# Investigating hydroxylation of the acetyltransferase NAA10 as a potential regulation mechanism

By
Karoline Krogstad


*Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science*

Department of Biological Sciences
Faculty of Mathematics and Natural Sciences
University of Bergen

November 2019

# TAKK TIL

Arbeidet med dette prosjektet har blitt gjennomført ved NAT-gruppen på Institutt for Biovitenskap, Universitetet i Bergen.

Aller først vil jeg gi en helt spesiell takk til mine to fantastiske medveiledere, Nina McTiernan og Rasmus Moen Ree. Tusen takk for all tålmodighet og positivitet dere har vist meg både på lab og under en lang og tung skriveprosess dette året. Jeg er utrolig takknemlig for all den tiden dere har brukt på retting, fiksing, mail og på å besvare de rareste spørsmål, det har betydd veldig mye for meg. Helger eller ukedager, morgen eller kveld – dere har stilt opp til enhver tid hvis jeg har hatt behov for hjelp, og det har jeg satt så utrolig stor pris på. Tusen takk for fantastisk god hjelp dette året! Jeg har lært så mye av dere, og uten dere hadde ikke dette vært mulig.

Tusen takk til Thomas Arnesen, min hovedveileder, som også har vært en fantastisk støtte gjennom dette prosjektet. Du er en utrolig motivator som virkelig vet å hente frem litt pågangsmot der det ikke er så mye igjen. Kontoret ditt har aldri vært langt unna, og veiledningsmøter hos deg har virkelig hjulpet meg til å holde motet oppe selv om ting innimellom ikke helt har gått veien. Takk for at du ga meg muligheten til å få være en del av dette utrolig spennende prosjektet, og for all inspirasjon og motivasjon du har gitt meg på veien. Jeg tar med meg utrolig mye verdifullt fra dette året hos dere!

Resten av NAT-gruppa skal også ha en stor takk for at det er så hyggelig og gøy å få være masterstudent i gruppa. Jeg har nok på en eller annen måte fått hjelp eller råd av hver og en av dere i løpet av dette året, og jeg er utrolig takknemlig for at det aldri har vært vanskelig å spørre om hjelp. Takk til min venninne og labpartner Christine som har sittet i samme båt som meg i både oppturer og nedturer. Dette hadde definitivt ikke vært det samme uten deg, og jeg ønsker deg masse lykke til med oppgaven din og med veien videre! I tillegg ønsker jeg å takke Helix og en stor gjeng med medstudenter for et godt og sosialt miljø som har gjort det så utrolig gøy å være masterstudent.

En stor takk må også rettes til familien min som har stilt opp hele veien og aldri tvilt på meg. Det har vært en uvurderlig støtte å ha dere i ryggen, selv om jeg vet at dette (hvertfall frem til nå) ikke akkurat har vært deres favorittfagfelt. Sist men ikke minst tusen takk til Andreas som har holdt ut med meg i dette kaoset, og som har fått meg til å glemme stress og bekymringer på litt mørke dager. Jeg er heldig som har deg!

**Bergen, november 2019**

**Karoline Krogstad**

# TABLE OF CONTENTS

# SELECTED ABBREVIATIONS

| | |
|---|---|
| ABC | Ammonium bicarbonate |
| Ac-CoA | Acetyl coenzyme A |
| ACN | Acetonitrile |
| AGC | Automatic gain control |
| DDA | Data-dependent acquisition |
| DTT | Dithiothreitol |
| *E. coli* | *Escherichia coli* |
| FA | Formic acid |
| FASP | Filter-aided sample preparation |
| FBS | Fetal bovine serum |
| FDR | False discovery rate |
| FIH | Factor inhibiting HIF1-α |
| GNAT | GCN5-related N-acetyltransferase |
| HAT | histone acetyltransferase |
| HIF1-α | Hypoxia-inducible factor 1α |
| HRP | Horseradish peroxidase |
| HYPK | Huntingtin-interacting protein K |
| IAA | Iodoacetamide |
| iMet | Initiator methionine |
| IP | Immunoprecipitation |
| KAT | Lysine acetyltransferase |
| kDa | Kilodalton |
| KDAC | Lysine deacetylase |
| LC-MS/MS | Liquid chromatography-tandem mass spectrometry |
| LFQ | Label-free quantification |
| Lys-Ac | Lysine acetylation |
| NAT | N-terminal acetyltransferase |
| NDAC | N-terminal deacetylase |
| Nt-Ac | N-terminal acetylation |
| ON | Overnight |
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phosphate-buffered saline |
| PBST | PBS-Tween |
| PTM | Post-translational modification |
| RT | Room temperature |
| SDS | Sodium dodecyl sulfate |
| WB | Western blotting |

# 1 ABSTRACT

Human N-terminal acetyltransferases (NATs) acetylate the N-termini of around 80% of the human proteome. The NATs identified in humans are termed NatA-F and NatH and display distinct substrate specificities (Aksnes, Ree and Arnesen, 2019). NatA is regarded the major NAT and is composed of the catalytic subunit NAA10 and the auxiliary subunit NAA15 (Arnesen *et al*., 2005a). Impaired NAA15 and NAA10 function and overexpression have been associated with intellectual and physical disabilities, as well as tumor development respectively (Myklebust, Støve and Arnesen, 2015; Kalvik and Arnesen 2013).

Studies have found that NAA10 may display functions that extends beyond N-terminal acetylation (Nt-Ac), and it has been suggested that it is also active as a lysine acetyltransferase (KAT) (Jeong *et al*., 2002). Research groups claim to have found NAA10-mediated lysine acetylation (Lys-Ac) of a variety of substrates (Lim *et al*., 2006; Shin *et al*., 2009; Wang *et al*., 2012; Shin *et al*., 2014; Yoon *et al*., 2014; Seo *et al*., 2016; Lee *et al*., 2017; Qian *et al*., 2017; Lu Vo *et al*., 2017). However, this is under debate, and the theory is challenged both by conflicting findings (Arnesen *et al*., 2005b; Murray-Rust *et al*., 2006) as well as structural features distinguishing the substrate specificity of the NATs from the specificity of the KATs (Liszczak *et al*., 2013). A recent attempt to explain this disagreement involves hydroxylation on Trp38 of monomeric NAA10 by factor inhibiting HIF-1α (FIH). This is thought to induce a conformational change in the NAA10 structure enabling it to acetylate Lys532 of HIF-1α (Kang *et al*., 2018).

In this study, the hydroxylation status of monomeric and NAA15-complexed NAA10 was investigated by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Hydroxylated Asp, Asn and Trp residues as well as oxidized Met residues of NAA10-derived peptides were identified, but the peptide of Trp38 was found to be unmodified. Supported by previous studies we concluded that a part of these modifications are likely artefacts of MS sample preparation (Froelich and Reid, 2008; Klont *et al*., 2018). Further study of the NAA10 and FIH relationship by Western blotting (WB) revealed a possible physical interaction, as well as increased expression levels of Xpress-NAA10 and FIH-V5 in presence of each other. However, i*n vitro* acetylation assays suggested that FIH has no clear impact on complexed NAA10 NAT activity, the amount of NatA in cells nor the NatA complex formation. In sum, our study does not lend support to the hypothesis that Trp38 hydroxylation of NAA10 is a key switch between its NAT and potential KAT activities.

# 2 INTRODUCTION

## 2.1 PROTEIN ACETYLATION

The human genome is comprised of about 20 000 protein-coding genes that give rise to the vast number of potential protein variants making up the complex human proteome. The diversity of human proteins is mainly caused by processes like alternative splicing, single amino acid polymorphisms and post-translational modifications (PTMs) such as methylation, phosphorylation, ubiquitination, glycosylation and acetylation (Verdin and Ott, 2015; Ponomarkenko *et al*., 2016; Boutureira and Bernardes, 2015). Among these, acetylation is known as one of the major PTMs, having essential roles in biological processes involved in metabolism, gene expression, and protein stability (Verdin and Ott, 2015).

Protein acetylation can be divided into N-terminal acetylation (Nt-Ac) and lysine acetylation (Lys-Ac). Nt-Ac has various molecular effects depending on the modified protein and may be involved in protein-protein interactions and correct subcellular localization of proteins (Aksnes, Ree and Arnesen, 2019). The acetyl-group may also act as a proteasomal degradation signal, thereby controlling the lifetime and stoichiometry of proteins (Aksnes, Ree and Arnesen, 2019). Lysine acetylation of proteins influence a range of cellular signaling pathways, such as metabolism, stress responses, apoptosis and membrane trafficking (Drazic *et al*., 2016). It also plays a crucial role in epigenetics and transcriptional regulation through acetylation of histones, where it affects the folding and interaction between the DNA and the histone (Drazic *et al*., 2016). Besides enzymatic acetylation, some acetylation events also occur chemically through direct interaction between Acetyl coenzyme A (Ac-CoA) and the receiving protein. This is typical for conditions where the pH and the concentration of Ac-CoA is high, like in mitochondria (Baeza *et al*., 2015; Wagner *et al*., 2013)

### 2.1.1 N-terminal acetylation

Nt-Ac refers to the covalent attachment of an acetyl group on the positively charged free α-amino group at the N-terminal end of polypeptides (Figure 2.1). With 80% of all human proteins being acetylated at their N-termini, protein Nt-Ac is one of the most common protein modifications in eukaryotic cells (Aksnes *et al*., 2016; Aksnes, Ree and Arnesen, 2019; Arnesen *et al*., 2009). The transfer is catalyzed by a number of enzyme complexes called N-terminal acetyltransferases (NATs), and Ac-CoA functions as an acetyl group donor (Aksnes *et al*., 2016). The NAT machinery is classically thought to act co-translationally as proteins emerge from the ribosome, but it has in recent years been shown to also act post-translationally for several proteins (Aksnes, Ree and Arnesen, 2019). So far, the reaction is considered irreversible as no N-terminal deacetylase (NDAC) has been discovered yet (Ree, Varland and Arnesen, 2018).

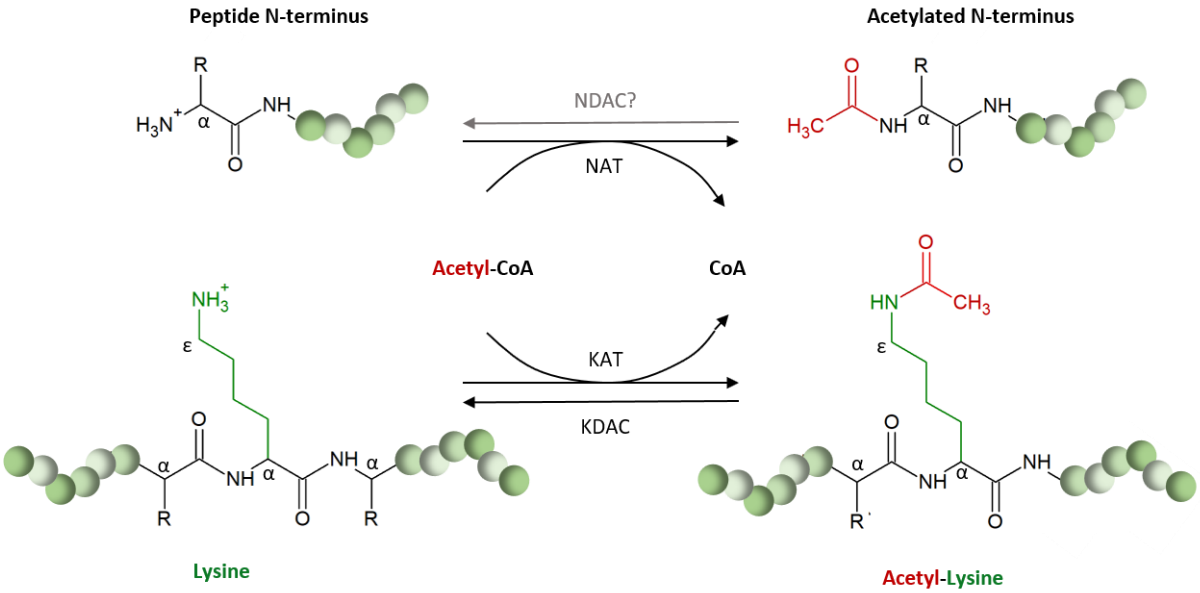


**Figure 2.1: Protein Nt-Acetylation and Lys-Acetylation.** NATs target the α-amino group of the N-termini of their protein substrates in order to catalyze the transfer of an acetyl group (red) from Ac-CoA. Lysine acetyltransferases (KATs) target the ε-amino group of an internal lysine residue of their substrates. So far Nt-Ac is considered to be irreversible as no NDAC has been identified, whereas lysine acetylation can be reversed by lysine deacetylases (KDACs). (Adapted from Aksnes, Ree and Arnesen, 2019)

Neutralization of the positively charged N-terminus of proteins changes their electrostatic properties (Ree, Varland and Arnesen, 2018) and depending on the receiving proteins the effects of Nt-Ac can be quite diverse (Aksnes, Ree and Arnesen, 2019). In some cases, the Nt-acetyl group has been found to regulate protein lifetime and stability by acting as a proteasomal degradation signal recognized by E3 ubiquitin ligases (Hwang, Shemorry and Varshavsky, 2010; Shemorry *et al.*, 2013). It has also been found to have an impact on protein folding and aggregation (Bartels *et al.*, 2011; Holmes *et al.*, 2014; Trexler *et al.*, 2012), complex formation (Monda *et al.*, 2013) and subcellular targeting (Behnia *et al.* 2004, Forte *et al.*, 2011, Setty *et al.*, 2004).

### 2.1.2 Lysine acetylation

Another form of protein acetylation is the acetylation of the ε-amino group of lysine residues (Figure 2.1) which is catalyzed by enzymes termed lysine acetyl transferases (KATs). Lys-Ac was first discovered in histones regulating gene transcription (Allfrey and Mirsky, 1964), and the enzymes responsible for catalyzing this modification were termed histone acetyltransferases (HATs). In 1997 it was discovered that HATs could also acetylate non-histone proteins (Imhof *et al.*, 1997), leading to the use of the term KATs instead of HATs. Lys-Ac is a reversible modification in contrast to Nt-Ac, and the reverse reaction is catalyzed by lysine deacetylases (KDACs) (Glozak *et al.*, 2005). The balance between acetylation and deacetylation of lysines is tightly regulated, as it functions as a switch for cellular signaling and metabolic pathways (Drazic *et al.*, 2016).

## 2.2 N-TERMINAL ACETYLTRANSFERASES

### 2.2.1 The family of NATs

The family of NATs are highly conserved through evolution, and Nt-Ac events occur in all kingdoms of life. Currently there are three known NATs existing in prokaryotes (RimI, RimJ and RimL) (Yoshikawa *et al*., 1987; Tanka *et al*., 1989). In eukaryotes, eight NATs termed NatA- NatG have been identified, all displaying distinctive features and substrate specificities (Figure 2.2). The ribosome-associated NatA-E can be found in the cytosol of yeast, plants and mammalian cells performing co-translational Nt-Ac of peptide substrates (Aksnes, Ree and Arnesen, 2019). NatF is integrated to the cytosolic surface of the Golgi membrane of multicellular eukaryotes such as animals and plants where it Nt-acetylates its substrates post-translationally (Aksnes *et al*., 2015). NatG, is localized in the chloroplasts of plants (Dinh *et al*., 2015). NatH is the most recently discovered NAT family member performing post-translational acetylation of the acidic N-terminal end of actin in the cytosol (Drazic *et al*., 2018).

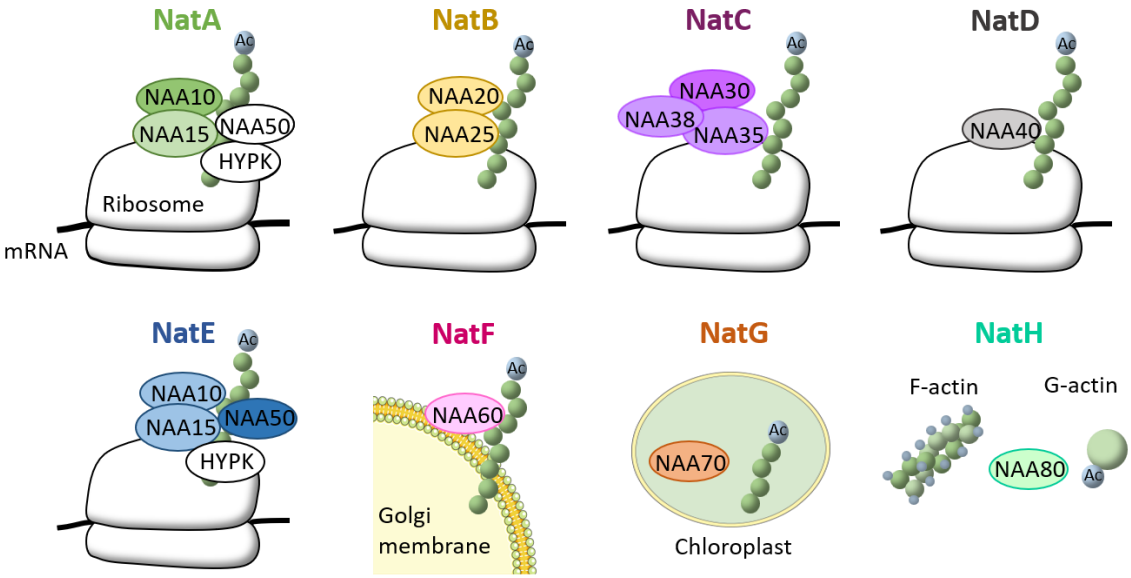


**Figure 2.2: Overview of the composition of the eight NATs known to date termed NatA-H.** The catalytic subunits (NAA10, NAA20, NAA30 and NAA50) typically associates with one or two auxiliary subunits (NAA15, NAA25, NAA35, NAA38, and NAA50) functioning as ribosomal anchors and/or modulators of substrate specificity. The catalytic subunits NAA40, NAA60, NAA70 and NAA80 is so far known to perform acetylation activities without complex formation. NatA-NatE and NatH all reside in the cytosol of cells, whereas NatF and NatG are organellar NATs integrated in the Golgi membrane and localized in plant chloroplasts respectively. (Adapted from Aksnes *et al*., 2016 and modified from Smart by Servial Medical Art)

### 2.2.2 Eukaryotic NATs

Most Nt-Ac in eukaryotes is performed by NatA-E, and higher eukaryotes also express NatF (Aksnes, Ree and Arnesen, 2019). The most recently described NAT, NatH, acts specifically on actin (Drazic *et al*., 2018). For most NATs, their function is dependent upon binding and complex formation between a catalytic subunit and auxiliary subunits with various functions (Varland, Osberg and Arnesen, 2015), except for NatD, NatF and NatH where only one catalytic subunit has been identified so far (Figure 2.2) (Hole *et al*., 2011; Van Damme *et al*., 2011b; Drazic *et al*., 2018). The identity of the two first amino acids located at the N-terminus mostly determines whether a protein becomes acetylated or not and by which NAT (Aksnes, Ree and Arnesen, 2019). The substrate specificities of the eukaryotic NAT enzymes are summarized in Table 2.1.

**Table 2.1:** Overview of the different NAT compositions and substrate specificities

| NAT | Subunits | Substrate specificity |
|---|---|---|
| **NatA** | NAA10, NAA15, HYPK | Ala-, Ser-, Thr-, Gly-, Val-, Cys- |
| **NatB** | NAA20, NAA25 | Met-(Asn/Asp/Gln/Glu) |
| **NatC** | NAA30, NAA35, NAA38 | Met-(Ile/Leu/Lys/Met/Phe/Tyr/Trp) |
| **NatD** | NAA40 | Ser-Gly (H2A and H4) |
| **NatE** | NAA50, NAA10, NAA15 | Met-(Ala/Ile/Leu/Lys/Met/Phe/Ser/Thr/Tyr/Val) |
| **NatF** | NAA60 | Met-(Ala/Gln/Gly/Ile/Leu/Lys/Met/Phe/Ser/Thr/Tyr/Val) |
| **NatH** | NAA80 | Asp/Glu-Asp/Glu-Asp/Glu (Actins) |

### 2.2.3 GNAT domain

The NATs as well as many KATs are all members of the GCN5 (general control non-derepressible 5)-related acetyltransferase (GNAT) superfamily and share a highly conserved structure despite a low overall sequence similarity (Aksnes *et al*., 2016; Drazic *et al*., 2018). The conserved core GNAT fold comprises seven β-strands and four α-helices (Figure 2.3), but some NATs may also have varying secondary structure elements on the C- or N-termini (Vetting *et al*., 2005). C-terminal elements are the most usual, but NAA40 and NAA60 are examples of NATs with varying secondary structure elements on the N-terminus (Magin, Liszczak and Marmorstein, 2015; Støve *et al*., 2016). The consensus Ac-CoA binding motif Q/RxxGxG/A is also a part of the core fold of all members of the GNAT superfamily which is located in the β4-α3 loop (Vetting *et al*., 2005). Most NATs are structurally distinguished from the KATs of the GNAT superfamily by having loops mediating specific substrate recognition. The loops constrict the peptide binding sites so that internal lysine substrates cannot be inserted, thereby restricting the NATs to N-terminal peptide acetylation (Magin, March and Marmorstein, 2016). An extended α1-α2 loop of NAA40 (Magin, Liszczak and Marmorstein *et al*., 2015) and the β6-β7 loop of NAA10, NAA20 from *C. albicans*, NAA50 and NAA60 (Liszczak *et al*., 2013; Hong *et al*., 2017; Liszczak *et al*., 2011; Støve *et al*., 2016) are responsible for this function. In the case of NAA80 both the α1-α2 and β6-β7 regions differ compared to other NATs in having more open conformations in the structure, resulting in a wider binding groove (Goris *et al*., 2018).

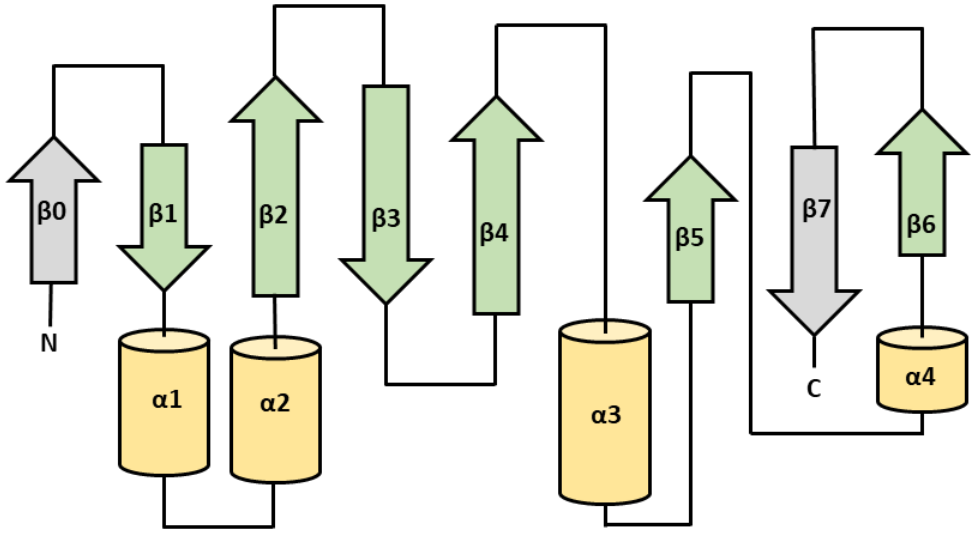


**Figure 2.3: Topology of the GNAT fold.** The conserved core GNAT fold is built up of seven β-strands (arrows) and four α-helices (cylinders), but some NATs also have varying secondary structure elements on the N- and/or C-terminus (grey). The highly conserved motif for Ac-CoA binding is located in the β4-α3 loop (Figure adapted from Vetting *et al*., 2005).

