation introduce loss of flexibility in α1 and α2 of NAA10. The human structure of NatA (PDB: 6C9M) was used for structural analysis by DynaMut. Amino acids are colored according to the vibrational entropy change ($\Delta\Delta S_{Vib}$, kcal.mol^{-1}.K^{-1}) upon mutation. Blue represents a rigidification of the structure.

The Western blot analysis of the V5-IP experiment revealed a 44 % higher yield of NAA10-V5 for NAA10 L11R-V5 were immunoprecipitated relative to NAA10 WT, still less NAA15 was co-immunoprecipitated with NAA10 L11R relative to WT (Figure 5.9A). This observation could imply that the mutation decreases the binding affinity for NAA15. This presumption was supported by the results from the NAA15-IP experiment. The Western blot analysis of the NAA15-IP experiment showed that less NAA10 L11R-V5 was co-immunoprecipitated with NAA15 compared to NAA10 WT-V5 (Figure 5.11A). This experiment was only done twice due to time limitations and the results were inconsistent, were one experiment showed that less NAA10 L11R-V5 was co-immunoprecipitated with NAA15 and the other experiment showed that less NAA10 L11R-V5 was co-immunoprecipitated with NAA15 relative to WT. Therefore, the experiment needs to be repeated several times in order to make any certain conclusions. When the measured catalytic activity was normalized to α-V5 bands, the NAA10-V5 L11R activity toward the monomeric NAA10 substrate EEEI was almost identical relative to WT. However, the NAA10-V5 L11R activity toward the NatA substrate SESS was reduced 1.8-fold relative to WT when normalized to α-NAA15 bands (Figure 5.9B). Product formation of the Nt-acetylated NatA substrate SESS is catalyzed with a 1000-fold higher activity when NAA10 is in complex with NAA15, relative to monomeric NAA10 (Myklebust et al., 2015). Because NAA10 and NAA15 tightly bind in a 1:1 relationship, the amount of NAA15 can serve as a measurement of the amount of NatA complex present in each sample (Liszczak et al., 2013). The 1.8-fold decrease in activity of NAA10 L11R toward SESS when correlated to α-NAA15 bands can be explained by the fact that NAA10 L11R has a reduced NatA catalytic activity as well as a reduced NatA complex formation. The impaired Nt-acetyltransferase activity and complex stability of NatA could be caused by the possibly reduced local flexibility affecting key regions for substrate binding in NAA10, resulting in a less optimal substrate binding. If the mutation slightly changes the flexibility of the conserved GNAT structure, this could possibly result in perturbations of the NAA10–NAA15 interaction interface explaining the reduced NatA complex formation.

The L11R mutant resembles S37P, Y43S, I72T, V107, F128I/L and V111G on the structural level where they likely promote changes in the hydrophobic surfaces that they occupy. Structural analysis of the S37P mutation revealed perturbations of the NAA10–NAA15 interaction interface and changes in the flexibility of key regions for substrate binding which would explain the impaired Nt-acetyltransferase activity and complex formation of NatA for the S37P mutant. The L11R mutation has corresponding results including reduced NatA activity and complex formation. Cheng et al. recently examined how NAA10 L11R affects the NatA enzymatic activity in presence and absence of HYPK and NatA thermostability in presence and absence of Ac-CoA (Cheng et al., 2019). They found that the mutation did not have a significant effect on NatA activity in presence of HYPK. However, the mutation had a decreased NatA activity in absence of HYPK relative to WT. As for the stability examination, they found that the protein variant was destabilized in the absence as well as the presence of Ac-CoA. They suggested that the NatA activity was unaffected in presence of HYPK due to packing-induced stabilization between the NAA15 helical stalk (α26-44) and the NAA10 α1-

helix. As for the stability of NAA10 L11R, Cheng et al., 2019 suggested that the mutation destabilize the conserved GNAT fold, resulting in a destabilization in the absence as well as presence of Ac-CoA. Hence, there are reasons to speculate that the insertion of an arginine will result in changes in the flexibility of key regions for substrate binding.

When considering the activity toward the monomeric NAA10 substrate EEEI (Figure 5.9B), it appears that the monomeric function of NAA10 L11R is unaffected. Whether the potential KAT and non-acetylating roles of NAA10 are impaired by the L11R mutation remains unknown. However, if L11R affects substrate binding or even global structural integrity there is reason to suspect that KAT activity and protein interactions may be impaired as well.

6.2.2 NAA10 H16P variant

An examination of the conservation of NAA10 H16 revealed some variations between species (Figure 5.2). However, the amino acid in position 16 in all species investigated, except *A. thaliana*, is a polar amino acid that can form hydrogen bonds. Thus, this suggests that amino acid variation in this position does not affect protein function and/or stability as long as the amino acid properties are conserved. According to the *in silico* analysis of NAA10 H16P, the mutation was predicted by PMut to have a pathological effect (Table 5.1) and CUPSAT predicted NAA10 H16P to be an overall destabilizing substitution (Table 5.3). In contrast, NAA10 H16P was considered tolerated by SIFT with a score of 0.06, just above the threshold (Table 5.2). According to SIFT, "tolerated" predictions are more reliable the closer to 1.0 they are. This particular score for H16P is right above the threshold for what SIFT predicts to be tolerated (above 0.05), and therefore not very reliable. The other bioinformatic tools applied in this study predict the H16P mutation to be deleterious (Table 8.2, Supplementary), which is also verified by the experimental data discussed below.

An examination of the human NatA structure revealed H16 to be positioned at the middle of the α1 helix with its side chain protruding out of the protein toward the groove between NAA10 and NAA15 where IP_6 is located. Histidine in position 16 is able to make contact with IP_6 and W486 and F551 of NAA15 (Figure 5.5A). By introducing a proline in position 16, the flexibility in this area will most likely increase and the polypeptide becomes more rigid. Proline is a nitrogen-containing ring where the side chain is connected to the protein backbone twice. Introducing a helix breaker such as proline at any other position than the first three positions of the helix may potentially introduce a bend. Proline is unable to occupy many of the main chain conformations easily adopted by all other amino acids, making the polypeptide chain more rigid. Proline can introduce a bend in the α1-helix that might increase vibrational entropy energy resulting in local flexibility around the mutation (Figure 6.2). The departure from the ideal helical structure and the reduction in hydrogen bonds can contribute to flexibility of a proline-containing α-helix. Underlying vibrational entropy contributes significantly to the binding free energies of proteins (Frappier et al., 2015; Goethe et

al., 2014). Local flexibility of the proline bend can confer with substrate binding or structural integrity related to functional mechanisms, which might again affect downstream pathways.