## 2.3 NatA

### 2.3.1 Specificity and composition

NatA, NatB and NatC are highly conserved from yeast to humans, where NatA is responsible for the majority of eukaryotic Nt-Ac. NatA is specific towards small amino acids such as Ser, Ala, Thr, Gly, Val and Cys at N-termini where the initiator methionine (iMet) has been cleaved off (Arnesen *et al*., 2009). In terms of number of substrates NatA is the major eukaryotic NAT (Starheim, Gevaert and Arnesen, 2012), acetylating approximately 40% of the proteome (Aksnes, Ree and Arnesen, 2019). NatA is composed of the catalytic subunit NAA10 and the ribosome binding subunit NAA15 (Liszczak *et al*., 2013), and is localized in the cytosol where it performs co-translational Nt-Ac in association with ribosomes (Arnesen *et al*., 2005a). Although most of the cellular NAA10 is complexed with NAA15 (Aksnes, Ree and Arnesen, 2019), it also exists in a monomeric state in both the cytoplasm and the nucleus. Upon NatA complex formation a conformational change in the α1-α2 loop of NAA10 causes an alteration in substrate specificity (Liszczak *et al*., 2013), which could indicate that monomeric NAA10 has an independent role in post-translational acetylation of proteins (Liszczak *et al*., 2013; Van Damme *et al*., 2011a). *In vitro* studies have shown that monomeric NAA10 targets and acetylates acidic N-termini rather than the known NatA substrates (Van Damme *et al*., 2011a), but since NAA80 has been demonstrated to acetylate the acidic N-termini of actin the natural substrates of monomeric NAA10 still remain unknown (Drazic *et al*., 2018).

Two additional auxiliary components of the NatA complex are NAA50 and the Huntingtin-interacting protein K (HYPK). NAA50 is also defined as NatE (Evjenth *et al*., 2009). HYPK is stabilized by association with the NatA complex, and it has been found that this interaction is also required for optimal NatA activity towards some substrates (Arnesen *et al*., 2010). In contrast, a recent study found HYPK to have an inhibitory effect on NatA activity *in vitro* and it was speculated that HYPK might have a proofreading mechanism *in vivo* (Gottlieb and Marmorstein, 2018).

### 2.3.2 NatA structure

The structure of the human NatA/HYPK complex was solved by Gottlieb and Marmorstein in 2018 (Figure 2.4), and it displays a high degree of conservation to the SpNatA structure solved by Liszczak *et al*. in 2013 (Gottlieb and Marmorstein, 2018). Human NAA10 is composed of 235 amino acids forming a secondary structure of five α-helices and seven β-strands. Human NAA15 consists of 866 amino acids making up its 45 α-helices that form 13 conserved helical bundle tetratricpeptide repeat (TPR) motifs. Such segments typically promote protein-protein interactions (Blatch and Lassle, 1999, Liszczak *et al*., 2013), and some of them are involved in interactions with NAA10 whereas some are suspected to be important for binding other interaction partners such as the ribosome, NAA50 and HYPK (Liszczak *et al*., 2013). HYPK is composed of five α-helices interacting with NatA by forming a clamp around the NAA15 auxiliary subunit (Gottlieb and Marmorstein, 2018) (Panel A). The NAA15-helices form a ring-like tertiary structure wrapping completely around NAA10. NAA10 adopts the conserved GNAT fold with the N- and C-terminal regions flanking the central Ac-CoA binding domain (Liszczak *et al*., 2013), and features an additional C-terminal α-helix (α5) as shown in Panel B (Gottlieb and Marmorstein, 2018). By comparing complexed and uncomplexed structures of NAA10, Liszczak *et al*. found that binding to NAA15 induces an allosteric change in the active site of NAA10 changing its specificity towards NatA substrates.

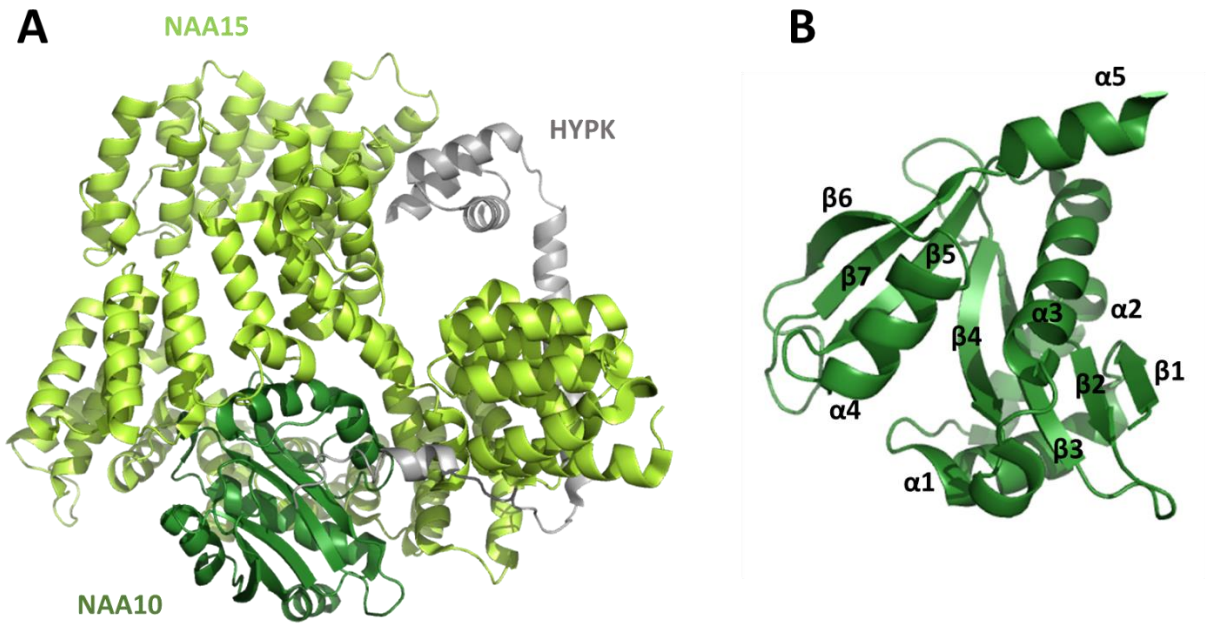


**Figure 2.4: Crystal structure of the human NatA/HYPK complex (PDB entry: 6C95).** α-helices are shown as cylinders and β-strands as arrows. (A) The crystal structure of the human NatA complex bound to HYPK was solved by Gottlieb and Marmorstein in 2018. NAA15 (light green) consists of 45 α-helices forming a ring-like structure around NAA10 (dark green). HYPK (grey) interacts with the NatA complex by forming a clamp around NAA15 (Gottlieb and Marmorstein., 2018). (B) NAA10 from the solved NatA crystal structure adopts the conserved GNAT fold with seven β-strands (β1-β7) and four α-helices and has one additional C-terminal α-helix (α5).

### 2.3.3 NAA10 as a lysine acetyltransferase (KAT)

Monomeric NAA10 has also been reported to acetylate lysine residues of a variety of substrates presented in Table 2.2. Jeong *et al*. found that NAA10 directly binds to and acetylates HIF-1α to regulate its stability and suggest that NAA10 functions as a negative regulator of the HIF-1α protein (Jeong *et al*., 2002). This theory has been challenged by several conflicting findings where the acetylation of HIF-1α has not been reproducible *in vitro* (Arnesen *et al*., 2005b; Murray-Rust *et al*., 2006). Shin *et al*. and Yoon *et al*. also reported that NAA10 mediates acetylation and thereby also regulates MLCK (Shin *et al*., 2009), represses the enzymatic function of MSRA (Shin *et al*., 2014) and inhibits Runx2 (Yoon *et al*., 2014). Magin *et al*. reconstituted and assessed Lys-Ac events *in vitro* by performing KAT assays for these reported substrates, but no difference in Lys-Ac was seen in presence and absence of NAA10 (Magin, March and Marmorstein, 2016). They therefore suggest that acetylation of the reported substrates occur chemically and not enzymatically.

**Table 2.2:** Overview of reported NAA10 KAT substrates

| Substrate | Reference |
|---|---|
| Hypoxia inducible factor 1α (HIF-1α) | Jeong *et al*., 2002 |
| β-catenin | Lim *et al*., 2006 |
| Myosin light chain kinase (MLCK) | Shin *et al*., 2009 |
| Androgen receptor | Wang *et al*., 2012 |
| Methionine sulfoxide reductase A (MSRA) | Shin *et al*., 2014 |
| Runt-related transcription factor 2 (Runx2) | Yoon *et al*., 2014 |
| Heat shock protein 70 (Hsp70) | Seo *et al*., 2016 |
| SAM domain and HD domain containing protein 1 (SAMDH1) | Lee *et al*., 2017 |
| Phosphoglycerate kinase 1 (PGK1) | Qian *et al*., 2017 |
| Aurora kinase A | Lu Vo *et al*., 2017 |

From a structural point of view, the KAT activity of NAA10 is quite controversial as the NAT structures determined to date, including the structure of NAA10, contain a loop blocking lysine residues from entering the substrate binding site of the enzymes as previously described. In the case of NAA10, the substrate gate located in the groove between the α1-α2 and β6-β7 loops is blocked by the β6-β7 loop, making the active site unable to accommodate a lysine substrate (Figure 2.5) (Gottlieb and Marmorstein, 2018 and Liszczak *et al*., 2013).

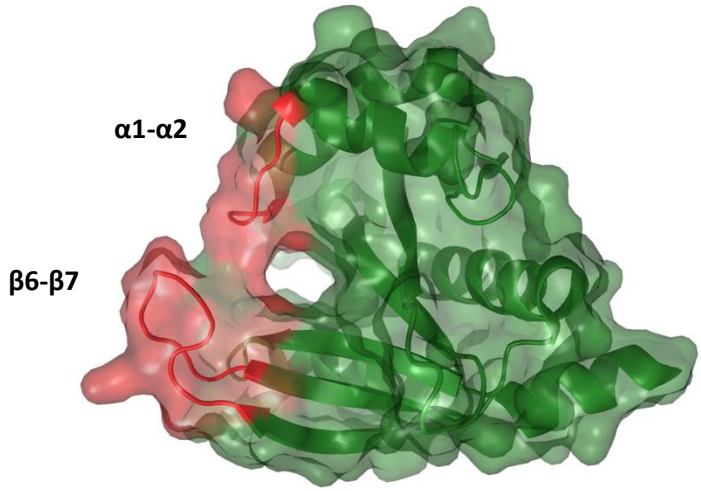


**Figure 2.5: The α1-α2 and β6-β7 loops of NAA10 partly blocks its substrate binding site (PDB entry: 6C95).** The studies reporting NAA10 KAT activity have been challenged by contradicting findings and structural restrictions. A structural feature separating the NATs from the KATs is secondary structure elements or loops partly blocking the substrate gate, making it too narrow to accommodate lysine residues. The active site of NAA10 is blocked by the α1-α2 and the β6-β7 loops.

### 2.3.4 NatA and NAA10 in disease and development

The importance of a well-functioning NAT machinery has been established through discovery of NAT mutations in various pathological conditions. The first human genetic disorder caused by a NAA10 mutation (Ser37Pro) was reported by Rope *et al*. and was termed lethal X-linked Oden syndrome as NAA10 is located on the X-chromosome (Rope *et al*., 2011). Several additional NAA10 missense mutations causing genetic disorders were reported after this finding (Casey *et al*., 2015; Popp *et al*., 2015; Saunier *et al*., 2016), sharing the phenotypes of intellectual disabilities, developmental delay as well as growth failure including cardiac and skeletal abnormalities (Myklebust, Støve and Arnesen, 2015). Furthermore, reduction of catalytic activity has been reported for the NAA10 variants, suggesting that NAA10 possesses an important role in the development of tissues and organs (Casey *et al*., 2015). Studies have shown that loss of NAA10 function is lethal in *Drosophila* (Wang *et al*., 2010), *Trypanosoma brucei* (Ingram *et al*., 2000) and *Caenorhabditis elegans (*Sönnichsen *et al*., 2005), and that it causes severe phenotypes also in *Danio rerio* (Ree *et al*., 2015). The effects of the Ogden syndrome strongly imply that loss of NAA10 function and impaired Nt-Ac is lethal also in humans (Rope *et al*., 2011; Myklebust, Støve and Arnesen, 2015). NAA10-mediated Lys-Ac has also been found to have an effect on oxidative stress. Opposing responses have been seen for Lys-Ac of the reported KAT substrates Hsp70 and MSRA, where the former has been found to protect the cells against stress conditions and damage by reactive oxygen species (Seo *et al*., 2016), whereas the latter was found to worsen the cellular damage (Shin *et al*., 2014). These results indicate that NAA10 also has an important role in cellular redox balance, but its functions seem to be dependent on the substrate (Vo *et al*., 2018).

Several studies have also linked NAA10 and NAA15 to tumor development and cell survival, and overexpression of the proteins in certain cancers seem to be correlated with tumor aggressiveness and low survival rate (Aksnes, Ree and Arnesen, 2019). Elevated NAA15 mRNA levels was first discovered in papillary thyroid carcinoma (Fluge *et al*., 2002), and subsequently in gastric cancer (Line *et al*., 2002) and neuroblastomas (Martin *et al*., 2007). Overexpression of NAA10 has also been reported in several types of cancer, such as colorectal cancer (Ren *et al*., 2008; Yu *et al*., 2009a), lung cancer (Lee *et al*., 2010), breast cancer (Yu *et al*., 2009a; Yu *et al*., 2009b) and hepatocellular carcinoma (Midorikawa *et al*., 2002). Several of the reported NAA10 KAT substrates have also been linked to cancer. For SAMHD1 and Aurora kinase A NAA10-dependent acetylation of specific lysine residues increases the activity and thereby promotes cancer cell proliferation (Lee *et al*., 2017; Lu Vo *et al*., 2017). The reported acetylation of PGK1 and β-catenin was also found promote brain tumor formation and lung cancer cell proliferation respectively (Qian *et al*., 2017; Lim *et al*, 2006). From these studies it has become evident that NAA10 displays important functions in normal development as well as in promoting cell proliferation and survival and regulation of cell metabolism (Drazic *et al*., 2016). Depending on tissue and tumor type, studies have found that NAA10 can act as both an oncoprotein and a tumor suppressor (Kalvik and Arnesen, 2013; Midorikawa *et al*., 2002), and questions regarding its complicated function in disease and tumor development remain to be answered. Since acetylation of several NAA10 target proteins have been reported to promote tumorigenesis, development of NAT-inhibitors has been proposed as a potentially effective method to target various types of cancer (Kalvik and Arnesen, 2013; Seo *et al*., 2010).

### 2.4 NAA10 and FIH in HIF-1α signaling

One of the reported lysine substrates of NAA10, HIF-1α, is one of the two subunits of the heterodimeric transcription factor HIF1 responsible for mediating the cellular response to reduced oxygen levels (hypoxia) in many mammals. While HIF-1β is constitutively expressed in the nucleus, HIF-1α levels are regulated by oxygen levels. Under normoxia, HIF-1α is rapidly degraded through the von Hippel-Lindau tumor suppressor gene product (pVHL)-mediated ubiquitin-proteasome pathway (Salceda and Caro, 1997), whereas under hypoxic conditions the HIF-1α level increases (Wang *et al*., 1995). The regulation of HIF-1α is so far known to involve multiple PTMs such as ubiquitinylation, hydroxylation and phosphorylation (Jeong *et al*., 2002). The rapid proteasomal degradation of HIF-1α under normoxic conditions is due to hydroxylation of two Pro residues within the oxygen-dependent degradation domain, which promotes the interaction with the pVHL ubiquitin ligase complex. The prolyl hydroxylase action is blocked under hypoxic conditions, and the stability of HIF-1α increases (Ivan *et al*., 2001; Jaakkola *et al*., 2001; Yu *et al*., 2001). Another reported HIF-1α regulator is the asparagine hydroxylase Factor Inhibiting HIF-1α (FIH). Through hydroxylation of Asn803 located in the C-terminal transactivation domain of HIF-1α, FIH is found to suppress its transcriptional activity under normoxic conditions. This is done by blocking further association with the transcriptional co-activators CBP/p300 (Lando *et al*., 2002).

Studies also show evidence of acetylation playing a role in HIF-1α regulation, as interaction was found between HIF-1α and NAA10. *In vitro* acetylation of Lys532 of HIF-1α was detected after incubation in presence of recombinant NAA10 (Jeong *et al*., 2002), a residue located on the oxygen-dependent degradation domain implicated in the regulation of HIF-1α stability (Bilton *et al*., 2005). Jeong *et al*. further reported the acetylation to have a destabilizing effect on HIF-1α by increasing its interaction with pVHL. Under hypoxia, Jeong *et al*. also detected less acetylated HIF-1α compared to normoxia which was consistent with the findings reported in another study (Fisher *et al*., 2005). The activity of acetyltransferases is not known to be dependent on the oxygen concentration (Arnesen *et al*., 2005b), and Jeong *et al*. claim that this can be explained by downregulation of NAA10 protein and mRNA under hypoxic conditions. These findings have been challenged by studies showing opposing results. Replication efforts failed to find an impact on HIF-1α stability by neither NAA10 overexpression nor silencing (Arnesen *et al*., 2005b, Bilton *et al*., 2005), or that HIF-1α is acetylated by NAA10 (Arnesen *et al*., 2005b; Murray-Rust *et al*., 2006). Bilton *et al*. also showed that neither NAA10 protein levels nor mRNA levels are regulated by hypoxia, which is supported by findings reported by Arnesen *et al*. Seo *et al*. have tried to explain these opposing results with autoacetylation, which they claim is required for NAA10-mediated KAT activity. This involve K29, K136 and K198, where K136 is claimed to be the most crucial acetylation site (Seo *et al*., 2010). However, this is inconsistent with a previous study by Murray-Rust *et al*. where liquid chromatography-mass spectrometry (LC-MS/MS) analyses provided evidence that none of the 16 lysine residues of NAA10 were found to be autoacetylated (Murray-Rust *et al*., 2006).

A recent study has reported new evidence that HIF-1α is acetylated by NAA10, proposing a mechanism of how NAA10 may overcome the blockage of the substrate gate in order to act as both a NAT and a KAT in cells (Figure 2.6) (Kang *et al*., 2018). In this study, Kang *et al*. found that the asparagine hydroxylase FIH binds to NAA10 and utilizes oxygen under normoxic conditions to hydroxylate Trp38 located in the α1-α2 loop of NAA10. They report an abundancy of 26% for the hydroxylated peptide (Panel A) which was also found to be hydroxylated at Asp47 and Asn49. According to Kang *et al*. the hydroxylation occurs on C-2 of the indole ring of Trp38, and they further tested the possibility that the substrate gate changes shape upon hydroxylation. Based on a molecular dynamics model this OH-group is thought to form an additional hydrogen bond with a carbonyl oxygen of Ala67 located at β3. They suggest that the hydrogen bonding is responsible for a conformational change where the α1-α2 loop bends backwards and widens the substrate gate. The molecular dynamics model indicates a substantially altered conformation for monomeric NAA10 and a more rigid and less affected conformation for NAA15-complexed NAA10. Kang *et al*. claim that monomeric and hydroxylated NAA10 is enabled to display KAT activity towards HIF-1α through this mechanism. In association with PHD-mediated hydroxylation of Pro402 and Pro564, acetylation of HIF-1α is further thought to recruit the pVHL ubiquitin ligase complex. HIF-1α is further degraded through the ubiquitin-proteasome system (Panel B). FIH also directly inhibits the transcriptional activity of HIF-1α by hydroxylating Asn803, thereby blocking the transcriptional co-activator protein p300 from binding (Kang *et al*., 2018). Under hypoxic conditions FIH cannot hydroxylate NAA10 or HIF-1α due to the lack of oxygen. In this case HIF-1α becomes stabilized and activated, enabling downstream expression of target genes.