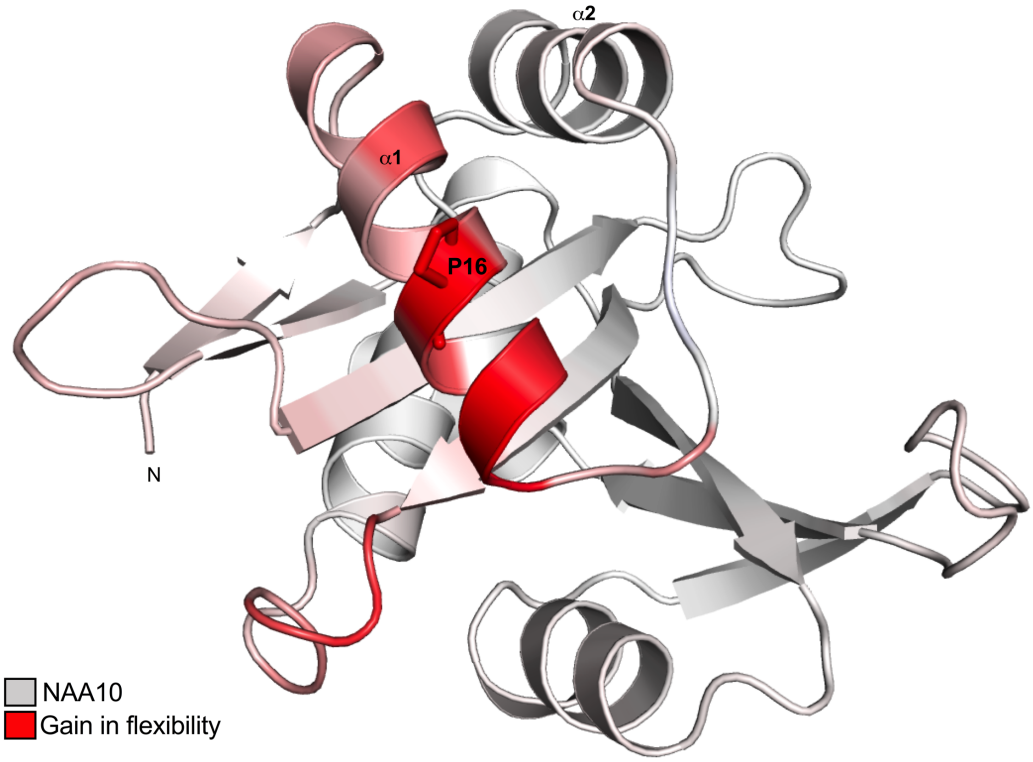

Figure 6.2: The H16P mutation cause gain of flexibility in α1 and α1- α2 loop of NAA10. The human structure of NatA (PDB: 6C9M) was used for structural analysis by DynaMut. Amino acids colored according to the vibrational entropy change ($\Delta\Delta S_{Vib}$, kcal.mol^{-1}.K^{-1}) upon mutation. Red is gain of flexibility.

Proline is introduced at position nine in the α1-helix and might not be able to make interactions with IP_6 and W486 and F551 of NAA15 (Figure 5.5B). Because ligand binding is often important for protein function, this function might be disturbed by this mutation. IP_6 is suggested to be important for stability of NAA10 and NAA15 and therefore might be important in NatA function (Gottlieb et al., 2018). IP_6 is the most abundant of the intracellular inositol phosphates in eukaryotes and is known to be important for maintaining homeostasis, store phosphate and act as a strong antioxidant and neurotransmitter (Kalam Shamsuddin and Bose, 2012). In addition, due to the rapid phosphate turnover properties, it plays a crucial role in cellular pathways involved in signaling transduction and ATP regeneration (Kalam Shamsuddin and Bose, 2012). The substitution of H16 with proline could potentially cause loss of interactions with NAA15 residues and/or IP_6 and impair NatA stability.

The predicted interaction of NAA10 H16 to NAA15 W486 and F551 might be important for stability of the NatA complex, and loss of interaction with these residues may lead to reduced stability of the NatA complex and/or disrupt NatA function. According to the results from the stability assay (Figure 5.8), the mutation is unlikely to affect neither monomeric NAA10 nor NatA stability. However, given the results from the complex formation and activity analysis of NatA, the mutation causes a reduced activity of NatA relative to NAA10 WT-V5 (Figure

5.10). The Western blot analysis revealed that more NAA10 H16P-V5 was immunoprecipitated than NAA10 WT-V5. Despite this, less NAA15 was co-immunoprecipitated with NAA10 H16P-V5 relative to NAA10 WT-V5. With overexpression there is often a greater chance to have more monomeric NAA10 WT than in complex, and consequently more monomeric NAA10 WT is pulled down by IP. However, since the expression of NAA10 WT and H16P variant is quite similar (lysate, Figure 5.10A), and the cellular stability of the H16P variant was not affected (Figure 5.8), there are reasons to assume that the amount of NAA10 WT and H16P in both complex and monomeric form is similar. The most obvious explanation is that NAA10 H16P-V5 has a lower affinity for NAA15. An IP experiment with α-NAA15 was performed to verify this. The results were unexpected and showed that more NAA10 H16P-V5 co-immunoprecipitated with NAA15 relative to NAA10 WT-V5 (Figure 5.11B). This result implies that H16P has a higher affinity for NAA15 relative to WT. This is inconsistent with the results from the V5-IP which indicated a reduced affinity of NAA10 H16P for NAA15. The structural analysis also suggested that the mutation might impair complex formation through loss of interactions with NAA15 residues and IP_6. Since the NAA15-IP was only performed twice and the results are contradicting the observations from both the V5-IP (performed 6 times) and structural analysis, this experiment should be repeated to exclude experimental errors before making any conclusions.

When the measured catalytic activity was normalized to α-V5 bands, the NAA10-V5 H16P activity toward the monomeric NAA10 substrate EEEI was almost identical relative to WT. Conversely, the NAA10-V5 H16P activity toward the NatA substrate SESS was reduced almost 4-fold compared to WT when normalized to α-NAA15 bands (Figure 5.10B). The fact that the H16P mutation possibly results in loss of interactions with IP_6 and NAA15 residues, believed to be involved in NAA10–NAA15 interaction interface, could explain the reduced NatA complex formation and the reduced NatA activity. The current results reveal that while the H16P variants has a reduced overall NatA activity, it is unlikely to affect neither monomeric NAA10 nor NatA stability (Figure 5.10B, Figure 5.8). Therefore, it is possible that the gain in local flexibility caused by the proline bend can confer with key regions and introduce a slight conformational change in NAA10, resulting in perturbations of the NAA10–NAA15 interaction interface. The NAA10 H16P variant showed a large reduction in NatA activity which resembles the NatA activity of the NAA10 S37P variant, in addition to the two newly identified NAA10 variants D10G and A104D reported by Cheng et al., 2019. Cheng et al, suggest that the protein changes of D10G and A104D likely alter numerous interactions including the loss of hydrogen bond interactions and the perturbation of hydrophobic pockets, consistent with the predictions of H16P. Proline and glycine is well known helix breakers (Imai et al., 2005). The NAA10 D10G mutation is found in the same α-helix as H16P. In contrast to proline that is to rigid to be incorporated in a local ordered structure, glycine is to flexible. Entropy effects drive the defined secondary structure into a loop (Imai et al., 2005).

6.3 Conclusion

During this study, newly identified *NAA10* mutations have been examined in terms of *in silico* structural and functional analysis. The cellular stability, *in vitro* NatA complex formation and intrinsic catalytic activity were analyzed as well. This investigation was performed to assess the mutations effect on NAA10 and to establish whether they are causative of disease. The NAA10 L11R showed a impaired NatA complex formation, whereas NAA10 H16P does not impair NatA complex formation. The NAA10 L11R and H16P variants were both shown to have a reduced NatA activity (Figure 6.3). The L11R variant seems to impair the NatA activity to a less extent than the H16P mutation. While examining the mutations in this thesis, Cheng et al, reported functional analysis of the NAA10 L11R variant. The NatA activity data of the L11R variant found in this thesis is consistent with the activity data reported by Cheng et al., 2019. In addition, the reduced NatA activity of NAA10 L11R is similar to the reduced NatA activity reported for the NAA10 Y43S and the I72T variant. The H16P variant showed a larger decrease in NatA activity, which is more similar to the D10G, S37P and A104D variants (Cheng et al., 2019). Cellular stability assays showed that NAA10 L11R is unstable, whereas NAA10 H16P is stable and is unlikely to affect neither monomeric NAA10 nor NatA stability. The data presented in this thesis support the hypothesis that a major factor contributing to disease in the two females, is loss of Nt-acetylation of specific NatA substrates due to reduced catalytic activity and reduced NatA complex formation. Further studies are however needed for both these variants in order to understand the consequences and underlying mechanism of these NAA10 variants *in vivo*.

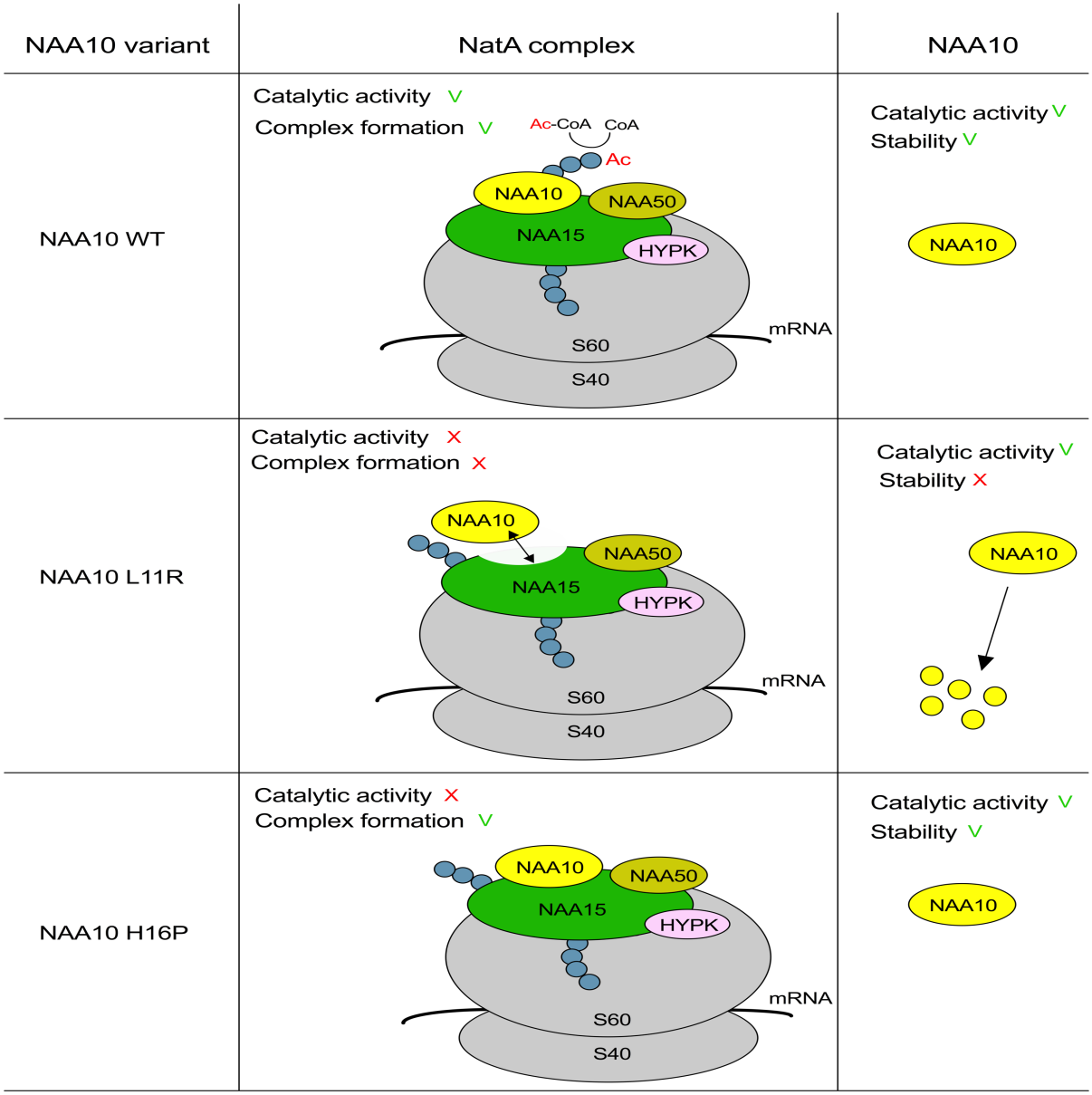

Figure 6.3: Graphical illustration of potential functional implications of NAA10 variants. NAA10 WT exerts Nt-acetylation of specific substrates when bound to NAA15 (the NatA complex). When NAA10 is not part of the NatA complex, monomeric NAA10 exerts NatA independent roles in the cells. NAA10 L11R has a reduced stability, NatA complex formation and a reduced *in vitro* NatA activity. NAA10 H16P is stable and is unlikely to affect neither monomeric NAA10 nor NatA stability. NAA10 H16P appears to not impair NatA complex formation, but the variant has a reduced activity toward NatA substrates.