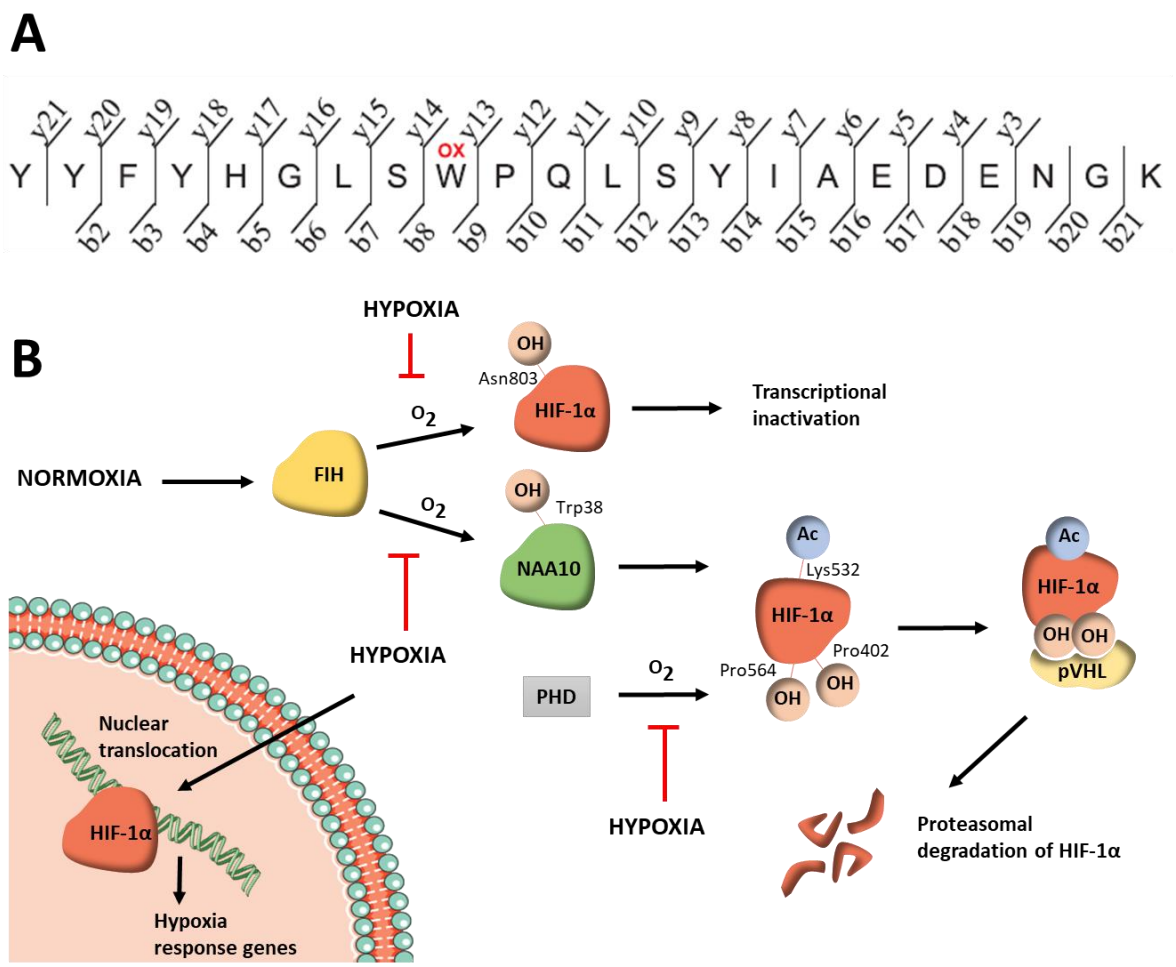


**Figure 2.6: NAA10 is claimed to be hydroxylated at Trp38 under normoxia by FIH.** (A) The peptide sequence with the hydroxylated Trp38 found by Kang *et al*. (2018). The peptide is positioned at 30-51 in the NAA10 protein sequence and was also found to be hydroxylated at Asp47 and Asn49. (B) The hydroxylation is thought to trigger a conformational change in the NAA10 structure, enabling it to accommodate a lysine residue and display KAT-activity towards HIF-1α. This contributes to the proteasomal degradation of the HIF-1α protein in addition to hydroxylation of Pro402 and Pro564 performed by HIF prolyl hydroxylase isozymes (PHD1-3), resulting in the short half-life of HIF-1α under normoxic conditions. FIH also directly hydroxylates HIF-1α, thereby inhibiting its transcriptional activity. Under hypoxic conditions, no hydroxylation occurs and HIF-1α is stabilized and free to translocate to the nucleus (Adapted from Aksnes, Ree and Arnesen, 2019 and modified from Smart by Servial Medical Art)

## 2.4 AIM OF STUDY

NAA10 was originally known as a eukaryotic NAT catalyzing the acetylation of N-terminal amino acids of newly synthesized protein substrates. Since the discovery of NAA10 KAT activity 15 years ago, several studies have reported NAA10-catalyzed acetylation of internal lysine residues of a variety of cellular proteins. It is debated whether NAA10 displays KAT activity or not, as structural and biochemical evidence speaks against it. Through its NAT and KAT activity, NAA10 is reported to play a physiological role in human disease and cancer even though its role in tumor development is far from understood. Increased NAA10 expression has in some cases been correlated with bad prognosis and increased tumor aggressiveness, but has in other cases had opposite effects in certain types of cancer (Zeng *et al*., 2014). This reflects the complexity of the signaling pathways through which NAA10 acts in cancer cells, and further research is needed to clarify the role of NAA10 in different contexts.

Autoacetylation of NAA10 and recently also specific hydroxylation of Trp38 of NAA10 by FIH have been proposed as ways to resolve these contradictions. Seo *et al*. claims that autoacetylation of K136 is crucial for the function of NAA10 as a KAT, as this is acting as a key regulatory switch (Seo *et al*., 2010). Kang *et al*. proposed a mechanism where the formation of a hydrogen bond between the hydroxylated tryptophan located in the α1-α2 loop and a carbonyl oxygen of Ala67 located in β3 induces a conformational change in the NAA10 structure. The hydrogen bonding is thought to bend the α1-α2 loop backwards and open the active site of NAA10, enabling it to acetylate internal lysine residues of substrate proteins. This is a potentially important and interesting finding in the process of better understanding the roles of NAA10 and its different modes of action. Thus, it was of great interest to further investigate this theory and the relationship between NAA10 and FIH. The aims of this thesis can be subdivided into four main questions:

### Is cellular NAA10 Trp38-hydroxylated?

Kang *et al*. recently proposed a mechanism of how NAA10 manages to overcome the structural restrictions in order to display KAT activity. They suggest that FIH contributes to this activity by hydroxylating NAA10 at Trp38 under normoxic conditions with an abundancy of 26%. This is thought to induce a conformational change in the loop blocking its substrate gate, enabling NAA10 to acetylate the Lys532 residue of HIF-1α. One aim has been to assess whether this hydroxylation can be detected with LC-MS/MS in different cell types.

### Does NAA10 physically interact with FIH?

Kang *et al*. found that both endogenous and ectopically expressed NAA10 and FIH interact with each other, and that they predominantly are localized in the cytoplasm. It was assessed whether NAA10 and FIH physically interact in order to further investigate the relationship between the two proteins by performing immunoprecipitation (IP) and Western blotting (WB) with the respective antibodies.

**Does the NAA10 expression level differ in presence and absence of FIH?**

As FIH is suggested to be critical for NAA10 KAT activity it was of interest to study the cellular expression of NAA10 under normal expression, overexpression and absence of FIH. A part of this thesis has therefore been focused on investigating protein expression with WB both for endogenous and ectopically expressed proteins. In order to do this, HEK293 cells were transfected with Xpress-NAA10 and/or FIH-V5 and the cell lysates were subject to IP, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and WB for further analysis.

**Does the activity of cellular NAA10 differ in presence and absence of FIH?**

FIH is claimed to be a requirement for NAA10 KAT activity by hydroxylating Trp38 of monomeric NAA10 under normoxic conditions. It has therefore been investigated whether FIH KO has an impact on the activity of both monomeric and complexed NAA10 towards a variety of substrates, and whether NAA10 displays KAT activity towards histone 4. This was done by performing $^{14}$C-Ac-CoA-based *in vitro* acetylation assay with monomeric and complexed NAA10 from HAP1 WT and FIH KO cells.

# 3 MATERIALS

## 3.1 SOLUTIONS, MEDIA AND BUFFERS

### 3.1.1 Buffers and solutions for bacterial work

**LB-media**
10 g/l Trypton
5 g/l Yeast extract
5 g/l sodium chloride
100 µg/ml ampicillin

**LB-Agar**
13.5 g/l granulated agar
10 g/l Trypton
5 g/l Yeast extract
5 g/l Sodium chloride
100 µg/ml ampicillin

**S.O.C. medium (NEB)**
2% (w/v) peptone
0.5% (w/v) Yeast extract
10 mM NaCl
2.5 mM KCl
10 mM MgCl2
10 mM MgSO4
20 mM glucose

### 3.1.2 Media for cell work

**Dulbecco's Modified Eagle's Medium (DMEM) (Sigma-Aldrich), D6546:**
4500 mg/L glucose, sodium pyruvate, and sodium bicarbonate
10% FBS
4 mM L-glutamine
2% penicillin-streptomycin (Pen-Strep)

**Iscove's Modified Dulbecco's Medium (IMDM) (Thermofisher Scientific), 12440061:**
4500 mg/L glucose, 4 mM L-glutamine and 25 mM HEPES
10% FBS
2% Penicillin-streptomycin (Pen-Strep)

### 3.1.3 Buffers and solutions for SDS-PAGE and Western blotting

**1X SDS electrophoresis buffer**
10% Bio-Rad 10x TGS, pH 8.3 (161-0772) containing:
25 mM Tris
192 mM glycine
0.1% SDS

**1X Blotting buffer**
10% Bio-Rad 10x TG, pH 8.3 (161-0771) containing:
25 mM Tris
192 mM glycine

**10X Phosphate buffered saline (PBS)**
80 g/l NaCl
2 g/l KCl
20 g/l Na2HPO4 x $2H_2O$
4 g/l KH2PO2 pH 7.4

**1X PBS-Tween (PBST)**
1X PBS
0.05% (v/v) Tween-20

**5% Blocking buffer**
5% (w/v) dry milk
1X PBS

### 3.1.4 Buffers and solutions for immunoprecipitation and $^{14}$C-acetylation assay

**IPH lysis buffer**
50 mM Tris-HCl, pH 8.0
150 mM NaCl
5 mM EDTA
0.5% NP-40
1X complete EDTA free protease inhibitor

**2x Acetylation buffer**
100 mM Tris-HCl
2mM EDTA
20% Glycerol
pH 8.5

**HEPES buffer**
10 mM HEPES
pH 7,4

### 3.1.5 Buffers and solutions for mass spectrometry analysis

**Wash solution**
250 µl 1M Ambic (frozen stock solution of 1M ammonium bicarbonate in MilliQ water, 80 mg/ml) added to 4750 µl MilliQ water and 5 ml ACN (HPLC grade)

**10 mM DTT in 100 mM Ambic**
10 µl 1M DTT (frozen stock solution of 154 mg DTT/ml MilliQ water) added to 890 µl MilliQ water and 100 µl 1M Ambic.

**55 mM IAA in 100 mM Ambic**
10 mg IAA added to 900 µl MilliQ water and 100 µl 1M Ambic.

**Digestion buffer**
50 µl 1M Ambic, 10 µl 0.1M $CaCl_2$ and 50 µl ACN added to 890 µl MilliQ water.

**6 ng/µl Trypsin**
10 µl Trypsin Promega Porcine stock solution (100 ng/µl dissolved in 50 mm acetic acid) and 160 µl digestion buffer.

**FASP buffer**
2% SDS
100 mM Tris-HCl pH 7.6
0.1 M DTT

**UA buffer**
8 M urea
100 mM Tris-HCl pH 8.0

## 3.2 ANTIBODIES

**Table 3.1:** Primary antibodies used for Western blot analysis

| Name | Description | Application (Dilution) | Concentration (µg/µl) | Supplier (Catalog #) |
|---|---|---|---|---|
| Anti-V5 | Mouse, monoclonal | IP<br>WB (1:5000) | 1.2 | Invitrogen (46-1157, R96025) |
| Anti-NAA15/NATH | Rabbit, polyclonal | IP<br>WB (1:2000) | 1 | Biogenes (Arnesen *et al*., 2005a) |
| Anti-NAA10 **(ARD1A)** | Rabbit, monoclonal | WB (1:1000) | | Cell Signaling (13357) |
| Anti-NAA10 | Rabbit, polyclonal | WB (1:1000) | 0.73 | Abcam (ab155687) |
| Anti-Xpress | Mouse, monoclonal | IP<br>WB (1:2000) | 1.5 | Invitrogen (46-0528, R91025) |
| Anti-Vinculin | Rabbit, monoclonal | WB (1:2000) | | Abcam (ab129002) |
| Anti-HA | Rabbit, polyclonal | IP<br>WB (1:2000) | 1 | Abcam (ab9110) |
| Anti-HIF-1α | Rabbit, monoclonal | WB (1:1000) | | Cell Signaling (36169T) |
| Anti-FIH | Mouse, monoclonal | WB (1:500) | 0.2 | Santa Cruz (sc271780) |
| Anti-β-tubulin | Mouse, monoclonal | WB (1:3000) | | Merck (T5293) |

**Table 3.2:** Secondary antibodies used for Western blot analysis

| Name | Description | Application (dilution) | Concentration (µg/µl) | Supplier (Catalog #) |
|---|---|---|---|---|
| Anti-rabbit IgG, HRP-linked whole antibody | Donkey | WB (1:5000) | 1 | GE Healthcare (NA934) |
| Anti-mouse IgG, HRP-linked whole antibody | Sheep | WB (1:5000) | 1 | GE Healthcare (NA931) |

## 3.3 APPARATUS AND INSTRUMENTS

**Table 3.3:** Apparatus and instruments

| Supplier | Name | Function |
|---|---|---|
| Bio-Rad | ChemiDoc XRS+ | Western blot membrane analyses |
| | Cell counter TC20 | Count cells for seeding |
| | Geldoc™ EZ Imager | Protein gel analyses |
| Eppendorf | Concentrator 5301 | Vacuum evaporator for drying gel pieces prior to mass spectrometry analysis |
| | Mastercycler Gradient | Thermocycler for sequencing PCR |
| Perkin Elmer | Tri-Carb 2900TR Liquid Scintillation Analyzer | $^{14}$C-signal measurement for $^{14}$C-acetylation assay |
| Saveen Werner | NanoDrop ND-2000 | Nucleic acid concentration measurement |
| Thermo Scientific | Q-Exactive HF mass spectrometer | Mass spectrometer to analyze hydroxylation status on NAA10 |

## 3.4 CENTRIFUGES

**Table 3.4**: Centrifuges

| Supplier | Name |
|---|---|
| Eppendorf | Eppendorf Centrifuge 5415R |
| Thermo Fisher | Heraeus Fresco 17 |

## 3.5 BACTERIAL STRAINS

**Table 3.5:** Bacterial strains used for protein expression and midi prep

| Name | Use | Supplier |
|---|---|---|
| NEB® 5-alpha Competent E. coli cells (NEBC2987) | Midi prep, protein expression | New England Biolabs |

## 3.6 COMMERCIAL KITS

**Table 3.6:** List of commercial kits

| Name | Use | Supplier |
|---|---|---|
| Mini-PROTEAN® TGX™ precast gels | SDS-PAGE | Bio-rad |
| NucleoBond® Xtra Midi Plus | Midi prep | Macherey-Nagel |
| SuperSignal™ West Pico PLUS Chemiluminescent Substrate | Chemiluminescence for Western blotting | Thermofisher-Scientific |

## 3.7 CHEMICALS AND REAGENTS

**Table 3.7:** List of chemicals and reagents used

| Supplier | Name | Cat. no. |
|---|---|---|
| **Amersham Biosciences** | Dithiothreitol | 171318-02 (PROBE) |
| **AppliChem GmbH** | Dry milk powder | A0830 |
| **Bio-Rad** | Bio-Safe™ Coomassie stain solution | 1610786 |
| | Laemmli SDS sample buffer, 4x | 161-0747 |
| | 10x TG | 1610771 |
| | 10x TGS | 1610772 |
| **Fluka analytical** | Ammonium bicarbonate ($(NH_4)HCO_3$) | 40867 |
| | Formic acid | 56302 |
| **Gibco** | Trypsin-EDTA (10X) | 15400-054 |
| **Honeywell** | Acetonitrile | 34967 |
| **Invitrogen** | Dynabeads™ Protein G | 10004D |
| **Lonza** | L-Glutamine | Be17-605E |
| **Merck Millipore** | EDTA | 1.08417.0250 |
| | Methanol | 32213 |
| | MG132 | 474790 |
| | Microcon 30 kDa MWCO filters | MRCF0R030 |
| | Potassium dihydrogen phosphate ($KH_2PO_4$) | 1.04873.1000 |
| | Ponceau staining solution | P7170 |
| | Sodium phosphate dibasic (Na2HPO4·2H2O) | 71645 |
| | Tris | 108382 |
| | Tryptone | 1.07213.1000 |
| **New England Biolabs** | S.O.C medium | B9020S |
| **Promega** | Trypsin/Lys-C mix, Mass Spec grade | V507A |
| | Trypsin Porcine | V511A (PROBE) |
| **Roche** | Complete EDTA free protease inhibitor cocktail tablet | 11 873 580 001 |
| | X-tremeGENE 9 DNA Transfection Reagent | 06 365 809 001 |
| **Sigma-Aldrich** | Agar | Y1500 |
| | Dithiotheritol | 43816 |
| | Fetal bovine serum | F7524 |
| | IGEPAL® CA-630 (NP-40) | 18896 |
| | Iodoacetamide | I-1149, I-6125 |
| | Isopropanol | I-9516 |
| | Glycerol | G5516 |
| | Penicillin-Streptomycin | P0781 |
| | Potassium chloride (KCl) | P9541 |
| | Sodium dodecyl sulfate (SDS) | 74255 |
| | Sodium chloride (NaCl) | 31434 |
| | Tween-20 | |
| | Yeast extract | 89526 |
| **Supelco** | Empore ™ Octadecyl C18 Extraction Disks | 66883-U |
| **Thermo Scientific** | Gibco™ Opti-MEM® reduced serum medium | 31985-070 |
| | PageRuler Plus, prestained protein ladder | 26619 |
| | Pierce ™ C18 tips, 100 µl | 87784 |
| | PageRuler™ Plus Prestained Protein Ladder | 26619 |
| **VWR Chemicals** | Ampicillin sodium salt | 0339-EU |
| | Ethanol (100%) | 20821.330 |

### 3.8 HUMAN CELL LINES

**Table 3.8:** Human cell lines used for transfection and protein expression

| Cell line | Species | Tissue | Cell type | Supplier |
|---|---|---|---|---|
| HeLa | Human | Cervix | Adenocarcinoma | |
| HEK293 | Human | Kidney | Embryonic kidney cells | |
| HAP1 FIH KO | Human | | Leukocytes | Horizon (HZGHC55662) |
| HAP1 WT | Human | | Leukocytes | Horizon (HD PAR-595) |

### 3.9 PLASMIDS

**Table 3.9**: Plasmids used for protein expression

| Name | Protein | Properties | Expression | Supplier |
|---|---|---|---|---|
| pcDNA3 | HA-hHIF1α | Ampicillin resistance, N-terminal HA-tag | Mammalian | Yu *et al*., 2001 |
| pcDNA3.1/V5 His TOPO | hNAA10-V5 | Ampicillin resistance, C-terminal V5-tag, 6x His-tag | Mammalian | Arnesen *et al*., 2005a |
| pcDNA3.1/V5 His TOPO | hFIH-V5 | Ampicillin resistance, C-terminal V5-tag, 6x His-tag | Mammalian | Scholz *et al*., 2016 |
| pcDNA4 HisMax TOPO | Xpress-hNAA10 | Ampicillin resistance, N-terminal Xpress-tag, 6x His-tag | Mammalian | Arnesen *et al*., 2005b |
| pcDNA3.1 V5-His-EMPTY | - | Ampicillin resistance, C-terminal V5-tag and His-tag | Mammalian | Invitrogen |

## 3.10 PRIMERS

**Table 3.10:** List of primers

| Name | Sequence (5'-3') | Tm (°C) |
|---|---|---|
| **oTA01, T7 F** | TAATACGACTCACTATAGGG | 49.0* |
| **oTA02, CMV F** | GAGGTCTATATAAGCAGAGC | 52.3 |
| **oTA12, Xpress F** | TATGGCTAGCATGACTGGT | 58.2 |
| **oTA35, BGH R** | TAGAAGGCACAGTCGAGG | 58.8 |
| **oTA1049, HIF-1α 459 F** | CACAGTGTGTTTGATTTTACTCATC | 61.3 |
| **oTA1050, HIF-1α 958 F** | CTAAAGGACAAGTCACCACAGG | 62.5 |
| **oTA1051, HIF-1α 1457 F** | TCGAAGTAGTGCTGACCCTG | 63.1 |
| **oTA1052, HIF-1α 1954 F** | CAAAAGACCGTATGGAAGACA | 61.5 |

* Calculated using the NEB Tm Calculator

## 3.11 SOFTWARE

**Table 3.11:** List of software used

| Name | Use | Supplier |
|---|---|---|
| **ImageLab (v. 6.0.1)** | Imaging of SDS-polyacrylamide gels and membranes | Bio-Rad |
| **MaxQuant (v. 1.6.8.0)** | Identification and quantification of peptides for LC-MS/MS analysis | Cox and Mann, 2008; Cox *et al.*, 2014; Cox *et al.*, 2011 |
| **Perseus (v. 1.6.5.0)** | Interpretation of RAW data retrieved from MaxQuant | Tyanova *et al.*, 2016 |
| **PyMOL (v. 2.3.2)** | Creating structural figures | Schrödinger |
| **Smart** | Creating illustrations | Servier Medical Art (Creative Common Attribution 3.0 Generic License) |
| **SnapGene (v. 3.4.1)** | Visualization of vector maps | GSL Biotech |

# 4 METHODS

## 4.1 GENERAL STRATEGY

The purpose of this thesis has been to functionally investigate whether cellular NAA10 is Trp38-hydroxylated in different cell lines, and to look for interactions and functional links between NAA10 and FIH. The methods applied in this study are presented in Figure 4.1.

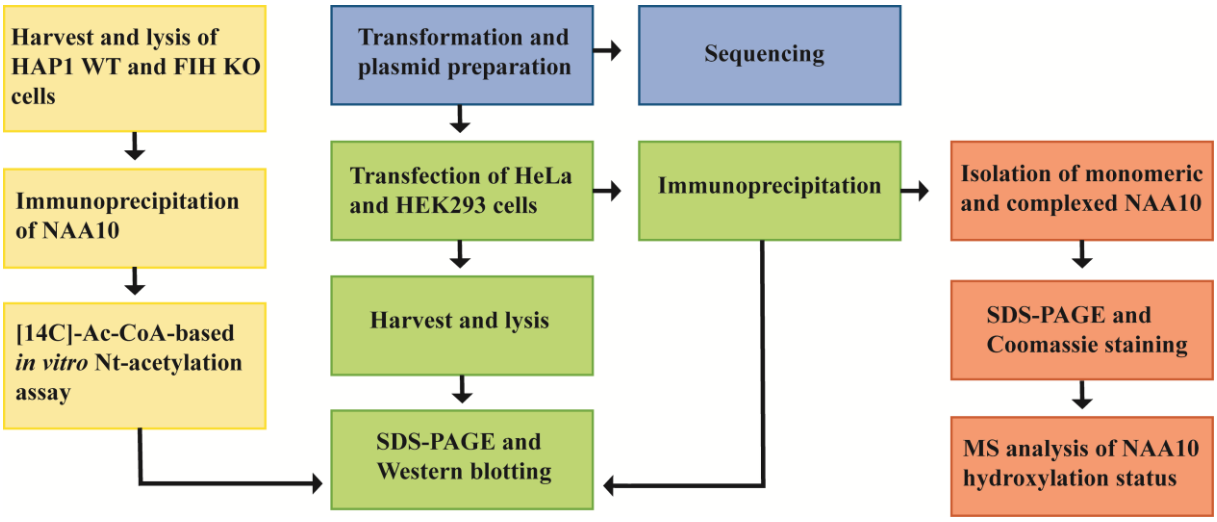


**Figure 4.1: General work flow summarizing the main steps of this project.** Methods applied in initial DNA work (blue), *in vitro* NatA and NAA10 activity assays in presence and absence of FIH (yellow), the study of protein expression and stability (green), and mass spectrometry analysis of the hydroxylation status of NAA10 (orange).

## 4.2 PLASMID PREPARATION

### 4.2.1 Transformation by heat shock

For plasmid preparation competent NEB® 5-alpha *Escherichia coli* (*E. coli*) cells (50 µl) were transformed with pcDNA3-HA-HIF1α and pcDNA4-Xpress-hNAA10 constructs through heat shock transformation. The cells were transformed with 3 µl of the respective plasmid constructs by 30 min incubation on ice followed by heat shock at 42°C for 30 sec. The cells were immediately placed on ice and mixed with 250 µl SOC medium, followed by 1 h of incubation at 37°C whilst shaking at 250 rpm. 40, 75 and 100 µl of the cells were spread out on LB agar plates containing 100 µg/ml ampicillin. The plates were incubated at 37°C overnight (ON).

### 4.2.2 Plasmid preparation

Colonies of *E. coli* expressing the pcDNA3-HA-HIF1α and pcDNA4-Xpress-hNAA10 constructs were picked and inoculated in 5 mL LB media for 8 h. The same applied for glycerol stocks of NEB® 5-alpha cells carrying the pcDNA3.1/V5 His TOPO vector with hFIH-V5 or hNAA10-V5. The cultures were then transferred to 100 mL LB media containing 100 µg/ml ampicillin and incubated ON at 37°C whilst shaking at 250 rpm. NucleoBond® Xtra Midi Plus plasmid Midiprep kit was used to isolate the plasmids as described by the manufacturer (Macherey Nagel), and the concentration of eluted DNA was measured by NanoDrop. Each plasmid sequence was verified by DNA sequencing.

### 4.2.3 Sequencing

The Big Dye terminator v3.1 sequencing kit was used to determine the plasmid DNA sequences. The samples were prepared as listed in Table 4.1a and the PCR was run using the PCR program described in Table 4.1b. 10 µl of $ddH_2O$ was added to each sample prior to sequencing performed at the DNA sequencing facility at the University of Bergen.

**Table 4.1a:** Sequencing reaction mix

| Component | Amount |
|---|---|
| Big Dye v3.1 | 1 µl |
| Sequencing buffer | 1 µl |
| Template | 200 ng |
| Primer | 3.2 pmol |
| $ddH_2O$ | Up to 10 µl |

**Table 4.1b:** Sequencing PCR program

| Step | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 96°C | 5 min |
| 2 | 25 | 96°C | 10 sec |
| | | 50°C | 5 sec |
| | | 60°C | 4 min |
| 3 | 1 | 4°C | ∞ |

## 4.3 MAINTENANCE OF CELLS

### 4.3.1 HEK293 and HeLa cells

HEK293 and HeLa cells were cultured in 10 cm dishes with 10 mL DMEM containing 10% fetal bovine serum (FBS), 3% L-glutamine, and 1% PenStrep at 37°C and 5% $CO_2$ under sterile conditions using a laminar flow hood to avoid contamination. When reaching confluency, the cells were washed with 1X PBS and detached by trypsinization at 37°C for 5-10 min. Detached cells were resuspended in appropriate amounts of fresh medium, and usually split 1:10 to new 10 cm dishes for further subculturing. 24 h prior to transfection the cells were split 1:3 in 10 cm dishes or 6-well plates.

### 4.3.2 HAP1 WT and FIH KO cells

HAP1 WT and FIH KO cells were cultured in 10 cm dishes with 10 mL IMDM containing 10% FBS and 1% PenStrep at the same conditions as described for HEK293 and HeLa cells. At 50-60% confluency the cells were washed and detached by trypsinization as previously described. Detached cells were resuspended in appropriate amounts of fresh medium and split 1:20 to new 10 cm dishes for further subculturing. 24 h prior to harvest and lysis $3x10^5$ cells were seeded in 6-well plates. Both HAP1 cell lines were passaged until diploidy was confirmed using flow cytometry.

### 4.3.3 Transfection

Transfection is the process of introducing foreign DNA to eukaryotic cells. Both HEK293 and HeLa cells were used for this purpose in order to analyze the hydroxylation status of monomeric and complexed NAA10, and to study FIH and NAA10 protein expression. For experiments testing protein expression from lysates the protein content was further analyzed by WB directly after harvest and lysis. For LC-MS/MS analysis and experiments investigating protein interactions the cell lysates were also subject to IP prior to WB analysis.

The cells were transfected at approximately 70% confluency when they had reached the log-phase of growth in order to increase the transfection efficiency. 24 h prior to transfection cells were seeded in 10 cm dishes for IP and in 6-well plates for lysis only. The transfection mixture was prepared by diluting 4 µg or 1.2 µg of vector in Opti-MEM, before adding the XtremeGene9 transfection reagent (Roche) in a 1:3 ratio (µg:µl). After 15 min of incubation at room temperature (RT) the cells were transfected by dropwise adding 500 µl of the transfection mix per 10 cm dish (150 µl for 6-well plates). The cells were incubated at 37°C with 5% $CO_2$, and the growth medium was replaced 24 h post transfection to avoid cell death due to cytotoxicity caused by the transfection reagent.

## 4.4 STUDY OF PROTEIN EXPRESSION AND INTERACTION

### 4.4.1 Immunoprecipitation

To study potential interactions between NAA10, FIH and HIF-1α proteins, as well as to enrich NAA10-V5 for LC-MS/MS analysis of PTM status, IP using specific antibodies was performed. IP is a technique used for isolation of specific proteins from cell lysates, which was done by using antibodies and magnetic Dynabeads™ Protein G (Invitrogen) in this study. The beads bind with high capacity to the heavy chains of antibodies, and can be isolated by the use of a magnet allowing for removal of the supernatant and bead washing (Thermofisher). HEK293 and HeLa cells were both used for IP experiments where they were transfected with 4 µg of either FIH-V5, Xpress-NAA10, HA-HIF 1α or NAA10-V5 vectors. For the experiments where a double transfection was carried out the same amount of an empty-V5 vector was co-transfected to make sure the conditions were equal.

Cells were washed and harvested 48 h post transfection in 1400 µl cold 1X PBS using a cell scraper, and the cells were centrifuged at 4 °C at 1000 x g. The pellets were then resuspended in IPH lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1X protease inhibitor) (350 µl per 10 cm dish and 105 µl per well in a 6-well plate) and was incubated on ice for 30 min. In order to remove cell debris, the lysate was centrifuged at 17000 x g for 5 min and the supernatant was transferred to a new tube. 30-50 µl of the lysate was saved for later WB analysis.

The cell lysate was incubated with 2.5 µg antibody per 10 cm dish at 4°C on a rotating wheel for 2-3 h. After preparing 30 µl Dynabeads™ per 10 cm dish by washing them three times in 1 ml IPH buffer they were added to each sample. In order to retrieve the immune complexes, the samples were incubated ON at 4 °C on a rotating wheel. The next day, the beads were isolated using a magnetic holder, and the supernatant was saved for later WB analysis. The beads were washed 3x in 1 ml IPH buffer, and resuspended in 100 µl 1X sample buffer followed by 5 min boiling at 95°C. In cases where the lysate was split in two for Xpress-IP and V5-IP, 30 µl and 20 µl Dynabeads™ were used respectively.

### 4.4.2 $^{14}$C-Ac-CoA-based *in vitro* acetylation assay

Immunoprecipitated NAA10 and NatA complexes were used in an *in vitro* $^{14}$C-Ac-CoA-based acetylation assay for measuring the catalytic activity. $^{14}$C-Ac-CoA is used as a donor upon acetylation of peptides by NAA10, and the product formation can be detected and quantitated due to radioactivity. A P18 phosphocellulose filter quenches the enzymatic reaction through binding of positively charged peptides, and excess $^{14}$C-Ac-CoA is washed off. By covering the filter in scintillation fluid, energy from β-particles is absorbed and results in photons being emitted. The acetylated peptides can be quantified based on the light emitted, and the measured activity is then adjusted to the amount of NatA or NAA10 used in the reactions, which is determined from WB analysis. Even low levels of acetylation can be detected with this highly sensitive method (Drazic and Arnesen, 2017).

$^{14}$C-acetylation assays were performed pairwise for FIH KO and WT HAP1 cells in order to compare the catalytic activity of immunoprecipitated NatA and NAA10, as well as evaluate the potential NAA10 KAT activity towards a Nt-acetylated histone 4 peptide. The DynaBeads™ with the immunoprecipitated NAA10 were washed three times in IPH lysis buffer and resuspended in 125 µl acetylation buffer (100 mM Tris-HCl pH 8.5, 2mM EDTA, 20% glycerol). Three replicates with 200 µM peptide (SESS, EEEI or Nt-acetylated histone 4), 50 µM $^{14}$C-labelled Ac-CoA, 10 µl IP beads and $dH_2O$ were mixed to a final volume of 25 µl. Two negative replicates without peptide were also prepared. All samples were incubated at 37°C and 1300 rpm on a thermoshaker for 30 min, except for the SESS samples which were incubated for 15 min. After incubation, the magnetic beads were isolated using a magnetic holder and 23 µl sample was transferred to P81 phosphocellulose filter squares. The filter squares were washed 3x5 min in 10 mM HEPES buffer (pH 7.4) followed by air drying on paper. The dried filter squares were placed in individual tubes and soaked in 5 ml scintillation fluid for $^{14}$C-signal measurement using a scintillation counter (Tri-Carb 2900TR Liquid Scintillation Analyzer, Perkin Elmer). The IP samples were further analyzed by WB (as described in Methods section 4.4.4) and the measured activity was adjusted to quantification of corresponding anti-NAA10 and anti-NAA15 bands.

### 4.4.3 SDS-PAGE

In sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), proteins are separated according to their size under denaturing conditions. The protein samples are mixed with a sample buffer containing the anionic detergent SDS, as well as dithiothreitol (DTT), glycerol and bromophenol blue. SDS binds and denatures the proteins by providing them with an overall negative charge proportional to their mass, and the DTT reduces the disulfide bonds. When applying an electric field to the gel the proteins are separated as they will migrate towards the positively charged anode. As the larger proteins encounter more resistance in the gel upon migration, they will migrate slower than the smaller proteins.

The lysates and IP samples collected 48 h post-transfection were mixed with sample buffer (4X and 1X respectively) and boiled for 5 min at 95°C. Most samples were loaded to a Mini-PROTEAN® TGX Stain-Free™ Precast Gel (Bio-Rad), whereas the samples for MS-analysis were loaded to an ordinary Mini-PROTEAN® TGX Precast Gel (Bio-Rad) followed by staining with Coomassie Brilliant Blue (Bio-Rad) for 2-3 h and destaining ON. PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific) was used as a molecular weight marker, and the gels were run in 1X TGS electrophoresis buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) at 100 V for 5 min and then at 200 V for 40 min.

### 4.4.4 Western Blotting

The cell lysates and IP samples were analyzed by WB for detection of specific proteins separated by SDS-PAGE. After separation the proteins are transferred from the gel to a nitrocellulose membrane in a blotting sandwich made up of filter paper and sponges by applying an electric field. By facing the membrane towards the anode, the negatively charged proteins in the gel will migrate towards the membrane (Towbin *et al*., 1979). The proteins can be detected by treating the membrane with specific antibodies with features allowing for detection through fluorescence or chemiluminescence, such as fluorophore- or enzyme-conjugated antibodies.

For WB a blotting sandwich was prepared and placed with the membrane facing the positive electrode in an electroblotting chamber. The proteins were transferred in 1X TG buffer (25 mM Tris, 192 mM glycine, pH 8.3) at 100 V for 40 min and the blotting was confirmed by adding Ponceau staining solution. To prevent unspecific binding of antibodies the membrane was blocked in 5% dry milk in 1X PBST at RT for 1 h with gentle shaking, before it was incubated with primary antibody in 1% dry milk ON at 4°C on a rocking shaker. The membrane was rinsed for unbound antibody by washing three times with 1X PBST for 5 min prior to 1 h incubation with secondary antibody in 3% dry milk at RT. The membrane was washed with 1X PBST three times and with 1X PBS once for 5 min and incubated in a 1:1 mix of the luminol and peroxide reagents from SuperSignal® West Pico Chemiluminescent Substrate (Thermo Scientific). This allowed for visualization of the protein of interest as the horseradish peroxidase (HRP) bound to the secondary antibody catalyzes the oxidation of the HRP substrate present (Figure 4.2). A chemiluminescent product is formed that can be detected with ChemiDoc™ XRS+ connected with ImageLab™ v. 6.0.1 (Bio-Rad).

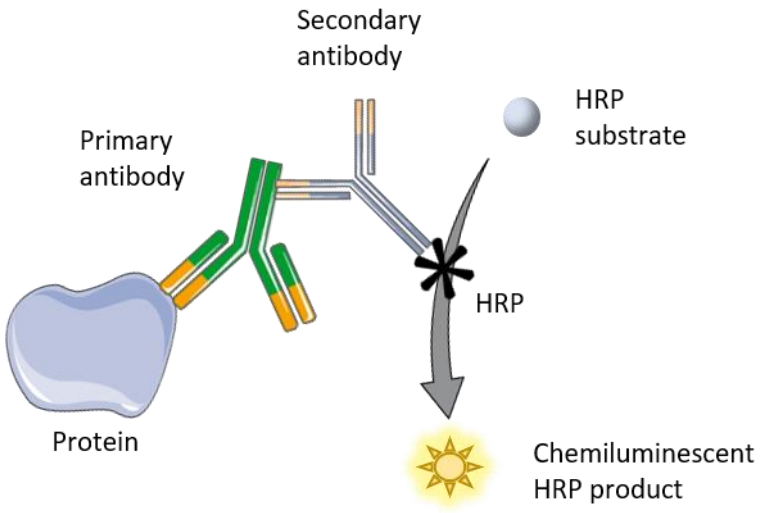


**Figure 4.2: Detection of proteins using chemiluminescence from HRP-linked secondary antibodies.** Primary antibodies recognize and bind to the target protein (blue), and the HRP-linked secondary antibodies recognize the primary antibodies. In the presence of HRP substrate, HRP catalyzes the oxidation of the substrate molecules and the product formed emits detectable light. (Modified from Smart by Servial Medical Art)

## 4.5 LC-MS/MS ANALYSIS

A total of 17 V5- and NAA15-IP samples were included in a dataset for analysis of NAA10 hydroxylation status. The samples were obtained from HAP1, HEK293 and HeLa cells and were prepared for LC-MS/MS analysis by following in-gel or filter-aided sample preparation (FASP) protocols. In order to analyze the hydroxylation status of NAA10 in monomeric form and in complex with NAA15, NAA10-V5 transfected cells and untransfected cells were harvested and lysed. The cell lysates were further subject to anti-V5 and anti-NAA15 IP as described in Methods section 4.4.1 and the samples that were prepared with in-gel digestion were loaded onto an SDS gel. The gel slice between 25 and 55 kilodalton (kDa) for each lane was excised. They were further processed at the Proteomics Unit at University of Bergen (PROBE) as described below. The FASP method is further explained in Methods section 4.5.2.

### 4.5.1 Alkylation, washing and in-gel digestion

At PROBE the gel from the samples was cut into 1 mm cubes and covered in wash solution (see Materials section 3.1.5) prior to 20 min incubation at RT with shaking. The washing step was repeated once, and a vacuum evaporator was used to dry the gel pieces. In order to reduce all disulfide bridges, the gel pieces were rehydrated in 50 µl 10 mM DTT (Amersham Biosciences) and incubated for 45 min at 56°C, followed by 30 min incubation at RT in the dark with 50 µl 55 mM iodoacetamide (IAA) (Sigma Aldrich) for cysteine alkylation to avoid reformation of disulfide bridges. The gel pieces were washed twice as described above and dried using the vacuum evaporator. For protein digestion PROBE rehydrated the Coomassie stained gel pieces in 6 ng/µl Trypsin Porcine (Promega) on ice for 30 min, and the samples were incubated at 37°C ON. After cooling and centrifugation, the supernatant was transferred to a new tube and the gel pieces were incubated once in 5% formic acid (FA) and once in 60% ACN/0.1% FA for 20 min at RT with shaking. Both supernatants were pooled with the first extraction. The samples were dried using the vacuum evaporator until 10-15 µl remained to remove all traces of ACN.

The samples not processed by PROBE were washed in 1:1 water/acetonitrile (ACN), followed by dehydration in 100% ACN and rehydration in 100 mM ammonium bicarbonate (ABC). An equal volume of 100% ACN was added after 5 min, and the solution was incubated for 15 min. The alkylation and subsequent washing steps were performed as described. The gel pieces were then covered in digestion solution (100mM ABC, 0.5 mM $CaCl_2$, 12.5 ng/µl modified sequencing-grade trypsin/LysC mix from Promega) prior to incubation as described above. The following day the gel pieces were covered and incubated in 100 mM ABC for 15 min prior to addition of equal volume of ACN and 15 min incubation for peptide extraction. This was repeated twice with 5% FA instead of ABC, and the supernatant was pooled and dried in a vacuum evaporator.

### 4.5.2 FASP

In order to process the samples for LC-MS/MS analysis some of the samples were prepared with FASP. After ON incubation of the samples with Dynabeads™ as described in Methods section 4.4.1 the beads were washed three times in IPH lysis buffer followed by elution of the bound immunocomplexes in 60 µl FASP buffer (2% SDS, 100 mM Tris-HCl pH 7.6, 0.1 M DTT). The samples were then heated for 5 min at 95°C. The protein samples were mixed with 200 µl UA buffer (8 M urea, 100 mM Tris-HCl pH 8.0), transferred to Microcon 30 kDa MWCO filters and centrifuged. All centrifugation steps were carried out at 23°C and 14 000 x g for 15 min. The filters were washed three times with 200 µl UA, and the proteins were subsequently Cys-alkylated by incubation with 100 µl 50 mM IAA in UA for 20 min. IAA was removed by centrifugation, and filters were washed three times with 100 µl UA. Subsequent washing with 100 µl 50 mM ABC three times to exchange the buffer was followed by addition of 500 ng trypsin (Promega) to the filters and ON incubation at 37°C. The next day the digested proteins were collected by centrifugation and the filter was washed once with 75 µl ABC. In order to acidify the peptides 5% FA was added, and the Pierce™ C18-Tips (Thermo Scientific) was used for desalting according to manufacturer´s protocol. The final eluates were dried by speed vacuum and diluted to desired concentrations with 5% FA.

### 4.5.3 Trapping and desalting

PROBE further prepared the samples by injecting 5 ng of the tryptic peptides dissolved in 2% ACN and 0.5% FA into an Ultimate 3000 RSLC system (Thermo Scientific) connected online to a Q-Exactive HF mass spectrometer (Thermo Scientific) equipped with EASY-spray nano-electrospray ion source (Thermo Scientific) by PROBE. The samples were loaded and desalted on a pre-column packed with 3µm C18 beads (Acclaim PepMap 100, 2cm x 75µm ID nanoViper column) at a flow rate of 5µl/min for 5 min with 0.1% trifluoroacetic acid.

For samples not processed by PROBE the remaining 1-2 µl peptides after in-gel digestion and vacuum evaporation were resuspended in 20 µl 5% FA. The samples were further desalted on Pierce™ C18 tips (Thermo Scientific). Buffer A and B were 1 % FA in water and 80% ACN respectively. The tips were conditioned with 200 µl B and equilibrated twice with 200 µl A before the samples were loaded. The flow through was reloaded to the tip followed by washing twice with A. In order to elute the bound peptides 200 µl B was loaded to the tips twice and the eluates were pooled and completely dried in a vacuum evaporator. The peptides were dissolved in 5% FA before they were sent to PROBE for LC-MS/MS analysis.

### 4.5.4 LC-run

The peptides were separated during a biphasic ACN gradient on a 50 cm analytical column packed with 2µm C18 beads (PepMap RSLC, 50cm x 75 µm i.d. EASY-spray column). Solvent A and B were 0.1% FA (vol/vol) in water and 100% ACN respectively. The gradient composition was 5% B during trapping (5 min) followed by 5-8% B over 0.5 min, 8–24% B for the next 109.5min, 24–35% B over 25 min, and 35–80% B over 15 min. Elution of very hydrophobic peptides and conditioning of the column were performed during 15 min isocratic elution with 80% B and 20 min isocratic conditioning with 5% B. The total LC-run time was 195 min.

### 4.5.5 Data-Dependent Acquisition (DDA)-run

The eluting peptides from the LC-column were ionized in the electrospray and analyzed by the Q-Exactive HF. The mass spectrometer was operated in the DDA-mode to automatically switch between full scan MS and MS/MS acquisition. Instrument control was through Q Exactive HF Tune 2.9 and Xcalibur 4.1. MS spectra ranging from m/z 375-1500 were acquired in the Orbitrap with resolution R = 120 000 at m/z 200, automatic gain control (AGC) target of 3 x $10^6$ and a maximum injection time of 100 ms. The 12 most intense eluting peptides (Top12 method) above intensity threshold 50 000 counts and charge states 2 to 5 were sequentially isolated to a target value (AGC) of $10^5$ and a maximum injection time of 110 ms in the C-trap, and isolation width maintained at 1.6 m/z (offset of 0.3 m/z), before fragmentation in the Higher-Energy Collision Dissociation cell. Fragmentation was performed with a normalized collision energy of 28%, and fragments were detected in the Orbitrap at a resolution of 30 000 at m/z 200, with first mass fixed at m/z 100. One MS/MS spectrum of a precursor mass was allowed before dynamic exclusion for 25s with “exclude isotopes” on. Lock-mass internal calibration (m/z 445.12003) was used. The spray and ion-source parameters were as follows: Ion spray voltage = 1800V, no sheath and auxiliary gas flow, and capillary temperature = 275 °C.

### 4.5.6 MS data analysis

The general work flow of the LC-MS/MS analysis of the NAA10 hydroxylation status is shown in Figure 4.3. MaxQuant (v. 1.6.2.6) with the integrated search engine Andromeda was used for analysis of the resulting RAW files in order to identify and quantify peptides and proteins (Cox and Mann, 2008). The data was searched against a database of Swiss-Prot reviewed sequences from the human proteome retrieved from Uniprot (downloaded Nov 9$^{th}$ 2017) containing a total of 20,431 entries. Protein Nt-Ac, Met oxidation as well as Asn, Asp and Trp hydroxylation were set as variable modifications, and carbamidomethylation of Cys residues was set as fixed. Peptide and protein identifications were filtered to a 1% false discovery rate (FDR). Specific trypsin (cleavage after Lys or Arg, except when followed by a Pro) was set as protease specificity, and a maximum of two missed cleavages was tolerated. Minimum peptide length was set to 7, with a maximum peptide mass of 4600 Da. Label-free quantification (LFQ) was included in the search where the "stabilize large LFQ ratios" and "require MS/MS for LFQ comparisons" options were checked (Cox *et al*., 2014). Further analysis of the MaxQuant data and determination of the NAA10 hydroxylation status was carried out using Perseus (v. 1.6.5.0).