6.4 Future perspective

Gene transcription, DNA-damage response cell cycle arrest, stress response, cell motility and bone and neuronal development are some of the implicated functions of NAA10 (Dörfel and Lyon, 2015). Still, the full extent of the many and diverse functions of NAA10 in the cell remain unknown. The NatA complex targets many substrates, expected to affect diverse cellular processes in many different organs. Various mutations may therefore cause distinct downstream Nt-acetylation defects that result in the wide spectrum of physiological

consequences. NatA independent functional roles of NAA10 can be affected by mutations, showing the massive complexity of NAA10 deficiency in human disease. Given that NAA10 carries out a wide spectrum of functions, the cellular-, tissue- and developmental stage-specific effects of NAA10 need to be investigated. The first focus of future work should be to finalize the functional characterization of the mutations studied in this thesis. In particular, due to the inconsistent results in α-NAA15 IP experiments, these experiments should be repeated to be able to verify if the NAA10 L11R mutation or the NAA10 H16P mutation has altered the binding of NAA15 to NAA10.

In order to investigate the impact of NAA10 H16P and NAA10 L11R on the structural level, potential changes in the secondary structure could be detected using circular dichroism (CD). Circular dichroism (CD) can be used to analyze whether the mutation affects its conformation or stability. In addition, it can be used to study protein interactions. This can be used to investigate the potential loss of interaction between NAA10 and NAA15.

Another goal of future studies will be to elaborate on the cellular/molecular specific activity and distinct pathways of NAA10. One potential approach can be to investigate the impact of *NAA10* variants at endogenous levels. This can be done by generating cell lines containing the specific *NAA10* mutation using mammalian cells such as HAP-1 cell lines with CRISPR-Cas9. This will ensure an identical genetic background between different NAA10 mutant cell lines. Then, the mutation's impact on NAA10's endogenous levels, interaction partners, subcellular localization, and cell growth and mobility can be examined. These cells can be submitted to detailed proteomic studies such as quantification of global protein levels to look for changes in signal pathways, investigations of N-terminal acetylation and analyzes of lysine acetylation events to find possible affected KAT events. Further, endogenous proteins can be fused to a fluorescent report protein created by employing CRISPR/Cas9 genome editing. These fluorescently tagged proteins can be used to visualize protein function and dynamics in real time as cells proliferate and differentiate.

Another potential study can be the same approach as Van Damme et al., 2014 used. They used ectopic expression of human NAA15 and NAA10 variant in *S. cerevisiae* where Naa10 and Naa15 have been deleted (Van Damme et al., 2014). Then, the proteomic method Nt-combined fractional diagonal chromatography (COFRADIC) (Arnesen et al., 2009), can be used to study the potential altered level of substrates or Nt-acetylation of the catalytic subunits NatA and NatE. This approach has been performed on the S37P variant (Myklebust et al., 2015, Van Damme et al., 2014).

Above all, *in vivo* studies are necessary to analyze the definite biological effects of NAA10. CRISPR generated model organism such as zebrafish and mice can be used for studying NAA10 *in vivo*. With the use of *in vivo* studies, proteome-wide N-terminal- and potential lysine acetylation patterns as well as non- acetylating functions of NAA10 can be investigated. Studying the effects of NAA10 deficiency *in vivo* may reveal underlying mechanism, as well as unknown functions of NAA10.

7 References

Adzhubei, I., Schmidt, S., Peshkin, L., Ramensky, V., Gerasimova, A., Bork, P., … Sunyaev, S. (2010). PolyPhen-2 : prediction of functional effects of human nsSNPs. *Curr Protoc Hum Genet*. https://doi.org/10.1017/CBO9781107415324.004

Aksnes H, Drazic A, Marie M, Arnesen T. First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. *Trends Biochem Sci*. 2016;41(9):746-760. doi:10.1016/j.tibs.2016.07.005.

Aksnes, H., Ree, R., & Arnesen, T. (2019). Co-translational, Post-translational, and Non-catalytic Roles of N-Terminal Acetyltransferases. *Molecular Cell*. doi.org/10.1016/j.molcel.2019.02.007

Aksnes, H., Van Damme, P., Goris, M., Starheim, K. K., Marie, M., Støve, S. I., … Arnesen, T. (2015). An organellar nα-acetyltransferase, naa60, acetylates cytosolic n termini of transmembrane proteins and maintains golgi integrity. *Cell Reports*. doi.org/10.1016/j.celrep.2015.01.053

Arnesen, T., Anderson, D., Baldersheim, C., Lanotte, M., Varhaug, J.E., and Lil- lehaug, J.R. (2005a). Identification and characterization of the human ARD1- NATH protein acetyltransferase complex. Biochem. J. *386*, 433–443.

Arnesen, T., Gromyko, D., Pendino, F., Ryningen, A., Varhaug, J. E., & Lillehaug, J. R. (2006). Induction of apoptosis in human cells by RNAi-mediated knockdown of hARD1 and NATH, components of the protein N-α-acetyltransferase complex. *Oncogene*. doi.org/10.1038/sj.onc.1209469

Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J.E., Vandekerckhove, J., Lillehaug, J.R., Sherman, F., and Ge- vaert, K. (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. Proc. Natl. Acad. Sci. USA *106*, 8157–8162.

Arnesen T, Starheim KK, Van Damme P, et al. The Chaperone-Like Protein HYPK Acts Together with NatA in Cotranslational N-Terminal Acetylation and Prevention of Huntingtin Aggregation. *Mol Cell Biol*. 2010;30(8):1898-1909. doi:10.1128/MCB.01199-09.

Asaumi M, Iijima K, Sumioka A, et al. Interaction of N-terminal acetyltransferase with the cytoplasmic domain of β-amyloid precursor protein and its effect on Aβ secretion. *J Biochem*. 2005;137(2):147-155.

Behnia R, Panic B, Whyte JRC, Munro S. Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p. *Nat Cell Biol*. 2004;6(5):405-413. doi:10.1038/ncb1120.

Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays *21*, 932–939.

Boutureira, Omar & Bernardes, Goncalo J. L. (2015) Advances in Chemical Protein Modification. *Chemical Reviews,* 115**,** 2174-2195.

Casey, J.P., Støve, S.I., McGorrian, C., Galvin, J., Blenski, M., Dunne, A., En- nis, S., Brett, F., King, M.D., Arnesen, T., and Lynch, S.A. (2015). NAA10 mu- tation causing a novel intellectual disability syndrome with Long QT due to N-terminal acetyltransferase impairment. Sci. Rep. *5*, 16022.

Cheng, H., Gottlieb, L., Marchi, E., Kleyner, R., Bhardwaj, P., Rope, A. F., … Lyon, G. J. (2019). Phenotypic and biochemical analysis of an international cohort of individuals with variants in NAA10 and NAA15. *Human Molecular Genetics*. https://doi.org/10.1093/hmg/ddz111

Cheol-Sang H, Shemorry A, Varshavsky A. N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals. *Science*. 2010;327(5968):973-977. doi:10.1126/science.1183147.N- Terminal.

Chien, M. H., Lee, W. J., Yang, Y. C., Tan, P., Pan, K. F., Liu, Y. C., … Hua, K. T. (2018). N-α-acetyltransferase 10 protein promotes metastasis by stabilizing matrix metalloproteinase-2 protein in human osteosarcomas. *Cancer Letters*. doi.org/10.1016/j.canlet.2018.06.033

Chiti F, Dobson CM. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu Rev Biochem*. 2006;75(1):333-366. doi:10.1146/annurev.biochem.75.101304.123901.

Chu, C.W., Hou, F., Zhang, J., Phu, L., Loktev, A.V., Kirkpatrick, D.S., Jackson, P.K., Zhao, Y., and Zou, H. (2011). A novel acetylation of beta-tubulin by San modulates microtubule polymerization via down-regulating tubulin incorpora- tion. Mol. Biol. Cell *22*, 448–456.

Deng, S., Magin, R. S., Wei, X., Pan, B., Petersson, E. J., & Marmorstein, R. (2019). Structure and Mechanism of Acetylation by the N-Terminal Dual Enzyme NatA/Naa50 Complex. *Structure*. doi.org/10.1016/j.str.2019.04.014

Dikiy I, Eliezer D. N-terminal Acetylation stabilizes N-terminal Helicity in Lipid- and Micelle-bound α- Synuclein and increases its affinity for Physiological Membranes. *J Biol Chem*. 2014;289(6):3652-3665. doi:10.1074/jbc.M113.512459.

Dinh, T. V., Bienvenut, W. V., Linster, E., Feldman-Salit, A., Jung, V. A., Meinnel, T., … Wirtz, M. (2015). Molecular identification and functional characterization of the first Nα-acetyltransferase in plastids by global acetylome profiling. *Proteomics*. doi.org/10.1002/pmic.201500025

Drazic, A., Myklebust, L. M., Ree, R., & Arnesen, T. (2016). The world of protein acetylation. *Biochimica et Biophysica Acta - Proteins and Proteomics*. doi.org/10.1016/j.bbapap.2016.06.007

Drazic, A., Aksnes, H., Marie, M., Boczkowska, M., Varland, S., Timmerman, E., … Arnesen, T. (2018). NAA80 is actin's N-terminal acetyltransferase and regulates cytoskeleton assembly and cell motility. *Proceedings of the National Academy of Sciences of the United States of America*. doi.org/10.1073/pnas.1718336115

Esmailpour, Taraneh, et al. (2014) A splice donor mutation in results in the dysregulation of the retinoic acid signalling pathway and causes Lenz microphthalmia syndrome. *Journal of Medical Genetics,* 51**,** 185-196.

Evjenth R, Hole K, Karlsen OA, Ziegler M, Amesen T, Lillehaug JR. Human Naa50p (Nat5/San) displays both protein Nα- and Nε-acetyltransferase activity. *J Biol Chem*. 2009;284(45):31122-31129. doi:10.1074/jbc.M109.001347.

Evjenth, R. H., Brenner, A. K., Thompson, P. R., Arnesen, T., Frøystein, N. Å., & Lillehaug, J. R. (2012). Human protein N-terminal acetyltransferase hNaa50p (hNAT5/hSAN) follows ordered sequential catalytic mechanism: Combined kinetic and NMR study. *Journal of Biological Chemistry*. doi.org/10.1074/jbc.M111.326587

Fluge Ø, Bruland O, Akslen LA, Varhaug JE, Lillehaug JR. NATH, a novel gene overexpressed in papillary thyroid carcinomas. *Oncogene*. 2002;21(33):5056-5068. doi:10.1038/sj.onc.1205687

Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R. C., Giglione, C., & Meinnel, T. (2006). The proteomics of N-terminal methionine cleavage. *Molecular and Cellular Proteomics*. doi.org/10.1074/mcp.M600225-MCP200

Gautschi, M., Just, S., Mun, A., Ross, S., Rucknagel, P., Dubaquie, Y., … Rospert, S. (2003). The Yeast N -Acetyltransferase NatA Is Quantitatively Anchored to the Ribosome and Interacts with Nascent Polypeptides. *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.23.20.7403-7414.2003

Glozak, M. A., Sengupta, N., Zhang, X., & Seto, E. (2005). Acetylation and deacetylation of non-histone proteins. *Gene*. doi.org/10.1016/j.gene.2005.09.010

Goetze, S., Qeli, E., Mosimann, C., Staes, A., Gerrits, B., Roschitzki, B., Mo- hanty, S., Niederer, E.M., Laczko, E., Timmerman, E., et al. (2009). Identifica- tion and functional characterization of N-terminally acetylated proteins in Drosophila melanogaster. PLoS Biol. *7*, e1000236.

Goris, M., Magin, R. S., Foyn, H., Myklebust, L. M., Varland, S., Ree, R., … Arnesen, T. (2018). Structural determinants and cellular environment define processed actin as the sole substrate of the N-terminal acetyltransferase NAA80. *Proceedings of the National Academy of Sciences of the United States of America*. doi.org/10.1073/pnas.1719251115

Gottlieb, L., and Marmorstein, R. (2018). Structure of Human NatA and Its Regulation by the Huntingtin Interacting Protein HYPK. Structure *26*, 925–935.e8.

Graille, M., Stura, E. A., Corper, A. L., Sutton, B. J., Taussig, M. J., Charbonnier, J. B., & Silverman, G. J. (2000). Crystal structure of a Staphylococcus aureus protein a domain complexed with the Fab fragment of a human IgM antibody: Structural basis for recognition of B-cell receptors and superantigen activity. *Proceedings of the National Academy of Sciences of the United States of America*. doi.org/10.1073/pnas.97.10.5399

Gromyko, D., Arnesen, T., Ryningen, A., Varhaug, J. E., & Lillehaug, J. R. (2010). Depletion of the human Nα-terminal acetyltransferase A induces p53-dependent apoptosis and p53-independent growth inhibition. *International Journal of Cancer*. doi.org/10.1002/ijc.25275

Hole, K., van Damme, P., Dalva, M., Aksnes, H., Glomnes, N., Varhaug, J. E., … Arnesen, T. (2011). The human N-Alpha-acetyltransferase 40 (hNaa40p/hNatD) is conserved from yeast and N-terminally acetylates histones H2A and H4. *PLoS ONE*. doi.org/10.1371/journal.pone.0024713

Holmes WM, Mannakee BK, Gutenkunst RN, Serio TR. Loss of amino-terminal acetylation suppresses a prion phenotype by modulating global protein folding. *Nat Commun*. 2014;5:1-11. doi:10.1038/ncomms5383.