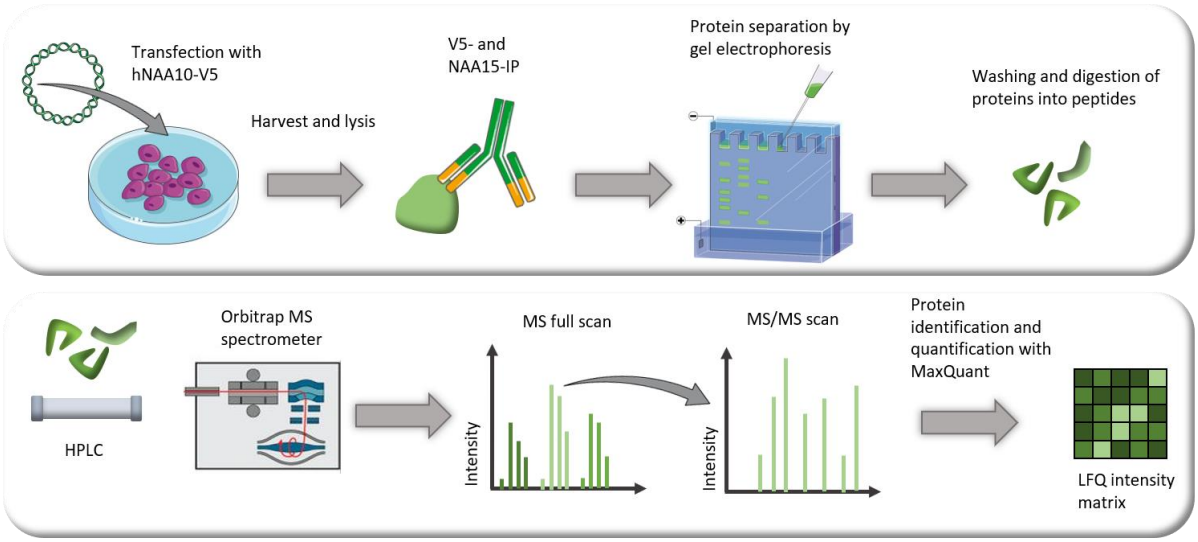


**Figure 4.3: General work flow of the LC-MS/MS analysis of NAA10 hydroxylation status.** HEK293, HeLa and HAP1 cells were transfected with NAA10-V5 and subsequent V5- and NAA15-IP to look for Trp38 hydroxylation. Some samples were subject to offline fractionation on protein level by gel electrophoresis and digested by in-gel, and some samples were processed by FASP. The proteins were digested into peptides for analysis using an Orbitrap MS spectrometer. The resulting RAW files were analyzed in order to identify and quantify proteins using MaxQuant, and the data was further interpreted using Perseus (Adapted from Keilhauer, Hein and Mann, 2015 and modified from Smart by Servial Medical Art)

### 4.5.7 NAA10 deep PTM analysis

An additional set of samples was prepared for deep PTM analysis. 10 dishes of HEK293 cells (10 cm) were transfected at approximately 80% confluence with 5 µg NAA10-V5 for 48 h. The cells were harvested with trypsin, washed twice with PBS and kept on ice for all subsequent steps. The cells were lysed in IPH buffer with 0.5% NP-40, and the lysate was used for IP with 20 µg anti-V5 and 200 µl Dynabeads™. The beads were preloaded with anti-V5 for 20 min and washed twice in IPH buffer. Then the IP proceeded at 4ºC for 1 h. The beads were washed three times with IPH buffer and three times with detergent-free IPH buffer to remove detergents which could interfere with downstream LC-MS/MS analysis. The beads were resuspended in a preheated elution buffer (6 M guanidine-HCl, 5 mM TCEP, 10 mM chloroacetamide, 100 mM Tris-HCl (pH 8.5) and containing 5 ng/µl LysC. The beads were incubated on a shaker for 1 h at 25ºC. The supernatant was transferred to a new tube and the beads were washed once with 250 µl 25 mM Tris-HCl (pH 8.5). The wash solution was combined with the eluate, trypsin was added to a final concentration of 5 ng/µl and incubated ON at 37ºC with shaking. Digestion was stopped by adding trifluoroacetic acid (TFA) to a final concentration of 1%. Peptides were desalted on a 50 mg SEP-PAK C18 column (Waters) on a vacuum manifold. The column was activated with 100% ACN and equilibrated three times with 0.1% TFA (all volumes 1 ml). The sample was loaded and washed three times with 0.1% TFA and eluted by 250 µl 40% ACN/0.1% TFA and then 250 µl 60% ACN/0.1% TFA. Peptides were dried in a Speedvac and resuspended in 100 µl 50 mM ABC for high-pH off-line stagetip fractionation. All stagetips steps were performed into an autosampler plate in a swing-bucket centrifuge at RT and 300 x *g*, with 50 µl volumes. A self-packed stagetip with four C18 disks were activated with 100% methanol and equilibrated three times with 50 mM ABC. The sample was loaded and the flowthrough fraction was collected. Then, elution was performed with increasing steps of ACN in 50 mM ABC (5%, 7%, 12%, 18%, 25%, 30% and 50% ACN) with each fraction collected separately. After the last elution, the plate was dried completely in the Speedvac and peptides were reconstituted in 10 µl 5% ACN/0.1% formic acid.

Each fraction was run in 30-min gradients on a Q Exactive-HF X mass spectrometer (Thermo Fisher Scientific) equipped with an in-house packed 15 cm, 75 µm ID capillary column with 1.9 µm Reprosil-Pur C18 beads (Dr. Maisch), as described (Kelstrup *et al*., 2018). The 1obile phases were 5% ACN and 0.1% FA in water or 100% ACN with 0.1% FA. The gradient ran from 8 to 24% ACN over 24.9 min and 24 to 36% ACN over 5.1 min, followed by a washout by ramping to 64% ACN over 0.5 min. The mass spectrometer was operated in DDA-mode with positive polarity, automatically switching between full MS and MS/MS acquisition. For full MS, the mass range was set to 350-1400 m/z, the AGC target to $3x10^6$, resolution to 60,000 at m/z = 200, and the top 10 most intense ions (Top10 method) selected for sequencing. For fragment spectra, AGC target was $1x10^5$, resolution was 45,000, and isolation width was 1.3 m/z. Fragmentation in the Higher-Energy Collision Dissociation cell was performed at 28% normalized collision energy. Capillary temperature was set to 275ºC. Raw files were searched in Maxquant (v. 1.6.8.0) against a human proteome retrieved from Uniprot on Nov 9th 2017 containing 20,431 protein sequences in the same manner as above, except the variable modifications were set to Lys-Ac, protein Nt-Ac, Met-O, Asn-OH, Asp-OH and Trp-OH.

# 5 RESULTS

## 5.1 LC-MS/MS ANALYSIS

### 5.1.1 Cellular NAA10 is not Trp38-hydroxylated

There have been several studies over the last years focusing on NAA10 and its potential role in Lys-Ac. The KAT-activity of NAA10 is still under debate as studies have shown contradicting results, and one of the aims of this thesis was to look further into the potential role of NAA10 as a KAT. Last year Kang *et al*. reported a potential mechanism of how NAA10 might display KAT activity towards HIF-1α. They found that FIH utilizes oxygen under normoxic conditions to hydroxylate NAA10 at Trp38, and that hydrogen bonding between the OH-group and a carbonyl oxygen of Ala67 causes a conformational change in the NAA10 structure. This is thought to bend the α1-α2 loop that is blocking the active site for Lys residues, and thereby opening the substrate gate of monomeric NAA10 enabling it to acetylate HIF-1α at Lys532. They reported a Trp38-OH abundancy of 26% which is based on their 28 identified Trp38-hydroxylated peptides over the total 106 MS/MS counts.

The hydroxylation status of monomeric and complexed NAA10 was investigated by first transfecting HEK293 and HeLa cells with NAA10-V5 as described in Methods section 4.3.3. The cells were harvested, and transfected and untransfected cells were further subject to V5-IP and NAA15-IP respectively as described in Methods section 4.4.1. The samples were subsequently prepared for MS-analysis by in-gel digestion as described in Methods section 4.5.1. In order to obtain a better coverage, additional V5 and NAA15-IP samples from HAP1, HEK293 and HeLa cells prepared with either in-gel or FASP methods from other experiments were included. All samples used in the dataset are listed in Table 5.1.

**Table 5.1:** List of samples for LC-MS/MS analysis

| Sample | Cell type | Transfection | Enrichment | Method | Prepared by |
|---|---|---|---|---|---|
| **1** | HeLa | NAA15 WT-V5 | V5-IP | FASP | N. McTiernan |
| **2** | HeLa | NAA15 L814P-V5 | V5-IP | FASP | N. McTiernan |
| **3** | HeLa | NAA15 T406Y-V5 | V5-IP | FASP | N. McTiernan |
| **4** | HEK293 | NAA10-V5 | V5-IP | In-gel | K. Krogstad |
| **5** | HeLa | NAA10-V5 | V5-IP | In-gel | K. Krogstad |
| **6** | HEK293 | - | NAA15-IP | In-gel | K. Krogstad |
| **7** | HeLa | - | NAA15-IP | In-gel | K. Krogstad |
| **8** | HEK293 | NAA10-V5 | V5-IP | In-gel | K. Krogstad /PROBE |
| **9** | HeLa | NAA10-V5 | V5-IP | In-gel | K. Krogstad /PROBE |
| **10** | HEK293 | - | NAA15-IP | In-gel | K. Krogstad /PROBE |
| **11** | HeLa | - | NAA15-IP | In-gel | K. Krogstad /PROBE |
| **12** | HAP1 | NAA10-V5 | V5-IP | FASP | N. McTiernan |
| **13** | HAP1 | - | NAA15-IP | FASP | N. McTiernan |
| **14** | HEK293 | NAA10-V5 | V5-IP | FASP | N. McTiernan |
| **15** | HEK293 | - | NAA15-IP | FASP | N. McTiernan |
| **16** | HeLa | NAA10-DelS205-V5 | V5-IP | FASP | R. Ree |
| **17** | HeLa | NAA10-WT-V5 | V5-IP | FASP | R. Ree |

The peptides of each sample were desalted and subsequently subject to LC-MS/MS analysis as described in Methods section 4.5.3-4.5.5. A Q-Exactive HF mass spectrometer was used to look for protein modifications and the raw data for each sample was analyzed using MaxQuant and Perseus as described in Methods section 4.5.6. It was searched against peptides of maximum 4600 Da with a minimum length of 7 amino acids. The variable modifications used were protein Nt-Ac, Met-O, Asp-OH, Asn-OH and Trp-OH, and the fixed modifications used was carbamidomethylation of Cys residues.

The TopN-method was used in these MS analyses where the m/z ratios were measured in cycles in the mass analyzer, and the mass spectrometer detected the N most abundant ions in each cycle for further fragmentation and a second round of MS (MS/MS). For the primary LC-MS/MS analysis, the Top12 method was used. After exclusion of potential contaminants and reversed peptides a total of 70,755 peptides and 5,612 protein groups were identified across the 17 samples. Out of these, approximately 15% (10,713 peptides) was found to be modified with Nt-Ac and/or hydroxylation of Asp, Asn and Trp, as well as oxidation of Met residues. Nt-Ac was detected for 1482 peptides (2%), Asp-OH for 1077 peptides (1.5%), Asn-OH for 641 peptides (0.9%), Trp-OH for 267 peptides (0.4%) and Met-O for 8275 peptides (11.7%) in total. The Trp38-peptide reported by Kang *et al*. was detected in 11 of the 17 samples and has the position 30-51 relative to the 235 amino acid NAA10 protein sequence. The peptide had a total MS/MS count of 37 and it was not found to be hydroxylated in any of the samples (Figure 5.1).

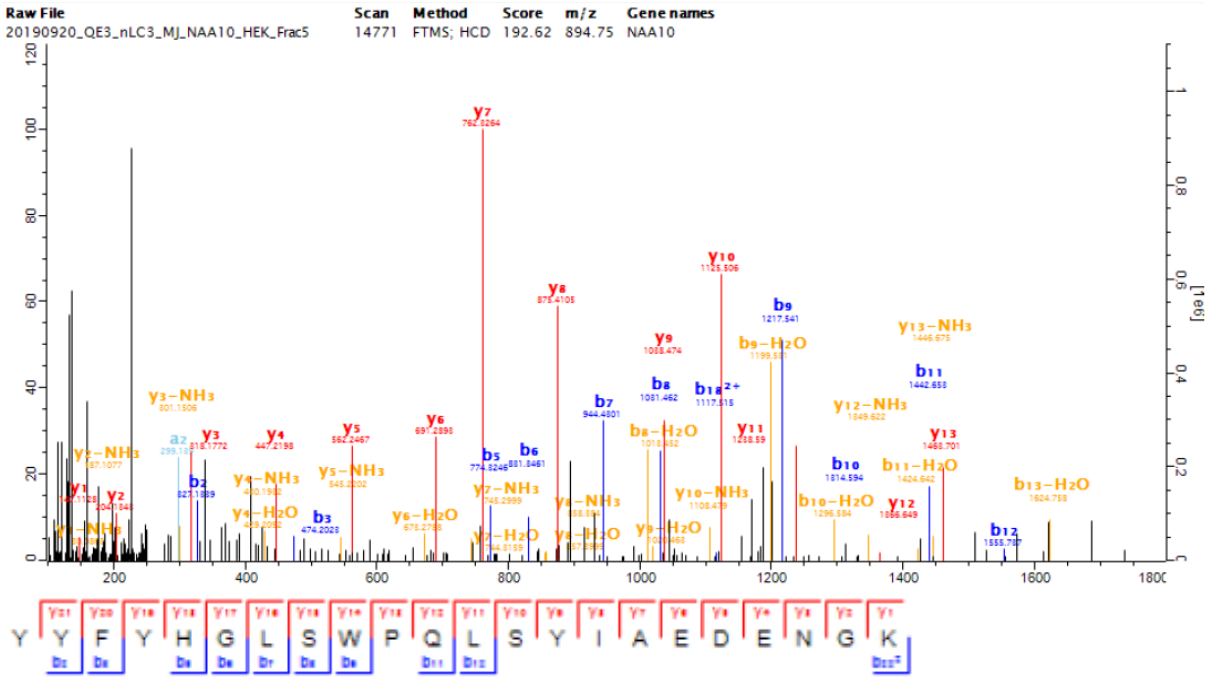


**Figure 5.1: LC-MS/MS spectrum showing the NAA10 peptide with unmodified Trp38.** Across the 17 samples included in the dataset 70,755 peptides were identified and 15% was found to be modified with Nt-Ac, Asn-OH, Asp-OH, Trp-OH or Met-O. The peptide of Trp38 with position 30-51 in the NAA10 amino acid sequence was detected in 11 samples with a total MS/MS count of 37 but was found to be unmodified in all samples.

Only 2.7% of the other peptides in the dataset had MS/MS counts higher than 37. This is not unexpected as V5-IP was performed to enrich for NAA10 in the samples. The distribution of MS/MS counts for all the detected peptides is shown in Figure 5.2.

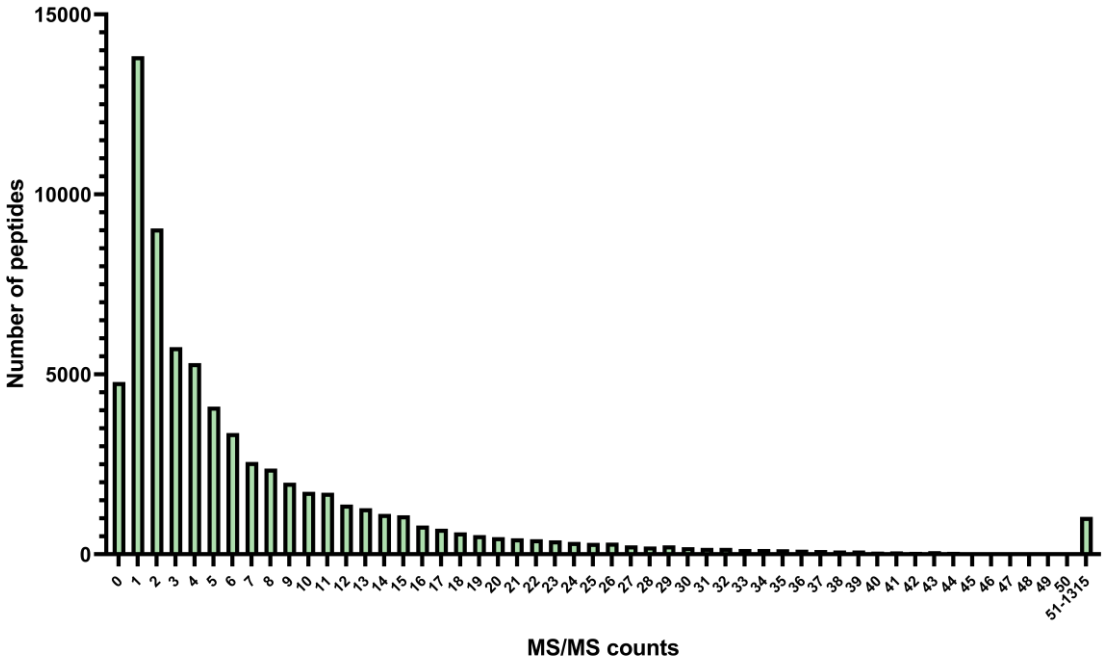


**Figure 5.2: Overview of the distribution of MS/MS counts with the number of detected peptides.** A total of 70,755 peptides were detected in the primary LC-MS/MS analysis. The peptide reported to be Trp38 hydroxylated was detected in 11 of the 17 samples with a total MS/MS count of 37 but was found to be unmodified. The MS/MS count for this peptide was higher than for 97% of the other peptides identified, which was expected due to the NAA10 enrichment by IP.

After looking into the LC-MS/MS analysis data from the study by Kang *et al*. it became evident that the majority of Trp38-hydroxylations were detected for a peptide 8 amino acids longer (bold) than the reported peptide due to a missed cleavage after Lys51 (YYFYHGLSWPQLSYIAEDENGK**IVGYVLAK**). Only 9 of the 106 detected MS/MS counts were of the reported sequence, whereas the remaining 97 peptides had a missed cleavage and a 30 amino acid peptide sequence. This peptide was not detected in our dataset. In three cases Kang *et al*. found the reported 22 amino acid peptide to be Trp38 hydroxylated, and in one case the same peptide was also Asn49 hydroxylated (Supplementary Table 8.1). For the 30 amino acid peptide Trp38-OH was identified in 25 cases, where five of the peptides were also Asn49 hydroxylated and five were Asp47 hydroxylated. Both peptides were also found to be singly Asp47 and Asn49 hydroxylated or with both modications at the same time.

Out of the 70,755 peptides detected in our dataset a total of 40 peptides were derived from NAA10. The NAA10 protein group had a sequence coverage of 92%, indicating that the peptides cover most of its primary structure. The NAA10 peptides detected are listed in Table 5.2 with their position relative to the 235 amino acid NAA10 sequence, the respective MS/MS counts, score and the detected modifications. For the peptides with only one residue corresponding to the identified modification, the modified residues are shown as bold letters. For some peptides several residues corresponded to the given modifications, and we were unable to determine the exact modified residue among these.

**Table 5.2:** List of NAA10 peptides detected in the primary LC-MS/MS analysis

| Peptide sequence * | Position | MS/MS count | Score | Modifications |
|---|---|---|---|---|
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 12 | 176 | Unmodified |
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 9 | 154 | Met-O |
| NARPEDL**M**N**M**QHCNLLCLPENYQ**M**K | 5-29 | 5 | 136 | 3 Met-O |
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 7 | 122 | 2 Met-O, Asn-OH |
| NARPE**D**LMNMQHCNLLCLPENYQMK | 5-29 | 1 | 72 | Asn-OH, Asp-OH |
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 1 | 55 | 2 Asn-OH |
| NARPE**D**LMNMQHCNLLCLPENYQMK | 5-29 | 1 | 111 | Asp-OH |
| NARPEDL**M**NMQHCNLLCLPENYQ**M**K | 5-29 | 2 | 84 | 2 Met-O |
| NARPE**D**LMNMQHCNLLCLPENYQMK | 5-29 | 1 | 41 | 2 Asn-OH, Asp-OH |
| YYFYHGLSWPQLSYIAEDENGK | 30-51 | 37 | 166 | Unmodified |
| IVGYVLAK | 52-59 | 36 | 103 | Unmodified |
| MEEDPDDVPHGHITSLAVK | 60-78 | 49 | 293 | Unmodified |
| **M**EEDPDDVPHGHITSLAVK | 60-78 | 36 | 193 | Met-O |
| RLGLAQK | 83-89 | 3 | 98 | Unmodified |
| LMDQASR | 90-96 | 8 | 168 | Unmodified |
| A**M**IENFNAK | 97-105 | 15 | 142 | Met-O |
| AMIENFNAK | 97-105 | 10 | 142 | Unmodified |
| A**M**IENFNAKYVSLHVR | 97-112 | 1 | 51 | Met-O |
| YVSLHVR | 106-112 | 15 | 143 | Unmodified |
| AALHLYSNTLNFQISEVEPK | 117-136 | 42 | 225 | Unmodified |
| YYADGEDAYA**M**K | 137-148 | 25 | 110 | Met-O |
| YYADGEDAYAMK | 137-148 | 20 | 87 | Unmodified |
| YYADGEDAYAMKR | 137-149 | 8 | 191 | Unmodified |
| YYADGEDAYA**M**KR | 137-149 | 14 | 93 | Met-O |
| RDLTQMADELRR | 149-160 | 5 | 133 | Unmodified |
| RDLTQ**M**ADELRR | 149-160 | 1 | 54 | Met-O |
| DLTQ**M**ADELR | 150-159 | 21 | 147 | Met-O |
| DLTQMADELR | 150-159 | 10 | 143 | Unmodified |
| DLTQMADELRR | 150-160 | 22 | 140 | Unmodified |
| DLTQ**M**ADELRR | 150-160 | 28 | 155 | Met-O |
| GRHVVLGAIENK | 168-179 | 6 | 69 | Unmodified |
| HVVLGAIENK | 170-179 | 30 | 210 | Unmodified |
| HVVLGAIENKVESK | 170-183 | 16 | 204 | Unmodified |
| GNSPPSSGEACR | 184-195 | 1 | 108 | Unmodified |
| GNSPPSSGEACREEK | 184-198 | 5 | 136 | Unmodified |
| GLAAEDSGGDSK | 199-210 | 9 | 153 | Unmodified |
| GLAAEDSGGDSKDLSEVSETTESTDVK | 199-225 | 11 | 258 | Unmodified |
| GLAAEDSGGDSKDLSEVSETTESTDVKDSS EASDSAS | 199-235 | 4 | 118 | Unmodified |
| DLSEVSETTESTDVK | 211-225 | 31 | 212 | Unmodified |
| DLSEVSETTESTDVKDSSEASDSAS | 211-235 | 2 | 100 | Unmodified |

*For the peptides with only one residue corresponding to the identified modification for the given peptide the modified residues are shown as bold letters.

Because the actual abundance of NAA10 Trp38-OH is unknown it is difficult to give the absolute probability of not detecting this modification if it is an actual PTM that occurs in the cell. According to our dataset the abundancy is 0% compared to 26% found by Kang *et al*., but the probability of detecting the modification if their estimation of 26% abundancy is correct can be estimated by plotting the binomial distribution (Equation 1).

$$P(B=k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{Equation 1}$$

Equation 1 gives the probability P(B) of finding k instances with a probability of p in n trials. The null hypothesis, which is the initial guess about the abundance of hydroxylated Trp38, was set to 26% and the binomial distribution was calculated using Equation 1. The number of trials (n) was set to 37 (the number of MS/MS counts for the peptide across all the samples in the dataset), with the number of successes (k) varying from 0 to 37. The hypothesized probability for success (p) was set to 26%. The probabilities given for each number of MS/MS counts are shown in Figure 5.3. The highest probability is seen for detecting Trp38-OH 9 and 10 times, and there is a 96% chance of detecting it between 5 and 15 times given a relative abundance of 26%. The probability of not detecting any Trp38-OH is calculated to be 0.001%.