Hou, F., Chu, C.-W., Kong, X., Yokomori, K., and Zou, H. (2007). The acetyl- transferase activity of San stabilizes the mitotic cohesin at the centromeres in a shugoshin-independent manner. J. Cell Biol. *177*, 587–597.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*, 973–977.

Idusogie, E. E., Presta, L. G., Gazzano-Santoro, H., Totpal, K., Wong, P. Y., Ultsch, M., … Mulkerrin, M. G. (2000). Mapping of the C1q Binding Site on Rituxan, a Chimeric Antibody with a Human IgG1 Fc. *The Journal of Immunology*. doi.org/10.4049/jimmunol.164.8.4178

Ingram, A.K., Cross, G.A., and Horn, D. (2000). Genetic manipulation indicates that ARD1 is an essential N(alpha)-acetyltransferase in Trypanosoma brucei. Mol. Biochem. Parasitol. *111*, 309–317.

International Human Genome Sequencing Consortium. (2004). International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. doi.org/nature03001 [pii]\r10.1038/nature03001

Jeong, J.W., Bae, M.K., Ahn, M.Y., Kim, S.H., Sohn, T.K., Bae, M.H., Yoo, M.A., Song, E.J., Lee, K.J., and Kim, K.W. (2002). Regulation and destabiliza- tion of HIF-1alpha by ARD1-mediated acetylation. Cell *111*, 709–720.

Kalvik, T.V., and Arnesen, T. (2013). Protein N-terminal acetyltransferases in cancer. Oncogene *32*, 269–276.

Kang, L., Moriarty, G.M., Woods, L.A., Ashcroft, A.E., Radford, S.E., and Baum, J. (2012). N-terminal acetylation of alpha-synuclein induces increased transient helical propensity and decreased aggregation rates in the intrinsically disordered monomer. Protein Sci. *21*, 911–917.

Kang, J., Chun, Y.S., Huh, J., and Park, J.W. (2018). FIH permits NAA10 to catalyze the oxygen-dependent lysyl-acetylation of HIF-1a. Redox Biol. *19*, 364–374.

Kats I, Khmelinskii A, Kschonsak M, et al. Mapping Degradation Signals and Pathways in a Eukaryotic N-terminome. *Mol Cell*. 2018;70(3):488-501.e5. doi:10.1016/j.molcel.2018.03.033.

Kemena, C., & Notredame, C. (2009). Upcoming challenges for multiple sequence alignment methods in the high-throughput era. *Bioinformatics*. doi.org/10.1093/bioinformatics/btp452

Kim, Min-Sik, et al. v(2014) A draft map of the human proteome. *Nature,* 509**,** 575-581.

Kim, Sung Chan, et al. (2006) Substrate and Functional Diversity of Lysine Acetylation Revealed by a Proteomics Survey. *Molecular Cell,* 23**,** 607-618.

Knorr, A.G., Schmidt, C., Tesina, P., Berninghausen, O., Becker, T., Beatrix, B., and Beckmann, R. (2018). Ribosome–NatA architecture reveals that rRNA expansion segments coordinate N-terminal acetylation. Nat. Struct. Mol. Biol. *26*, 35–39.

Lee, C.F., Ou, D.S., Lee, S.B., Chang, L.H., Lin, R.K., Li, Y.S., Upadhyay, A.K., Cheng, X., Wang, Y.C., Hsu, H.S., et al. (2010). hNaa10p contributes to tumor- igenesis by facilitating DNMT1-mediated tumor suppressor gene silencing. J. Clin. Invest. *120*, 2920–2930.

Lee, C.C., Peng, S.H., Shen, L., Lee, C.F., Du, T.H., Kang, M.L., Xu, G.L., Upadhyay, A.K., Cheng, X., Yan, Y.T., et al. (2017). The Role of N-alpha-acetyltransferase 10 Protein in DNA Methylation and Genomic Imprinting. Mol. Cell *68*, 89–103.e7.

Lee, C.-C., Shih, Y.-C., Kang, M.-L., Chang, Y.-C., Chuang, L.-M., Devaraj, R., & Juan, L.-J. (2019). Naa10p Inhibits Beige Adipocyte-Mediated Thermogenesis through N-α-acetylation of Pgc1α. *Molecular Cell*. doi.org/10.1016/j.molcel.2019.07.026

Linē A, Stengrēvics A, Slucka Z, Li G, Rees RC. Serological identification and expression analysis of gastric cancer-associated genes. Br *J Cancer.* 2002;86(11):1824-1830. doi:10.1038/sj.bjc.6600321

Linster, E., Stephan, I., Bienvenut, W.V., Maple-Grødem, J., Myklebust, L.M., Huber, M., Reichelt, M., Sticht, C., Møller, S.G., Meinnel, T., et al. (2015). Downregulation of N-terminal acetylation triggers ABA-mediated drought re- sponses in Arabidopsis. Nat. Commun. *6*, 7640.

Liszczak, G., Goldberg, J. M., Foyn, H., Petersson, E. J., Arnesen, T., & Marmorstein, R. (2013). Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. *Nature Structural and Molecular Biology*. doi.org/10.1038/nsmb.2636

López-Ferrando, V., Gazzo, A., De La Cruz, X., Orozco, M., & Gelpí, J. L. (2017). PMut: A web-based tool for the annotation of pathological variants on proteins, 2017 update. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkx313

Lozada, E.M., Andrysik, Z., Yin, M., Redilla, N., Rice, K., and Stambrook, P.J. (2016). Acetylation and deacetylation of Cdc25A constitutes a novel mecha- nism for modulating Cdc25A functions with implications for cancer. Oncotarget *7*, 20425–20439.

Magin, R.S., March, Z.M., and Marmorstein, R. (2016). The N-terminal acetyl- transferase Naa10/ARD1 does not acetylate lysine residues. J. Biol. Chem. *291*, 5270–5277.