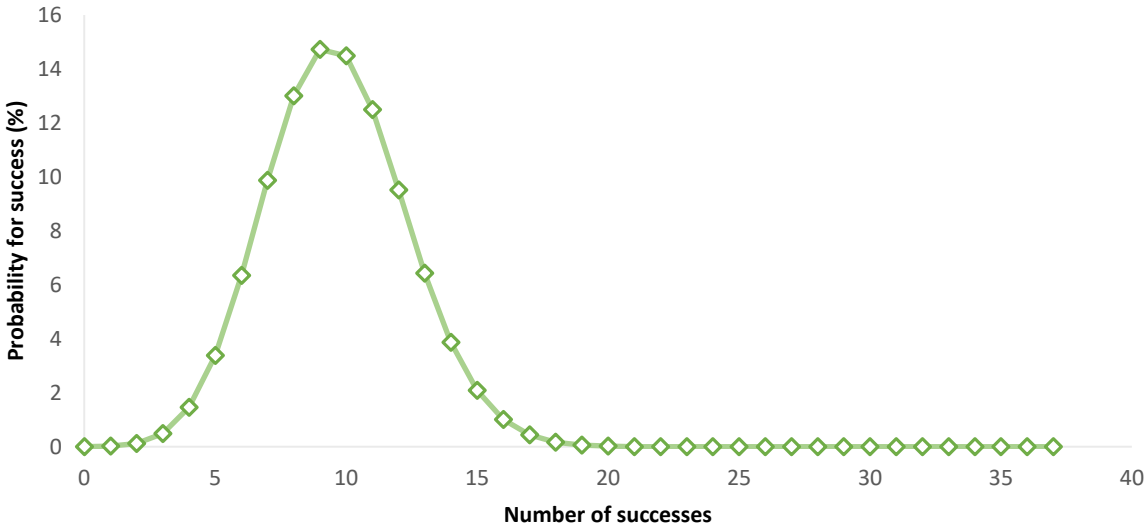


**Figure 5.3: The binomial distribution of the probability for detecting Trp38-OH given the null hypothesis assuming an abundancy of 26%.** The hypothesized probability for success (p) was set to 26%. The number of trials (n) describing the MS/MS count for the peptide was set to 37, and the number successes (k) describing the number of times the Trp38-OH is detected varied from 0 to 37. The highest probability is given for detecting 9 and 10 Trp38-OH, but the chance of not finding any was calculated to be 0.001%.

Given that the actual abundancy is lower than what Kang *et al.* reported, it was of interest to also determine the abundancy (p) for which the probability for being true is 95% (P(B=k)) given that no Trp38-OH events (k) were detected over 37 MS/MS events (n). As k is set to zero Equation 1 can be simplified and transformed to Equation 2 because both $\binom{37}{0}$ and $p^0$ equals 1. Based on no Trp38-OH events detected over 37 MS/MS events this gives that it is a 95% chance that 0.14% is the fraction of Trp38-peptides are actually hydroxylated.

$$p = 1-(P(B = k))^{\frac{1}{n}} \qquad \text{Equation 2}$$

A total of 40 NAA10, 60 NAA15, 9 NAA50 and 10 HYPK peptides were identified across all samples in the dataset. The LFQ intensity was measured for each protein group in each sample, and the percentile scores for the NatA components and FIH for each sample are color graded in Table 5.3. The numbers denote the percentage of protein groups with a lower LFQ intensity in that sample. We expected to see more NAA15, NAA50 and HYPK in the FASP samples compared to the in-gel samples where only the gel-region comprising NAA10 was further processed. The high percentile scores for NAA10 indicate a high amount of NAA10 in each sample compared to the rest of the proteins, supporting that the V5-IP was successful as plenty of NAA10 had been isolated. Successful coprecipitation of NAA15 was also seen, mostly for the FASP samples as expected but also for some of the in-gel samples that had been subject to NAA15-IP. NAA50 and HYPK coprecipitated for most of the FASP samples, and the percentile scores indicate that there are less of these components compared to NAA10 and NAA15. The samples showed no evidence of NAA10 interacting with FIH as no FIH-derived peptides were detected. Also, no peptides derived from HIF-1α were detected in this dataset.

**Table 5.3:** Percentile scores for LFQ intensities for the NatA components in each sample

| Sample | Transfection | IP | Method | LFQ intensity percentile score | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | NAA10 | NAA15 | NAA50 | HYPK | FIH |
| 1 | NAA15 WT-V5 | V5 | FASP | 82% | 90% | 22% | 61% | 0% |
| 2 | NAA15 T406Y-V5 | V5 | FASP | 82% | 90% | 0% | 64% | 0% |
| 3 | NAA15 L814P-V5 | V5 | FASP | 74% | 85% | 31% | 39% | 0% |
| 4 | NAA10-V5 | V5 | In-gel | 94% | 0% | 0% | 0% | 0% |
| 5 | NAA10-V5 | V5 | In-gel | 98% | 0% | 0% | 0% | 0% |
| 6 | Untransfected | NAA15 | In-gel | 89% | 3% | 0% | 0% | 0% |
| 7 | Untransfected | NAA15 | In-gel | 93% | 58% | 0% | 0% | 0% |
| 8 | NAA10-V5 | V5 | In-gel | 88% | 0% | 0% | 0% | 0% |
| 9 | NAA10-V5 | V5 | In-gel | 95% | 0% | 0% | 0% | 0% |
| 10 | Untransfected | NAA15 | In-gel | 87% | 33% | 0% | 0% | 0% |
| 11 | Untransfected | NAA15 | In-gel | 79% | 48% | 0% | 0% | 0% |
| 12 | NAA10-V5 | V5 | FASP | 85% | 86% | 49% | 50% | 0% |
| 13 | Untransfected | NAA15 | FASP | 89% | 98% | 28% | 82% | 0% |
| 14 | NAA10-V5 | V5 | FASP | 76% | 79% | 39% | 37% | 0% |
| 15 | Untransfected | NAA15 | FASP | 90% | 97% | 30% | 84% | 0% |
| 16 | NAA10-V5 | V5 | FASP | 89% | 3% | 0% | 0% | 0% |
| 17 | NAA10-V5 | V5 | FASP | 91% | 36% | 0% | 0% | 0% |

For the additional deep PTM analysis performed to increase the coverage the Top10 method was used, resulting in 3,437 detected peptides and 565 protein groups across the eight fractionated and one unfractionated sample. 30% of these were found to be modified with protein Nt-Ac (167 peptides), Asn-OH (153 peptides), Asp-OH (186 peptides), Trp-OH (104 peptides), Met-O (665 peptides) and/or Lys-Ac (62 peptides). A total of 86 peptides derived from NAA10 were detected (Table 5.4), and the sequence coverage of 94.5% indicated a very high coverage of the primary structure. 136 NAA15 peptides, 6 NAA50 peptides and 20 HYPK peptides were detected across the samples, but there was no sign for interactions between NAA10 and FIH or HIF-1α. Both the 22 and 30 amino acid peptides containing Trp38 were detected in the deep PTM analysis with MS/MS counts of 3 and 2 respectively, but both were found to be unmodified. In addition to Met-O, Asn-OH, Asp-OH, Trp-OH and protein Nt-Ac it was searched for Lys-Ac to look for the reported autoacetylation of NAA10 at Lys136. Only one acetylated lysine was detected for the peptide MEEDPDDVPHGHITSLAVK at position 60-78 relative to the NAA10 protein sequence.

**Table 5.4:** List of NAA10 peptides detected in the deep PTM analysis

| Peptide sequence * | Position | MS/MS count | Score | Modifications |
|---|---|---|---|---|
| **M**NIRNARPEDLMNMQHCNLLCLPENYQM**K** | 1-29 | 1 | 127 | 3 Met-O, Nt-Ac, Asp-OH |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 1 | 126 | 2 Asn-OH, Nt-Ac, Met-O |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 1 | 124 | 2 Met-O, Nt-Ac, Asn-OH |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 1 | 165 | Nt-Ac, Asn-OH, Met-O |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 2 | 194 | 2 Met-O, Nt-Ac |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 1 | 101 | 3 Met-O, Nt-Ac, Asn-OH |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 1 | 93 | 3 Met-O, Nt-Ac |
| **M**NIRNARPE**D**LMNMQHCNLLCLPENYQMK | 1-29 | 1 | 110 | 2 Met-O, Nt-Ac, Asp-OH |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 2 | 122 | 2 Asn-OH, Nt-Ac |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 1 | 60 | Nt-Ac, Asn-OH |
| **M**NIRNARPEDLMNMQHCNLLCLPENYQMK | 1-29 | 1 | 118 | 2 Met-O, Nt-Ac, Asn-OH, Asp-OH |
| NARPE**D**LMNMQHCNLLCLPENYQMK | 5-29 | 6 | 176 | 2 Met-O, Asn-OH |
| NARPEDL**M**N**M**QHCNLLCLPENYQ**M**K | 5-29 | 2 | 158 | 3 Met-O |
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 12 | 203 | 2 Met-O |
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 2 | 199 | Asn-OH, Met-O |
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 9 | 210 | Met-O |
| NARPEDLMNMQHCNLLCLPENYQMK | 5-29 | 4 | 272 | Unmodified |
| NARPE**D**LMNMQHCNLLCLPENYQMK | 5-29 | 1 | 67 | Asn-OH, Asp-OH |
| NARPE**D**LMNMQHCNLLCLPENYQMK | 5-29 | 0 | 40 | 2 Asn-OH, Asp-OH |
| NARPE**D**LMNMQHCNLLCLPENYQMK | 5-29 | 1 | 95 | Asp-OH, Met-O |
| YYFYHGLSWPQLSYIAEDENGK | 30-51 | 3 | 193 | Unmodified |
| YYFYHGLSWPQLSYIAEDENGKIVGYVLAK | 30-59 | 2 | 211 | Unmodified |
| IVGYVLAK | 52-59 | 3 | 108 | Unmodified |
| IVGYVLAK**M**EEDPDDVPHGHITSLAVKR | 52-79 | 2 | 164 | Met-O |
| **M**EEDPDDVPHGHITSLAVK | 60-78 | 16 | 312 | Met-O |
| MEEDPDDVPHGHITSLAVK | 60-78 | 8 | 312 | Unmodified |
| **M**EEDPDDVPHGHITSLAVKR | 60-79 | 24 | 363 | Met-O |
| MEEDPDDVPHGHITSLAVKR | 60-79 | 15 | 380 | Unmodified |
| **M**EEDPDDVPHGHITSLAV**K**R | 60-79 | 1 | 93 | Ac-Lys, Met-O |
| MEEDPDDVPHGHITSLAVKR | 60-79 | 1 | 87 | Asp-OH |
| RLGLAQK | 83-89 | 5 | 136 | Unmodified |
| RLGLAQKL**M**DQASR | 83-96 | 4 | 86 | Met-O |
| LGLAQKL**M**DQASR | 84-96 | 9 | 164 | Met-O |
| LGLAQKL**MD**QASR | 84-96 | 3 | 89 | Asp-OH, Met-O |
| LGLAQKLMDQASR | 84-96 | 8 | 289 | Unmodified |
| LGLAQKLM**D**QASR | 84-96 | 1 | 79 | Asp-OH |
| L**MD**QASR | 90-96 | 1 | 111 | Asp-OH, Met-O |
| L**M**DQASR | 90-96 | 5 | 134 | Met-O |
| LMDQASR | 90-96 | 1 | 134 | Unmodified |
| A**M**IENFNAK | 97-105 | 12 | 141 | Met-O |
| AMIENFNAK | 97-105 | 3 | 122 | Unmodified |
| A**M**IENFNAKYVSLHVR | 97-112 | 8 | 136 | Met-O |
| AMIENFNAKYVSLHVR | 97-112 | 3 | 98 | Unmodified |
| A**M**IENFNAKYVSLHVRK | 97-113 | 3 | 104 | Met-O |

| | | | | |
|---|---|---|---|---|
| YVSLHVR | 106-112 | 6 | 142 | Unmodified |
| YVSLHVRK | 106-113 | 4 | 173 | Unmodified |
| KSNRAALHLYSNTLNFQISEVEPK | 113-136 | 3 | 153 | Unmodified |
| SNRAALHLYSNTLNFQISEVEPK | 114-136 | 8 | 214 | Unmodified |
| AALHLYSNTLNFQISEVEPK | 117-136 | 11 | 433 | Unmodified |
| AALHLYSNTL**N**FQISEVEPK | 117-136 | 0 | 42 | Asn-OH |
| AALHLYSNTLNFQISEVEPKYYADGEDAYA **M**K | 117-148 | 5 | 190 | Met-O |
| AALHLYSNTLNFQISEVEPKYYADGEDAYA **M**KR | 117-149 | 8 | 235 | Met-O |
| AALHLYSNTLNFQISEVEPKYYADGEDAYA MKR | 117-149 | 5 | 289 | Unmodified |
| AALHLYSNTLNFQISEVEPKYYADGEDAYA MKR | 117-149 | 1 | 131 | Asp-OH |
| YYADGEDAYA**M**K | 137-148 | 2 | 125 | Met-O |
| YYADGEDAYAMK | 137-148 | 1 | 93 | Unmodified |
| YYADGEDAYA**M**KR | 137-149 | 1 | 61 | Asp-OH, Met-O |
| YYADGEDAYA**M**KR | 137-149 | 8 | 98 | Met-O |
| YYADGEDAYAMKR | 137-149 | 8 | 294 | Unmodified |
| YYADGEDAYAMKRDLTQMADELR | 137-159 | 1 | 69 | Asp-OH, Met-O |
| RDLTQ**M**ADELRR | 149-160 | 4 | 93 | Met-O |
| RDLTQMADELRR | 149-160 | 2 | 182 | Unmodified |
| DLTQ**M**ADELR | 150-159 | 5 | 150 | Met-O |
| DLTQ**M**ADELR | 150-159 | 1 | 119 | Asp-OH, Met-O |
| DLTQMADELR | 150-159 | 3 | 168 | Unmodified |
| DLTQ**M**ADELRR | 150-160 | 17 | 185 | Met-O |
| DLTQ**M**ADELRR | 150-160 | 6 | 152 | Asp-OH, Met-O |
| DLTQMADELRR | 150-160 | 8 | 266 | Unmodified |
| DLTQ**M**ADELRRHLELK | 150-165 | 5 | 138 | Met-O |
| DLTQMADELRRHLELK | 150-165 | 2 | 83 | Unmodified |
| RHLELKEK | 160-167 | 3 | 186 | Unmodified |
| HLELKEK | 161-167 | 1 | 136 | Unmodified |
| HLELKEKGR | 161-169 | 7 | 155 | Unmodified |
| EKGRHVVLGAIENK | 166-179 | 3 | 155 | Unmodified |
| GRHVVLGAIENK | 168-179 | 10 | 222 | Unmodified |
| GRHVVLGAIENKVESK | 168-183 | 20 | 302 | Unmodified |
| HVVLGAIENK | 170-179 | 7 | 220 | Unmodified |
| HVVLGAIENKVESK | 170-183 | 13 | 240 | Unmodified |
| HVVLGAIENKVESKGNSPPSSGEACR | 170-195 | 1 | 223 | Unmodified |
| GNSPPSSGEACR | 184-195 | 2 | 118 | Unmodified |
| GNSPPSSGEACREEK | 184-198 | 4 | 105 | Unmodified |
| GNSPPSSGEACREEKGLAAEDSGGDSK | 184-210 | 5 | 274 | Unmodified |
| EEKGLAAEDSGGDSK | 196-210 | 1 | 197 | Unmodified |
| GLAAEDSGGDSK | 199-210 | 5 | 150 | Unmodified |
| GLAAEDSGGDSKDLSEVSETTESTDVK | 199-225 | 3 | 294 | Unmodified |
| DLSEVSETTESTDVK | 211-225 | 4 | 222 | Unmodified |

*For the peptides with only one residue corresponding to the identified modification for the given peptide the modified residues are shown as bold letters.

## 5.2 WESTERN BLOT ANALYSIS

### 5.2.1 Does NAA10 physically interact with FIH?

Kang *et al*. performed *in vitro* binding analyses and reported interactions between both endogenous and ectopically expressed NAA10 and FIH. Although the mass spectrometry data did not reveal any evidence of binding between NAA10 and FIH this was further investigated by WB analysis as described in Methods section 4.4.4. HEK293 cells were transfected with Xpress-NAA10 and/or FIH-V5 plasmids and grown for 48 h before they were lysed. The lysates were subject to IP with anti-Xpress and anti-V5 and subsequent WB analysis with anti-NAA10, anti-Xpress and anti-V5 to see whether the proteins coprecipitated with each other (Figure 5.4).

The molecular masses of endogenous NAA10, Xpress-NAA10 and FIH-V5 are 26, 27 and 40 kDa respectively. Still the signal from endogenous NAA10 (red asterisk) was seen at 35 kDa, the signal from Xpress-NAA10 (green asterisk) around 39 kDa, and the signal from FIH-V5 was seen around 50 kDa (yellow asterisk) in Figure 5.4. This discrepancy was seen for all blots for unknown reasons, but detection of both Xpress-NAA10 and FIH-V5 confirmed successful IP's. Coprecipitation of Xpress-NAA10 with FIH-V5 was seen for the double transfected cells when using anti-NAA10, but it was not detected with anti-Xpress. An unspecific band of the same size as Xpress-NAA10 is also seen for the Xpress-NAA10 transfected cells, which could be due to background binding or contamination between the wells when loading the gel. No coprecipitation of FIH-V5 was detected with anti-V5 in the Xpress-IP samples.

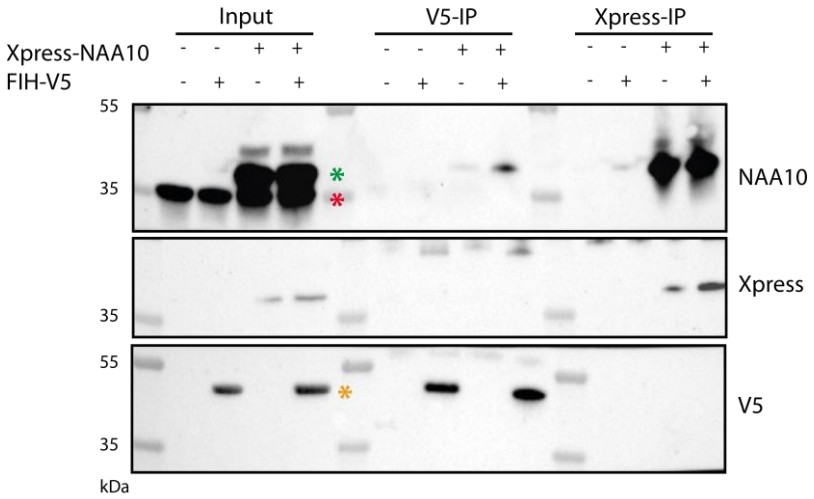


**Figure 5.4: Study of interactions between Xpress-NAA10 and FIH-V5 in HEK293 cells.** HEK293 cells were transfected with Xpress-NAA10 and/or FIH-V5 (4 µg of each plasmid per 10 cm dish) and lysed 48 h post-transfection. The cell lysates were subject to IP using Xpress- and V5-antibodies. Control cells were transfected with the same amount of the empty-V5 vector. The blots were incubated with anti-Xpress, anti-NAA10 and anti-V5 for WB analysis, and the band sizes are indicated with a green, red and yellow asterisk respectively. Coprecipitation of Xpress-NAA10 (27 kDa) with FIH-V5 was detected with anti-NAA10 (top panel) but not with anti-Xpress (middle panel). No coprecipitation of FIH-V5 (40 kDa) with Xpress-NAA10 was detected with anti-V5 (bottom panel).

To look for interactions between endogenous NAA10 and FIH, HAP1 WT and FIH KO cells were lysed 24 h post-seeding and subject to IP using anti-FIH and anti-NAA10 antibodies (Figure 5.5). WB analysis revealed that FIH was only detectable in the WT cells, thereby confirming the FIH KO. Coprecipitation was not detected for neither NAA10 (26 kDa) nor FIH (40 kDa) in the FIH-IP and NAA10-IP respectively, revealing no evidence of physical interactions between the proteins.

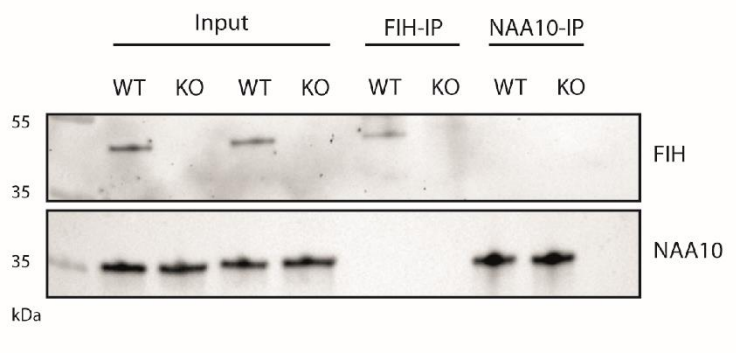


**Figure 5.5: Study of interactions between endogenous NAA10 and FIH in HAP1 WT and FIH KO cells.** HAP1 WT and FIH KO cells were harvested and lysed 24 h post-seeding, and endogenous proteins were isolated by FIH- and NAA10-IP. WB analysis against anti-FIH and anti-NAA10 was carried out, but no coprecipitation was detected for neither FIH (40 kDa) nor NAA10 (26 kDa) indicating no interaction between the endogenous proteins.

HEK293 cells were also transfected with HA-HIF-1α in order to look for interactions with FIH and/or NAA10. None of the available antibodies worked properly for this purpose, and we did not manage to isolate HA-HIF-1α with anti-HA or anti-HIF-1α.

### 5.2.2 Does NAA10 expression differ in presence and absence of FIH?

To investigate whether FIH has an impact on NAA10 protein expression, lysates from FIH-V5 and Xpress-NAA10 transfected cells as well as untransfected HAP1 WT and FIH KO cells were analyzed by WB (Figure 5.6). WB analysis against anti-Xpress with lysates from transfected HEK293 cells is shown in Panel A. In four independent experiments a 19, 36, 46 and 65% increase in Xpress signal was seen in presence of FIH-V5 (Panel A). In three of the same four experiments the signal for FIH-V5 also increased with 35-40% in presence of Xpress-NAA10, and in one experiment no difference was detected. The increase in signal strength for both proteins in double transfected cells could suggest that Xpress-NAA10 and FIH-V5 may be directly or indirectly stabilized by each other. However, when using anti-NAA10 the Xpress-NAA10 signal only showed an increase of 5-17% for double transfected cells compared to the 19-65% increase observed when using anti-Xpress. The reasons for this difference remain unclear. In the case for endogenous NAA10 the signal showed an increase similar to Xpress-NAA10 ranging from 5-11%.