Martin DT, Gendron RL, Jarzembowski JA, et al. Tubedown expression correlates with the differentiation status and aggressiveness of neuroblastic tumors. *Clin Cancer Res*. 2007;13(5):1480- 1487. doi:10.1158/1078-0432.CCR-06-1716

McTiernan, N., Støve, S. I., Aukrust, I., Mårli, M. T., Myklebust, L. M., Houge, G., & Arnesen, T. (2018). NAA10 dysfunction with normal NatA-complex activity in a girl with non-syndromic ID and a de novo NAA10 p.(V111G) variant - a case report. *BMC Medical Genetics*. https://doi.org/10.1186/s12881-018-0559-z

Midorikawa, Y., Tsutsumi, S., Taniguchi, H., Ishii, M., Kobune, Y., Kodama, T., Makuuchi, M., and Aburatani, H. (2002). Identification of genes associated with dedifferentiation of hepatocellular carcinoma with expression profiling anal- ysis. Jpn. J. Cancer Res. *93*, 636–643.

Mittl, P. R. E., & Schneider-Brachert, W. (2007). Sel1-like repeat proteins in signal transduction. *Cellular Signalling*. doi.org/10.1016/j.cellsig.2006.05.034

Mullen, J. R., Kayne, P. S., Moerschell, R. P., Tsunasawa, S., Gribskov, M., Colavito-Shepanski, M., … Sternglanz, R. (1989). Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. *The EMBO Journal*. doi.org/10.1002/j.1460-2075.1989.tb03615.x

Myklebust, Line M., et al. (2015) Biochemical and cellular analysis of Ogden syndrome reveals downstream Nt-acetylation defects. *Human Molecular Genetics,* 24**,** 1956- 76.

Neubauer, J.L. (2012). Structural Analysis of the N-Terminal Acetyltransferase A Complex (Duke University), p. 136, Dissertation.

Nguyen KT, Mun S-H, Lee C-S, Hwang C-S. Control of protein degradation by N-terminal acetylation and the N-end rule pathway. *Exp Mol Med*. 2018;50(7):91. doi:10.1038/s12276-018-0097-y.

Park, E. C., & Szostak, J. W. (1992). ARD1 and NAT1 proteins form a complex that has N-terminal acetyltransferase activity. *The EMBO Journal*. doi.org/10.1002/j.1460-2075.1992.tb05267.x

Pertea, Mihaela (2012) The Human Transcriptome: An Unfinished Story. *Genes,* 3**,** 344- 360.

Ponomarenko, Elena A., et al. (2016) The Size of the Human Proteome: The Width and Depth. *International Journal of Analytical Chemistry,* 2016.

Popp, Bernt, et al. (2015) De novo missense mutations in the NAA10 gene cause severe non-syndromic developmental delay in males and females. *European Journal of Human Genetics,* 23**,** 602–609.

Rathore, Om Singh, et al. (2016) Absence of N-terminal acetyltransferase diversification during evolution of eukaryotic organisms. *Scientific Reports,* 6**,** 21304.

Raychaudhuri, S., Sinha, M., Mukhopadhyay, D., and Bhattacharyya, N.P. (2008). HYPK, a Huntingtin interacting protein, reduces aggregates and apoptosis induced by N-terminal Huntingtin with 40 glutamines in Neuro2a cells and exhibits chaperone-like activity. Hum. Mol. Genet. *17*, 240–255.

Ree, R., Geithus, A. S., Tørring, P. M., Sørensen, K. P., Damkjær, M., Lynch, S. A., & Arnesen, T. (2019). A novel NAA10 p.(R83H) variant with impaired acetyltransferase activity identified in two boys with ID and microcephaly. *BMC Medical Genetics*. doi.org/10.1186/s12881-019-0803-1

Ree, R., Myklebust, L. M., Thiel, P., Foyn, H., Fladmark, K. E., & Arnesen, T. (2015). The N-terminal acetyltransferase Naa10 is essential for zebrafish development. *Bioscience Reports*. doi.org/10.1042/BSR20150168

Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., and Yao, X. (2008). CSS-Palm 2.0: an updated software for palmitoylation sites prediction. Protein Eng. Des. Sel. *21*, 639–644.

Ribeiro, A.L., Silva, R.D., Foyn, H., Tiago, M.N., Rathore, O.S., Arnesen, T., and Martinho, R.G. (2016). Naa50/San-dependent N-terminal acetylation of Scc1 is potentially important for sister chromatid cohesion. Sci. Rep. *6*, 39118.

Rope, Alan F., et al. (2011) Using VAAST to Identify an X-Linked Disorder Resulting in Lethality in Male Infants Due to N-Terminal Acetyltransferase Deficiency. *American Journal of Human Genetics,* 89**,** 28-43.

Saunier, Chloé, et al. (2016) Expanding the Phenotype Associated with NAA10-Related N-Terminal Acetylation Deficiency. *Human Mutation,* 37**,** 755-764.

Scott DC, Monda JK, Bennett EJ, Harper JW, Schulman BA. N-Terminal Acetylation Acts as an Avidity Enhancer Within an Interconnected Multiprotein Complex. *Science*. 2011;334:674-678. doi:10.1126/science.1209307.

Seo, J.H., Park, J.H., Lee, E.J., Vo, T.T., Choi, H., Kim, J.Y., Jang, J.K., Wee, H.J., Lee, H.S., Jang, S.H., et al. (2016). ARD1-mediated Hsp70 acetylation balances stress-induced protein refolding and degradation. Nat. Commun. *7*, 12882.

Shemorry A, Hwang C-S, Varshavsky A. Control of Protein Quality and Stoichiometries by N-Terminal Acetylation and the N-End Rule Pathway. *Mol Cell*. 2013;50(4):540-551. doi:10.1016/j.molcel.2013.03.018.

Shin, S. H., Yoon, H., Chun, Y. S., Shin, H. W., Lee, M. N., Oh, G. T., and Park, J. W. (2014) Arrest defective 1 regulates the oxidative stress response in human cells and mice by acetylating methionine sulfoxide reductase A. *Cell death & disease* **5**, e1490

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., … Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*. https://doi.org/10.1038/msb.2011.75

Soppa, J. (2010). Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea*. doi.org/10.1155/2010/820681

Song, O. K., Wang, X., Waterborg, J. H., & Sternglanz, R. (2003). An Nα-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A. *Journal of Biological Chemistry*. doi.org/10.1074/jbc.C300355200

Starheim, K. K., Arnesen, T., Gromyko, D., Ryningen, A., Varhaug, J. E., & Lillehaug, J. R. (2008). Identification of the human Nα-acetyltransferase complex B (hNatB): A complex important for cell-cycle progression. *Biochemical Journal*. doi.org/10.1042/BJ20080658

Starheim, K. K., Gromyko, D., Evjenth, R., Ryningen, A., Varhaug, J. E., Lillehaug, J. R., & Arnesen, T. (2009). Knockdown of Human N -Terminal Acetyltransferase Complex C Leads to p53-Dependent Apoptosis and Aberrant Human Arl8b Localization. *Molecular and Cellular Biology*. doi.org/10.1128/mcb.01909-08

Støve, S. I., Blenski, M., Stray-Pedersen, A., Wierenga, K. J., Jhangiani, S. N., Akdemir, Z. C., … Arnesen, T. (2018). A novel NAA10 variant with impaired acetyltransferase activity causes developmental delay, intellectual disability, and hypertrophic cardiomyopathy. *European Journal of Human Genetics*. doi.org/10.1038/s41431-018-0136-0

Ud-Din, A. I. M. S., Tikhomirova, A., & Roujeinikova, A. (2016). Structure and functional diversity of GCN5-related n-acetyltransferases (GNAT). *International Journal of Molecular Sciences*. doi.org/10.3390/ijms17071018

Van Damme, P., Evjenth, R., Foyn, H., Demeyer, K., De Bock, P. J., Lillehaug, J. R., … Gevaert, K. (2011a). Proteome-derived peptide libraries allow detailed analysis of the substrate specificities of Nα-acetyltransferases and point to hNaa10p as the post-translational actin Nα-acetyltransferase. *Molecular and Cellular Proteomics*. doi.org/10.1074/mcp.M110.004580

Van Damme, P., Hole, K., Pimenta-Marques, A., Helsens, K., Vandekerckhove, J., Martinho, R. G., … Arnesen, T. (2011b). NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation. *PLoS Genetics*. doi.org/10.1371/journal.pgen.1002169

Van Damme, P., Lasa, M., Polevoda, B., Gazquez, C., Elosegui-Artola, A., Kim, D. S., … Aldabe, R. (2012). N-terminal acetylome analyses and functional insights of the N-terminal acetyltransferase NatB. *Proceedings of the National Academy of Sciences of the United States of America*. doi.org/10.1073/pnas.1210303109

Van Damme, P., Hole, K., Gevaert, K., & Arnesen, T. (2015). N-terminal acetylome analysis reveals the specificity of Naa50 (Nat5) and suggests a kinetic competition between N-terminal acetyltransferases and methionine aminopeptidases. *Proteomics*. doi.org/10.1002/pmic.201400575

Van Damme, P., Kalvik, T. V., Starheim, K. K., Jonckheere, V., Myklebust, L. M., Menschaert, G., … Arnesen, T. (2016). A role for human N-alpha acetyltransferase 30 (Naa30) in maintaining mitochondrial integrity. *Molecular and Cellular Proteomics*. doi.org/10.1074/mcp.M116.061010

Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. *Proceedings of the National Academy of Sciences of the United States of America*. doi.org/10.1073/pnas.1816596116

Verdin, E., & Ott, M. (2015). 50 years of protein acetylation: From gene regulation to epigenetics, metabolism and beyond. *Nature Reviews Molecular Cell Biology*. doi.org/10.1038/nrm3931

Wang, Y., Mijares, M., Gall, M.D., Turan, T., Javier, A., Bornemann, D.J., Manage, K., and Warrior, R. (2010). Drosophila variable nurse cells encodes arrest defective 1 (ARD1), the catalytic subunit of the major N-terminal acetyl- transferase complex. Dev. Dyn. *239*, 2813–2827.

Wang Z, Wang Z, Guo J, et al. Inactivation of androgen-induced regulator ARD1 inhibits androgen receptor acetylation and prostate tumorigenesis. *PNAS*. 2012;109(8):3053-3058. doi:10.1073/pnas.1113356109.

Weyer, F. A., Gumiero, A., Lapouge, K., Bange, G., Kopp, J., & Sinning, I. (2017). Structural basis of HypK regulating N-Terminal acetylation by the NatA complex. *Nature Communications*. https://doi.org/10.1038/ncomms15726

Wiame, E., Tahay, G., Tyteca, D., Vertommen, D., Stroobant, V., Bommer, G. T., & Van Schaftingen, E. (2018). NAT6 acetylates the N-terminus of different forms of actin. *FEBS Journal*. doi.org/10.1111/febs.14605

Williams, B.C., Garrett-Engele, C.M., Li, Z., Williams, E.V., Rosenman, E.D., and Goldberg, M.L. (2003). Two putative acetyltransferases, san and deco, are required for establishing sister chromatid cohesion in Drosophila. Curr. Biol. *13*, 2025–2036.

Yoon, H., Kim, H.L., Chun, Y.S., Shin, D.H., Lee, K.H., Shin, C.S., Lee, D.Y., Kim, H.H., Lee, Z.H., Ryoo, H.M., et al. (2014). NAA10 controls osteoblast differentiation and bone formation as a feedback regulator of Runx2. Nat. Commun. *5*, 5176.

Yu M, Ma M, Huang C, et al. Correlation of expression of human arrest-defective-1 (hard1) protein with breast cancer. *Cancer Invest*. 2009;27(10):978-983. doi:10.3109/07357900902769723.

Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kübler, J., Lozajic, M., … Alva, V. (2018). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of Molecular Biology*. https://doi.org/10.1016/j.jmb.2017.12.007

8 Supplementary

Table 8.1: *In silico* prediction tools applied to predict the possible impact of the NAA10 L11R variant.

NAA10 L11R		
Tool	Prediction	Probability / ΔΔG (kcal/mol)
PolyPhen-2	Probably damaging	0.989
PMut	Disease	0.85
PROVEAN	Deleterious	-7.226
SNP&GO3D	Disease	0.94
SIFT	Affect protein function	0.01
CUPSAT	Favorable, stabilizing	-0.42 kcal/mol
HOPE	Destabilizing	-
DUET	Destabilizing	-1.959 kcal/mol
DynaMut	Destabilizig	-0.825 kcal/mol

Table 8.2: *In silico* prediction tools applied to predict the possible impact of the NAA10 H16P variant

NAA10 H16P		
Tool	Prediction	Probability / ΔΔG (kcal/mol)
PolyPhen-2	Probably damaging	0.837
PMut	Disease	0.70
PROVEAN	Deleterious	-4.743
SNP&GO^{3D}	Disease	0.69
SIFT	Tolerated	0.06
CUPSAT	Unfavorable, destabilizing	1.25 kcal/mol
HOPE	Destabilizing	-
DUET	Destabilizing	-0.62 kcal/mol
DynaMut	Stabilizing	0.261 kcal/mol