HAP1 WT and FIH KO cells were seeded and lysed in triplicates in order to look for differences in NAA10 and NAA15 expression in presence and absence of FIH. The lysates were analyzed by WB against anti-FIH, anti-NAA10 and anti-NAA15 as shown in Panel B, and β-tubulin was used as a loading control. When considering the loading control there is no detectable difference in NAA10 and NAA15 signals between WT and FIH KO cells. In this case FIH does not seem to have any detectable impact on the amount of NatA in the cells.

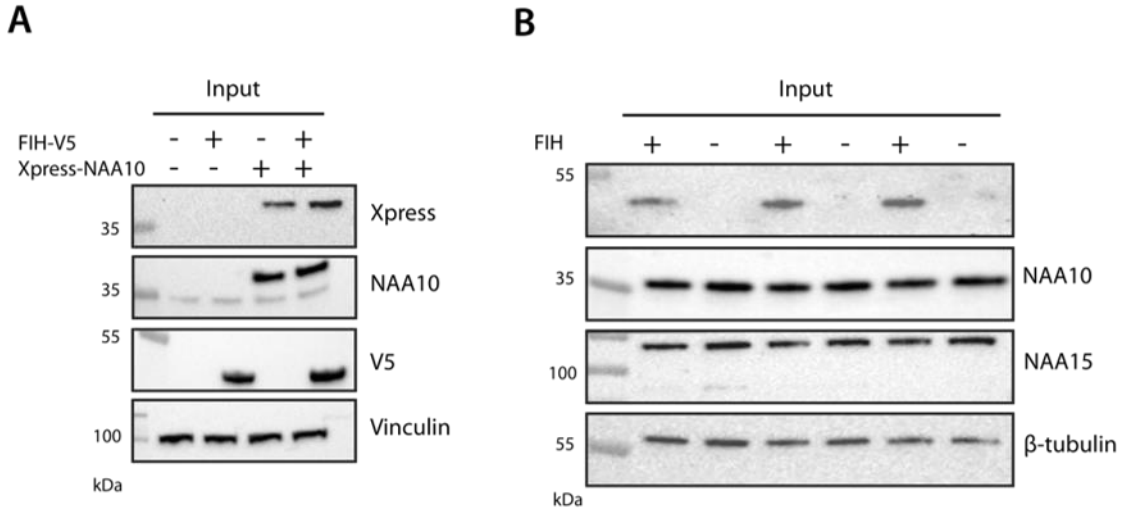


**Figure 5.6: Study of NAA10 and NAA15 expression in presence and absence of FIH.** (A) Lysate from FIH-V5 and Xpress-NAA10 transfected HEK293 cells were subject to WB analysis against anti-Xpress and anti-V5 to investigate the difference in NAA10 expression. The blot revealed a 65% increase in Xpress-NAA10 signal and 35% increase in FIH-V5 signal for double transfected cells. An increase of only 5-17% was seen for both Xpress-NAA10 and endogenous NAA10 in the presence of FIH when using anti-NAA10. (B) HAP1 WT and FIH KO cell lysates were analyzed in triplicates by WB against anti-NAA10, anti-NAA15 and anti-FIH to study differences in NAA10 and NAA15 expression. No difference was detected between HAP1 WT and FIH KO cells when comparing the signal to β-tubulin expression, indicating that FIH has no detectable impact on the amount of NatA in the HAP1 cells.

## 5.3 *IN VITRO* ACETYLATION ASSAY

### 5.3.1 Does cellular NAA10 activity differ in presence and absence of FIH?

In order to assess whether the presence of FIH impacts the cellular NatA or NAA10 catalytic activity, an activity analysis was performed in three independent experiments. Upon NatA complex formation the substrate specificity of NAA10 shifts towards N-termini starting with Ala, Ser, Thr, Gly, Val and Cys (Liszczak *et al*., 2013). HAP1 WT and FIH KO cells were harvested and lysed 24 h after seeding, and the lysates were subject to NAA10-IP with anti-NAA10 and magnetic DynaBeads™ as described in Methods section 4.4.1. The immunoprecipitated NAA10 was further used in $^{14}$C-Ac-CoA-based *in vitro* acetylation assays with the peptide substrates SESS, EEEI and histone 4 as described in Methods section 4.4.2, and the NAT activity towards SESS and EEEI and potential KAT activity towards histone 4 was measured (Figure 5.7). The lysates, IP samples and unbound-samples were further analyzed by SDS-PAGE and WB with anti-NAA10, anti-NAA15 and anti-FIH for quantification of NAA10 and NAA15 (Panel A). Here, quantification of NAA15 is equivalent to the amount of NatA complexes, whereas NAA10 both exists in monomeric form and complexed with NAA15. The measured catalytic activity towards the NatA substrate SESS was normalized to NAA15, and the catalytic activity towards the monomeric NAA10 substrate EEEI and the potential NAA10 KAT substrate histone 4 was normalized to NAA10 (Panel B).

The WB analysis in Panel A revealed slightly higher signal for NAA10 and NAA15 in the FIH KO cell lysate compared to the WT HAP1 cells that might be due to loading differences. Equal amounts of both proteins were immunoprecipitated between WT and FIH KO cells, and anti-FIH was used to confirm the presence and absence of FIH in the cell lines. For all three experiments the measured catalytic activity for NatA towards SESS appeared to be similar in presence and absence of FIH (Panel B). NAA10 from WT cells was found to have a slightly increased activity towards EEEI compared to NAA10 from FIH KO cells in two of the experiments, but for the third one similar activity towards EEEI for NAA10 from both cell types was seen. In all three experiments a higher activity was seen for the control towards histone 4 compared to SESS and EEEI. This may be caused by background binding or autoacetylation as no enzyme is expected to be isolated during the control-IP. A higher activity towards histone 4 was also seen for NAA10 from WT cells compared to the FIH KO cells in two of the experiments, whereas in one case it was found to be higher for NAA10 from FIH KO cells. Altogether, the data suggests that FIH has no detectable impact on the NatA activity, but when it comes to the monomeric NAA10 activity towards EEEI and histone 4 our results are inconclusive.

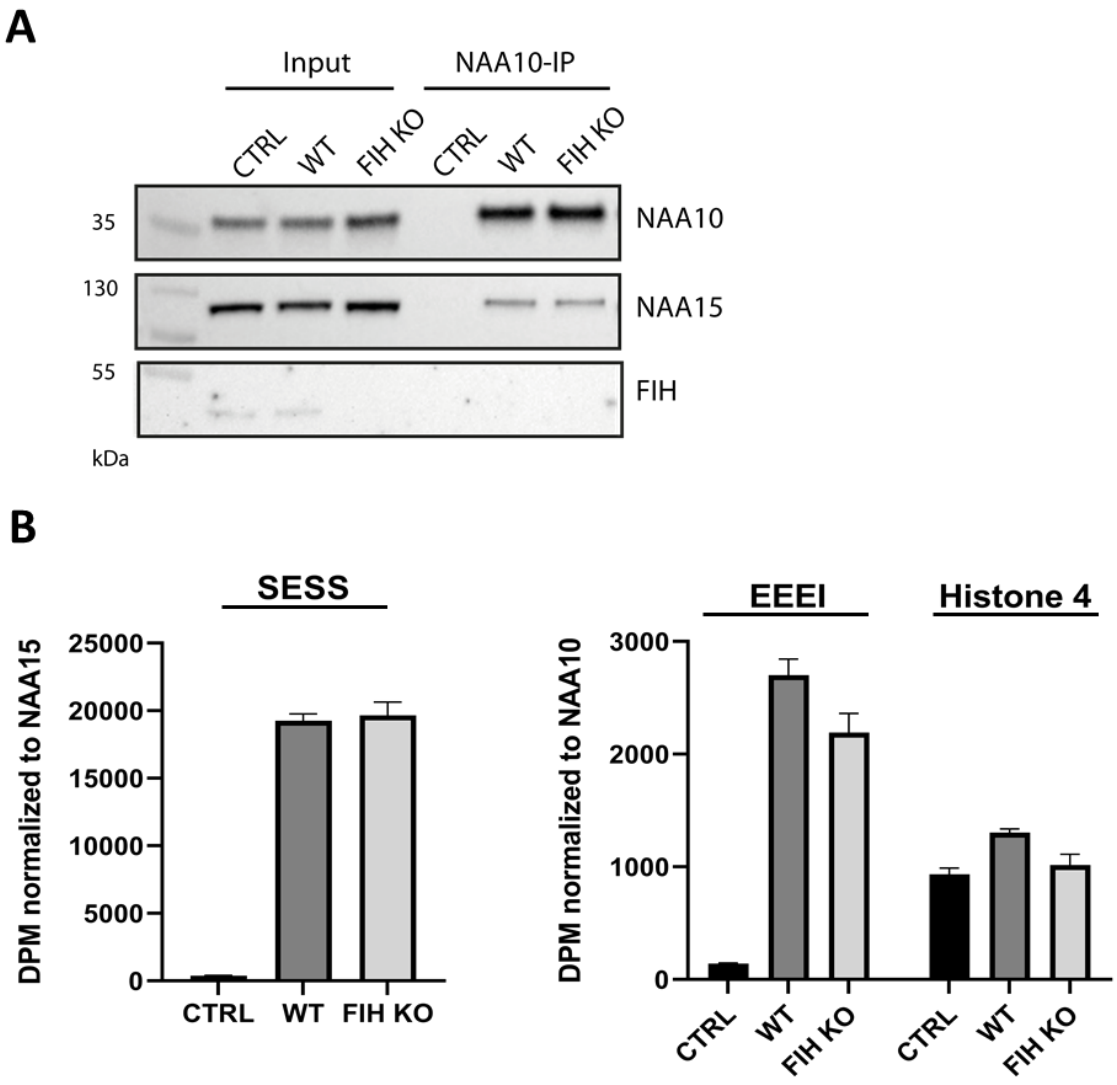


**Figure 5.7: $^{14}$C-acetylation assay measuring NatA and monomeric NAA10 catalytic activity.** (A) WB analysis of NAA10 immunoprecipitated from HAP1 WT and FIH KO cells using anti-NAA10 antibody. Co-IP of NAA15 was detected with anti-NAA15, and anti-FIH was used to confirm presence and absence of FIH. (B) NatA and monomeric NAA10 activity measurement from $^{14}$C-acetylation assay of immunoprecipitated NAA10 and coprecipitated NAA15. The activity towards SESS was normalized to NAA15, and the activity towards EEEI and histone 4 was normalized to NAA10 to show relative product formation per enzyme.

# 6 DISCUSSION

As much as 80% of the human proteome is Nt-acetylated, affecting protein half-life, protein-protein interactions, subcellular targeting, folding and aggregation (Hwang, Shemorry and Varshavsky, 2010; Shemorry *et al*., 2013; Monda *et al*., 2013; Behnia *et al*., 2004; Forte *et al*., 2011; Setty *et al*., 2004; Bartels *et al*., 2011; Holmes *et al*., 2014; Trexler *et al*., 2012). The NATs also have important roles in development, stress signaling and metabolism as well as in tumorigenesis (Rope *et al*., 2011; Myklebust, Støve and Arnesen, 2015; Seo *et al*., 2016; Shin *et al*., 2014; Aksnes, Ree and Arnesen, 2019; Fluge *et al*., 2002; Line *et al*., 2002; Martin *et al*., 2007; Ren *et al*., 2008; Yu *et al*., 2009a; Lee *et al*., 2010; Yu *et al*., 2009b; Midorikawa *et al*., 2002). Monomeric NAA10 has, in addition to perform Nt-Ac, been found to acetylate internal Lys residues of a variety of substrates (Lim *et al*., 2006; Shin *et al*., 2009 and 2014; Wang *et al*., 2012; Yoon *et al*., 2014; Seo *et al*., 2016; Lee *et al*., 2017; Qian *et al*., 2017; Lu Vo *et al*., 2017). One of the reported substrates is HIF-1α, where NAA10-mediated acetylation of Lys532 was found to play a central role in HIF-1α stability (Jeong *et al*., 2002). This was mainly seen under normoxic conditions as both NAA10 mRNA levels and its affinity to HIF-1α was found to decrease under hypoxia. This theory is still under debate as the results were found to not be reproducible *in vitro*, and conflicting findings have been reported by several research groups (Arnesen *et al*., 2005b; Murray-Rust *et al*., 2006). Structural and biochemical evidence also speak against the ability of NAA10 to act on internal Lys residues (Liszczak *et al*., 2013; Gottlieb and Marmorstein, 2018). Seo *et al*. have tried to explain these conflicting findings with autoacetylation, stating that only autoacetylated NAA10 displays KAT activity (Seo *et al*., 2010 and 2015). Recently, Kang *et al*. suggested that hydroxylation of NAA10 Trp38 by FIH induces a conformational change to the NAA10 structure, potentially enabling it to display KAT-activity by widening the substrate gate (Kang *et al*., 2018). NAA10-mediated Lys-Ac of several KAT substrates is also reported to play a role in human disease and cancer development (Lee *et al*., 2017; Lu Vo *et al*., 2017; Qian *et al*., 2017; Lim *et al*, 2006), emphasizing the importance of further research on potential moonlighting functions of NAA10.

The aim of this project was to look further into the reported KAT activity of NAA10, and to investigate its relationship to the proposed NAA10 hydroxylase, FIH. In order to do so, the NAA10 hydroxylation status was investigated by isolating monomeric and NAA15-complexed NAA10 from NAA10-V5 transfected and untransfected HEK293 and HeLa cells by IP. It was also assessed whether NAA10 and FIH physically interacts by isolating FIH and NAA10 by IP from FIH-V5 and Xpress-NAA10 transfected HEK293 cells, and subsequently perform WB analysis against anti-FIH-V5 and Xpress-NAA10. The lysates from these experiments as well as untransfected HAP1 WT and FIH KO lysates were further analyzed by WB to investigate whether Xpress-NAA10 or NAA10 expression differed in presence and absence of FIH. Furthermore, HAP1 WT and FIH KO cells were subject to NAA10-IP to assess the activity of NAA10 and NatA in presence and absence of FIH by $^{14}$C-Ac-CoA-based *in vitro* acetylation assays. The classical NatA substrate SESS and the acidic EEEI peptide to which monomeric NAA10 displays catalytic activity were used to measure NAT activity. A Nt-acetylated histone 4 peptide, which is known to be a KAT substrate of other enzymes, was also included as a potential but not reported KAT substrate of NAA10 to see if Lys-Ac could be detected.

## 6.1 STUDY OF NAA10 AND FIH RELATIONSHIP

### 6.1.1 LC-MS/MS analysis of NAA10 hydroxylation status

Today mass spectrometry is a widely used tool for large-scale protein analysis, and it has rapidly improved over the last decade with respect to resolution, mass accuracy, sensitivity, and scan rate of mass spectrometers. The method has been widely used by researchers for different purposes, which include but are not limited to investigation of protein-protein interactions, protein quantification, protein modification and proteome profiling (Zhang *et al*., 2013). An alternative to MS in detection of PTMs is the use of antibodies such as anti-hydroxyproline or anti-acetyl lysine, but due to a very high sensitivity MS may still be the preferred method in many cases. Despite the advantages it has compared to other methods, mass spectrometry also has its limitations that are important to consider. The identification of protein modifications with a high degree of certainty relies on high resolution MS, a criterion which is met in this study by using a Q-Exactive HF mass spectrophotometer for the primary MS analysis as well as a Q-Exactive HF X mass spectrometer for the deep PTM analysis. The stoichiometry of protein modifications may also be a limitation as some modifications, such as Lys-Ac, may only occur for a small fraction of a specific protein. Implementation of enrichment steps during sample preparation may result in improved detection of such modifications. In this case an anti-Ac-Lys antibody can be used for IP prior to LC-MS/MS analysis in order to increase the probability of detecting the Lys acetylated peptides. Another important consideration is the quality of the chromatography in DDA LC-MS, which is important to ensure detection of all the relevant peptides.

Mass spectrometry was used in this study to investigate the hydroxylation status of monomeric and complexed NAA10. Top12 and Top10 methods were used for the primary LC-MS/MS analysis and the deep PTM analysis respectively to select ions for MS/MS. A total of 17 samples (Table 5.1) from HEK293, HeLa and HAP1 cells subjected to V5- and NAA15-IP were further prepared for LC-MS/MS analysis through FASP or in-gel methods. The raw data was analyzed using MaxQuant where it was searched for peptides of maximum 4600 Da and minimum 7 amino acids. LC-MS/MS data published by Kang *et al*. revealed that the reported peptide (Figure 2.6 Panel A) was found to be both Trp38-, Asp47- and Asn49-hydroxylated with a total MS/MS count of 106. As 28 Trp-OH events were detected they reported an abundancy of 26%. By further inspecting their data it was found that this peptide was only detected in 9 of the 106 cases, whereas a longer peptide (YYFYHGLSWPQLSYIAEDENGK**IVGYVLAK**), resulting from a missed cleavage during trypsin digestion, was detected in the remaining 97 cases (Supplementary Table 8.1). Most of the reported hydroxylation events were seen for this long peptide, and not for the 22 amino acid peptide presented in the paper (Kang *et al*., 2018). Since Kang *et al*. found Trp-OH, Asn-OH and Asp-OH for the Trp38-peptide, these were set as variable modifications in addition to protein Nt-Ac and Met-O in our search. For the deep PTM analysis Lys-Ac was also added as a variable modification to look for acetylation at the reported autoacetylation site of NAA10 (Lys136). IAA treatment during sample preparation for LC-MS/MS caused all cysteine residues to be alkylated, and cysteine carbamidomethylation was set as a fixed modification adding additional 57 Da to each cysteine residue (160 kDa).

The data from the primary LC-MS/MS analysis was further interpreted using Perseus (v. 1.6.5.0) and a total of 70,755 peptides and 5,612 protein groups were detected after excluding reverse peptides and potential contaminants. Approximately 2.8% of these peptides were hydroxylated Trp, Asn and/or Asp residues, but since the abundancy of other modifications in this dataset is unknown it is difficult to say whether this is a small or a large fraction of the peptides. Generally, a lot of Met-O was detected in this search applying for approximately 12% of the peptides identified. A deep PTM analysis was also performed in order to increase the coverage, where a total of 3,437 peptides and 565 protein groups were detected. For these samples 13% of the peptides identified were hydroxylated on Asn, Asp and/or Trp residues, and 19% were found to be modified with Met-O.

Protein hydroxylation is a PTM which is known to be introduced to proteins by cellular enzymes and reagents (Shacter, 2000). Tryptophan is usually hydroxylated by the tryptophan hydroxylase enzyme, which is one of three enzymes in the aromatic amino acid hydroxylase family. During the first step in the biosynthesis of serotonin and the catecholamine neurotransmitters this enzyme catalyzes Trp hydroxylation at position C-5 forming 5-hydroxytryptophan (Figure 6.1 Panel A) (Roberts and Fitzpatrick, 2013). So far FIH is only known to hydroxylate Asp, Asn (Yang *et al*., 2011a) and His residues (Yang *et al*., 2011b), but has not been reported to act on Trp. According to Kang *et al*. the FIH-directed hydroxylation of Trp38 of monomeric NAA10 occurs on C-2 of the indole ring to form 2-hydroxytryptophan (Figure 6.1 Panel B), but the reason for this assumption is unknown as this is not described in their paper. However, hydroxylation of protein residues has previously also been found to be an artefact of the sample preparation prior to LC-MS/MS analysis such as during in-gel digestion method (Klont *et al*., 2018) as well as the electrospray ionization (Boys *et al*., 2009). This has been shown to be more prevalent during in-gel digestion than in-solution digestion (Klont *et al*., 2018), and cysteine and methionine residues are most prone to such oxidative modifications (Shacter, 2000). Singly hydroxylated Trp, or hydroxytryptophan, has also been reported to be an artefact of in-gel digestion (Froelich and Reid, 2008) exhibiting a mass shift of 16 Da similar to the shift reported by Kang *et al*. This might explain the high occurrence of Met-O in both our analyses, as well as the high Trp-OH abundance reported by Kang *et al.* Such artificial oxidation events can be reduced by degassing of buffer prior to SDS-PAGE (Kim *et al*., 2015).

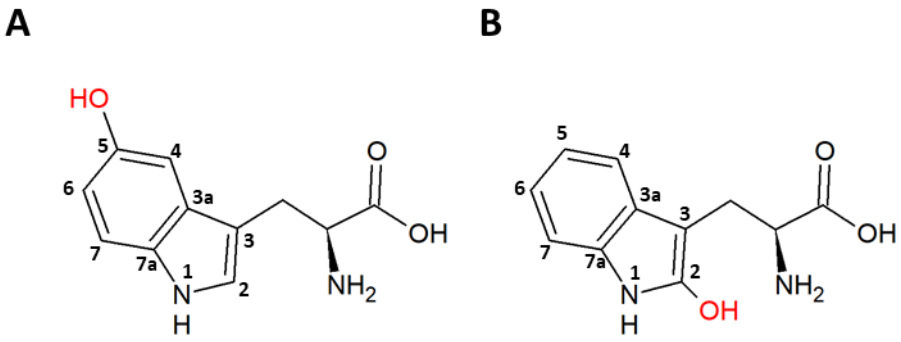


**Figure 6.1: Hydroxylation of Tryptophan.** Hydroxylation of amino acid residues has been described as a result of enzymatic action but also as an artefact of sample preparation for LC-MS/MS analysis. A) Tryptophan hydroxylase is an enzyme responsible for the formation of 5-hydroxytryptophan by introducing a hydroxyl group at position C-5 (red). B) Kang *et al.* claim that FIH under normoxic conditions is responsible for hydroxylating Trp38 of monomeric NAA10 on C-2 of the indole ring, thereby forming 2-hydroxytryptophan.

The 40 NAA10-derived peptides identified in the primary LC-MS/MS analysis are listed in Table 5.2. None of the peptides were Trp hydroxylated, but 36% of the peptides were Met oxidized. Asn-OH and Asp-OH were only seen for one NAA10-derived peptide of region 5-29 relative to the 235 amino acid NAA10 protein sequence. The reported 22 amino acid peptide containing Trp38 was detected in 11 of the 17 samples with a total MS/MS count of 37, but it was not found to be modified in any way (Figure 5.1). The peptide of 30 amino acids was not detected in this dataset most likely due to the use of a mix of trypsin and LysC for most samples, making it more unlikely to miss a cleavage after a Lys residue. In the deep PTM analysis 86 NAA10-derived peptides were identified (Table 5.4). 45% were found to be Met oxidized, 25% Asp hydroxylated and 7% being Asn hydroxylated, and a lot of the modifications were found to be associated with the same peptide from region 1-29 of the NAA10 protein sequence. Both the 22 and 30 amino acid peptides were detected in the deep PTM analysis with MS/MS counts of 3 and 2 respectively, but neither of them were found to be modified. Compared to the rest of the peptides found in the primary search, 37 MS/MS counts was found to be relatively high as only 2.7% of the other peptides had higher MS/MS counts (Figure 5.2). As we enriched for NAA10 by doing V5-IP this is expected and the high number indicates that the IP has worked properly. This is also reflected in the high sequence coverage of 92% across the NAA10 peptides (95% coverage for the deep PTM analysis), as well as the LFQ intensity percentile scores for NAA10 shown in Table 5.3. The LFQ intensity is a measure of protein abundance based on raw peptide intensities (Cox *et al.*, 2014), and for all samples at least 73% of the protein groups had lower LFQ intensities than NAA10.

The LFQ intensity percentile scores also reveal successful co-IP of the NatA components. NAA15, NAA50 and HYPK were all present in the FASP-samples, but not in the in-gel samples which was expected due to the specific exclusion of the gel regions above 55 kDa and below 25 kDa during the in-gel digestion. A few exceptions were seen for sample 7, 10 and 11 subjected to NAA15-IP where NAA15 (93 kDa) was detected to a higher degree compared to the other in-gel samples. One reason for this may be contamination during sample preparation. Another possible explanation may be that the NAA15 peptides detected are degradation products of sizes ranging from 25 to 55 kDa. If the protease inhibitor is not working, added in a low concentration or if the lysate has not been kept cold the proteases might become active after lysis and start degrading proteins. Because we enriched for NAA15 in these samples this may be the best target for the proteases, which can result in degradation products that will be detected in the LC-MS/MS analysis. As NAA15 was only present in the NAA15-IP in-gel samples this may be a more reasonable explanation than contamination. Based on the identification of peptides derived from the other NatA components it is likely to think that we would also detect other interactor proteins of NAA10. Among the peptides identified in this search none of them were HIF-1α specific, which is expected due to its short half-life under normoxic conditions. FIH specific peptides were also not identified during this search, which may either indicate that FIH does not bind to NAA10 or that the interaction is transient.

Based on the 28 cases of Trp38-OH over the 106 peptide identifications Kang *et al*. claim a 26% abundancy of Trp38-OH. This way of measuring peptide abundancy is referred to as spectral counting and requires that the peptides compared are chemically identical, which is not the case for the unmodified and hydroxylated Trp38-peptides. Their ionization efficiency may differ, and methods such as selected reaction monitoring (SRM) and SILAC should be used in this case to obtain reliable data. These methods are based on stable isotope labelling of specific amino acids on peptides and allow for direct comparison of chemically identical peptides within samples. As the peptides of one sample have different masses, they can be distinguished from the peptides of another sample during MS analysis and directly compared. In order to determine the probability of detecting the hydroxylated Trp38 in our search if this abundancy is correct, a binomial test was performed. The probability was calculated by setting the probability for success to 26%, the number of experiments to 37, and by varying the number of successes from 0 to 37. Figure 5.3 shows the probability distribution for the different MS/MS counts. The figure shows many possible proportions where the highest probability was seen for 9 and 10 MS/MS counts, which is expected as these values are close to 26% of the 37 MS/MS counts. Some of the proportions are also extremely unlikely, such as the probability of not detecting any hydroxylation at Trp38 if the abundancy is actually 26% which was found to be 0.001%. This extreme unlikeliness of not detecting any Trp38 hydroxylation in our dataset indicates that the hypothesis is wrong and that the abundancy of Trp-OH is unlikely to be 26%. It was therefore also calculated what Trp38-OH abundancy for which the probability for being true was 95%. Given that no Trp38-OH events were detected over 37 MS/MS events it was calculated a 95% probability that 0.14% of the peptides are Trp38 hydroxylated.

It was also searched for Lys-Ac in the deep PTM analysis in order to test the hypothesis by Seo *et al*. stating that autoacetylation of Lys136 on NAA10 is crucial for its ability to display KAT activity. This is also suggested to involve Lys29 and Lys198 (Seo *et al*., 2010). Only one acetylated lysine was found among the NAA10-derived peptides, which was detected for Lys78 and none of the previously reported autoacetylation sites (Table 5.4). Seo *et al*. compared the human NAA10 structure with the structure of several histone acetyltransferases and emphasized structure similarities with the GCN5 histone acetyltransferase from *Tetrahymena* solved by Clements *et al.* (Clements *et al.,* 2003). A high degree of structure similarity was seen when comparing the two structures, and Seo *et al*. further suggested that NAA10 may be a member of the same GCN5/PCAF histone acetyltransferase family. The structure of GCN5 bound to a 19-residue histone 3 peptide (PDB entry: 1PU9) is shown in Panel A of Figure 6.2 and was compared to a model of NAA10 (Panel B). The model was made by superimposing the histone 3 peptide with NAA10 from the crystal structure of human NatA solved by Gottlieb and Marmorstein (PDB entry: 6C95) in order to investigate how the peptide may be oriented in the active site of NAA10. As the peptide is just superimposed and not docked to NAA10 this is just a model that does not necessarily display the exact binding between NAA10 and the histone 3 peptide. Despite the overall structure similarity, the substrate binding sites seemed to have substantial differences. The active site of GCN5 seemed to be more open and accessible compared to the one of NAA10, which is partially blocked by the α1-α2 and the β6-β7 loops as described by Liszczak *et al*. and Gottlieb and Marmorstein. The family of GCN5/PCAF KATs are also known to exhibit specificity for Lys14 of the histone 3 peptide which is shown in yellow. This model may further support that structural limitations prevents monomeric NAA10 from acting on lysine residues, making it questionable whether NAA10 is autoacetylated and if it is associated with KAT activity.

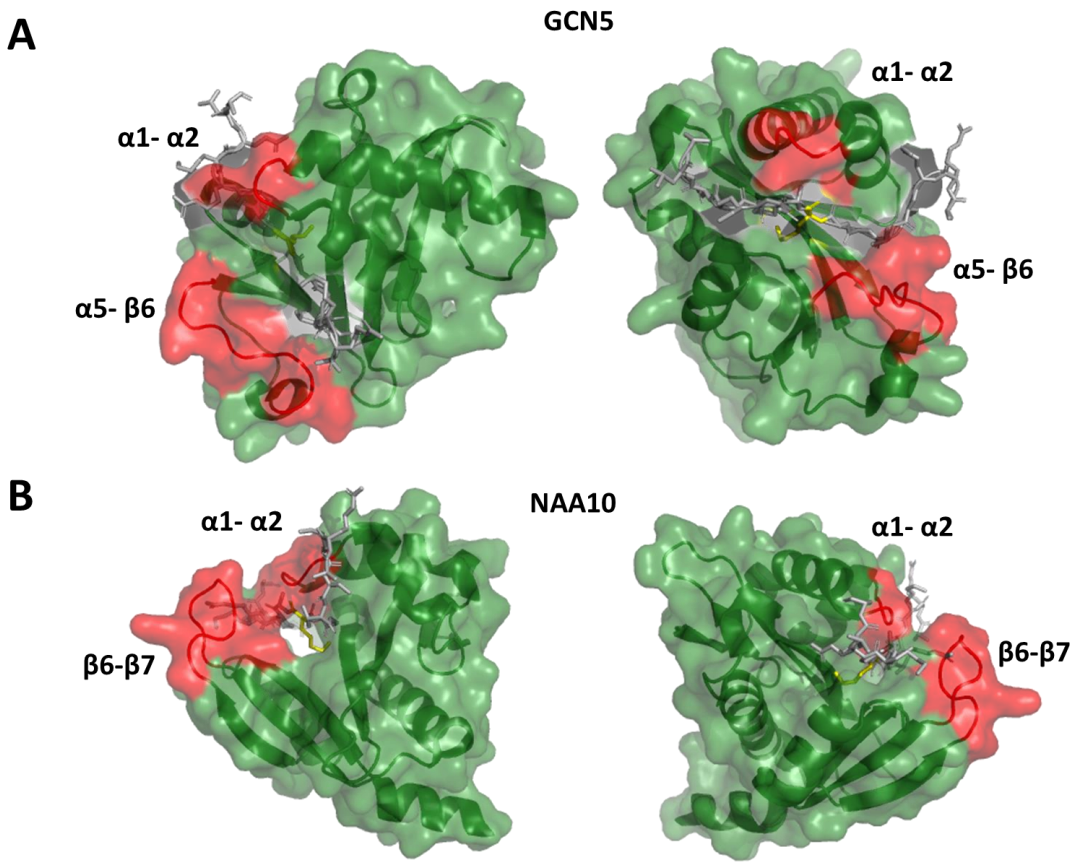


**Figure 6.2: Comparison of GCN5 from *Tetrahymena* bound to a histone 3 peptide and human NAA10 structures.** A: Structure of GCN5 from *Tetrahymenta* (PDB entry: 1PU9). GCN5 (green) is bound to a histone 3 peptide (gray), and is known to act on Lys14 (yellow) B: A model was made by superimposing the NAA10 structure (PDB entry: 6C95) with the histone 3 peptide substrate (yellow) of GCN5 to simulate how it may fit in the NAA10 substrate gate and the orientation of Lys14 (yellow).

Altogether, our MS analyses provide evidence that a fraction of the peptides identified in our samples are modified with Asn-OH, Asp-OH, Trp-OH and/or Met-O. Trp38 on the other hand was found to be unmodified in this analysis. A high MS/MS count, a high sequence coverage and the LFQ intensity percentile counts for both NAA10 and the other NatA components indicate a successful IP of monomeric and NAA15-complexed NAA10, as well as successful coprecipitation of HYPK and NAA50. A binomial test also revealed an extremely low probability for not detecting Trp38-OH if the abundancy is 26%, and it was estimated a probability of 95% that the abundancy is 0.14% based on no Trp38 hydroxylation events detected over 37 MS/MS counts. Based on previous studies where such oxidative modifications have been introduced as artefacts of in-gel digestion it can be speculated that the hydroxylation detected are artefacts of sample preparation rather than actual PTMs. If Trp38-OH had a biological function it would be likely to think that it would be independent of sample preparation and identified in several samples, and the lack of systematic detection makes it less likely of being a biological PTM. Despite the identification of peptides derived from the other NatA interaction partners, the MS analyses revealed no traces of FIH or HIF-1α peptides. No detected autoacetylation implied that this probably either occurs under specific conditions that were not met in these experiments, or that it occurs to a very small fraction of NAA10 which was not detected during the LC-MS/MS analysis.

### 6.1.2 Investigation of physical interactions between NAA10 and FIH

Kang *et al*. also reported coprecipitation of both endogenous and ectopically expressed FIH and NAA10, strongly indicating a physical interaction between the two proteins. Despite no sign of interactions between NAA10 and FIH in the LC-MS/MS analysis, HEK293 cells were transfected with FIH-V5 and/or Xpress-NAA10 and V5- and Xpress-IP were carried out to further look for physical interactions (Figure 5.4). WB analysis against anti-NAA10 revealed a band corresponding to the size of Xpress-NAA10 in the V5-IP from double transfected cells, indicating coprecipitation of Xpress-NAA10 with FIH-V5. An additional weak band of the same size was also detected in the same blot for Xpress-NAA10 transfected cells. Since there was no FIH-V5 to be immunoprecipitated in this sample, this band is probably caused by background binding or contamination from another sample during gel loading. The Xpress-NAA10 signals were weak, and the coprecipitation of Xpress-NAA10 detected with anti-NAA10 was not detected with anti-Xpress. This could be due to poor quality of the Xpress antibody and/or experimental errors impacting the concentration or binding of antibodies.

No coprecipitation was detected for FIH-V5 in the Xpress-IP with anti-V5 (Figure 5.4). One possible explanation for coprecipitation being detected only for Xpress-NAA10 and not for FIH-V5 might be that the accessibility of the tags has changed upon complex formation. When Xpress-NAA10 and FIH-V5 bind, it might cause the Xpress-tag to be less exposed, thereby making it difficult to isolate complexed Xpress-NAA10 using anti-Xpress (Figure 6.3). Thus, it may be that only monomeric Xpress-NAA10 is isolated during IP. The V5-tag on the other hand may be more exposed, enabling for anti-V5 binding and isolation of the FIH-V5-Xpress-NAA10 complex. Kang *et al*. used HA-FIH and FLAG/SBP-NAA10 for their experiments where both tags are assumed to be located at the N-terminal end of the proteins. It is unknown why their results could not be reproduced in this experiment, but the use of different tags may have an impact.

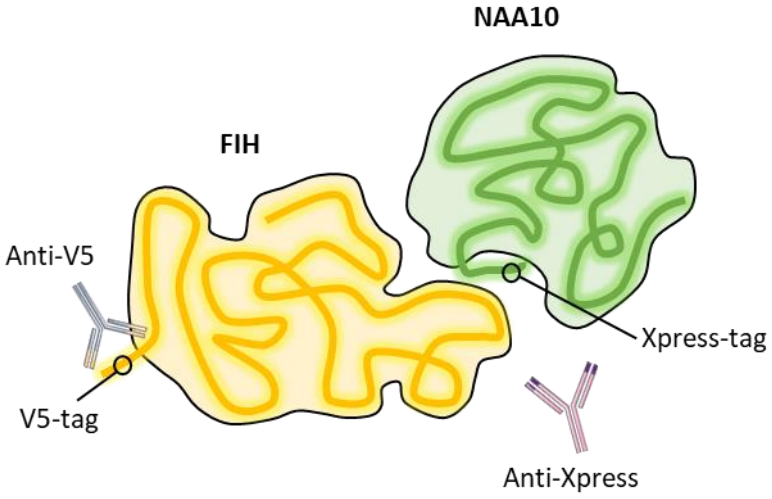


**Figure 6.3: Potential Xpress-NAA10/FIH-V5 complex formation may alter the accessibility of the Xpress-tag.** WB analyses of V5-IP samples of FIH-V5 and Xpress-NAA10 transfected cells revealed a potential physical interaction between Xpress-NAA10 and FIH-V5. The coprecipitation of Xpress-NAA10 with FIH-V5 was detected using anti-NAA10 and not with anti-Xpress, which may be due to the Xpress-tag becoming less exposed upon complex formation.

FIH- and NAA10-IP were also carried out for lysates from HAP1 WT and FIH KO cells to look for interactions between the endogenous protein (Figure 5.5). No coprecipitation was seen for neither the FIH-IP nor the NAA10-IP, and the blot presented was representative for two experiments. We know that anti-NAA10 binds on the C-terminal end of the protein, and if NAA10 binds FIH in the same region this will affect our ability to isolate the complex. Despite the long exposure time, the FIH-signal seemed quite weak in this assay, which might be due to the anti-FIH quality. It is uncertain whether the signal is not detected due to no coprecipitation or the quality of the antibodies used. The stoichiometry of monomeric NAA10 and NAA10 in complex with NAA15 is unknown, and if the majority of NAA10 exists in NatA complexes it might be difficult to isolate and detect monomeric NAA10.

Based on our results we conclude that a physical interaction between Xpress-NAA10 and FIH-V5 may occur, and that the challenges related to isolation of the protein complex by Xpress-IP may be related to a less exposed Xpress-tag upon complex formation. This conclusion is supported by a study by Rodriguez *et al*. where NAA10 was detected as a possible substrate of FIH using a quantitative interaction proteomics approach (Rodriguez *et al.,* 2016). However, several large-scale studies where NAA15 and NAA50 have been identified as interactors of NAA10 have not reported identification of FIH as a possible binding partner (Huttlin *et al*., 2017; Hein *et al*., 2015). This indicates that a potential interaction between NAA10 and FIH is most likely less stable than the interaction between NAA10 and the other NatA subunits.

### 6.1.3 Analysis of NAA10 expression levels in presence and absence of FIH

The lysates from the transfected HEK293 cells were also analyzed by WB to study the Xpress-NAA10 and FIH-V5 expression in presence and absence of each other (Figure 5.6 Panel A). An empty vector was used in this experiment to ensure equal amounts of DNA and transfection reagent introduced to the cells in each sample. The majority of the resulting blots with anti-Xpress revealed an increased Xpress-NAA10 signal when FIH-V5 was present, and the increase varied from 19-65% across four experiments performed under similar conditions. Over the same experiments the FIH-V5 signal also increased with 35-40% for double transfected cells. WB analysis of the same samples with anti-NAA10 only revealed a slight increase of 5-17% in Xpress-NAA10 and endogenous NAA10 signals, but the reason for this remains unknown.

Untransfected HAP1 WT and FIH KO cells were also lysed and subject to FIH- and NAA10 IP. WB analyses against anti-FIH, anti-NAA10 and anti-NAA15 revealed no detectable difference in NAA10 and NAA15 expression when compared to the expression of β-tubulin as a loading control (Panel B). Based on these results FIH does not seem to have any detectable impact on the amount of NatA expressed in these cells. Altogether, the results indicate that it is unclear but possible that Xpress-NAA10 and FIH-V5 become either directly or indirectly stabilized or upregulated by the presence of each other. They might also have other functions in the cell that indirectly leads to an increased expression through either stabilization or upregulation. The results also showed that FIH has little or no impact on endogenous NAA10, and reproducibility is therefore needed to further support any of the results presented before making any conclusions.

### 6.1.4 Analysis of NAA10 activity in presence and absence of FIH

It was further investigated in a $^{14}$C-Ac-CoA-based *in vitro* acetylation assay whether the NatA or NAA10 enzymatic activity towards the peptide substrates SESS and EEEI differed in presence and absence of FIH (Figure 5.7). Nt-acetylated histone 4 peptide, which contains several internal lysine residues, was also included as a substrate to test potential monomeric NAA10 KAT activity. FIH did not seem to have any impact on complex formation between NAA10 and NAA15 as the WB analysis in Panel A revealed similar amounts of immunoprecipitated NAA10 and NAA15 for both WT and FIH KO cells. FIH did also not seem to influence the NatA activity towards SESS, as the activity data in Panel B revealed that NatA from both WT and FIH KO cells displayed similar levels of activity in three experiments. A small decrease in the activity of NAA10 from FIH KO cells towards EEEI was seen compared to NAA10 from WT cells, indicating that FIH could have a small positive impact on monomeric NAA10 NAT activity in this case. However, this was only observed in two experiments whereas a third experiment showed similar activity, and the experiment should be repeated before making any conclusions.

Little activity was seen for the control IP both towards SESS and EEEI substrate peptides in Panel B. The control activity represents both background binding of enzymes to the magnetic Dynabeads™ and antibodies as well as chemical acetylation mediated by Ac-CoA in absence of any enzyme. For the histone 4 peptide, the control activity was higher than for the other two peptides. Since the same immunoprecipitates were used for all three peptide reactions, and therefore contain the same background binders, the high control activity for histone 4 is most likely caused by chemical acetylation. This indicates that Lys residues are more prone to chemical acetylation compared to protein N-termini. Also, a slightly higher activity for NAA10 from WT cells is seen compared to both the control and the FIH KO cells. A small part might be NAA10-mediated even though histone 4 is not a reported NAA10 KAT substrate. A more optimal experiment would involve peptides from a previously reported KAT substrate, but since this was not available during this work it will remain an experiment for a future study. In two out of three experiments a small decrease in activity towards histone 4 for NAA10 from FIH KO cells was seen, whereas a third experiment indicated an increased activity. As no clear trend was seen for these results, we are unable to conclude whether NAA10 activity towards histone 4 was influenced by FIH. Given that FIH-mediated Trp38 hydroxylation of NAA10 induces a conformational change to the NAA10 structure, it is unlikely that NAA10 is able to maintain its optimal NAT activity. For this experiment this was not the case, as its NAT activity towards EEEI is even higher in the presence of FIH. Alterations in NAA10 structure could also impair its interaction to NAA15, which is also found not to be the case in this experiment. The NAT activity of NAA10 in complex with NAA15 could also be impaired by FIH-mediated hydroxylation as a fraction of NAA10 would be in a monomeric state rather than in complex with NAA15. The activity assays towards SESS confirmed that NAA10 from WT and FIH KO cell displayed similar activity. Due to the small differences seen between WT and FIH KO cells and the low activity towards histone 4 in the reported experiments, further studies on this need to be undertaken with suggested KAT substrates.

For many studies where NAA10 KAT activity has been reported the Ac-CoA concentrations used have been artificially high (often 10-20 mM) compared to the *in vivo* cellular Ac-CoA concentration measured of about 10-20 µM for whole-cell samples and 20-30 µM for mitochondrial matrix (Chen *et al.*, 2016). It can therefore be speculated that this can result in chemical acetylation as well as NAA10-mediated acetylation that is not real or not necessarily occurring *in vivo* because high Ac-CoA levels pushes the system towards acetylation. Repeating the experiments with amounts of Ac-CoA closer to cellular concentrations might give different and more reliable results.

## 6.2 CONCLUSION

The NAT machinery comprises eight NATs known to date, all displaying distinct features and substrate specificities. Their role in Nt-Ac is evolutionary conserved, but over the past years several studies have also reported NAA10-dependent Lys-Ac of substrate proteins. This theory has been challenged by contradicting findings by several research groups. The NAT structure is also distinguished from the KAT structure by loops partly blocking the substrate gate, thereby restricting the NATs from acetylating Lys substrates. Seo *et al*. attempted to explain this with the requirement of NAA10 autoacetylation (Seo *et al*., 2010), whereas Kang *et al*. recently proposed a mechanism for this NAA10 duality involving hydroxylation of Trp38 (Kang *et al*., 2018). The hydroxylated NAA10 is suggested to undergo a conformational change and thereby mediate acetylation which is claimed to direct degradation of HIF-1α.

Despite limited information available regarding the conditions Kang *et al.* used for in-gel digestion and chromatography in their study it was attempted to reproduce their experiments to further investigate the NAA10 hydroxylation status. Our findings are summarized in a revised overview of NAA10 and FIH in HIF-1α signaling in Figure 6.4. Altogether, our primary MS analysis and the deep PTM analysis revealed that 36 and 47% respectively of the NAA10-derived peptides identified were hydroxylated on Asn, Asp, Trp and/or oxidized on Met residues, but the reported peptide of Trp38 was not found to be modified in any of the samples. Based on the claimed Trp38-OH abundancy of 26% the probability for not detecting any Trp38-OH was extremely low. In conclusion, our findings suggest that a fraction of monomeric and complexed NAA10 isolated from HEK293, HeLa and HAP1 cells under normoxic conditions are modified on Asn, Asp, Trp and Met residues, but that these modifications are likely to be artefacts of in-gel digestion. NAA10 autoacetylation of Lys136 has also been proposed as a requirement for KAT activity, which was also not detected in our dataset. Based on the artificially high levels of Ac-CoA used in certain experiments reporting NAA10-mediated Lys-Ac it is therefore unclear whether the Lys-Ac of the reported substrates are NAA10-mediated or chemical. High Ac-CoA levels may have induced misleading results and could potentially have pushed NAA10 to catalyze acetylation that is not naturally occurring *in vivo*.

It can also be speculated that NAA10 and FIH physically interacts, but the link between them remains unknown. FIH does not seem to have any detectable impact on NAA10/NAA15 complex formation, the amount of NatA in the cells nor its activity towards SESS. The results regarding its impact on monomeric NAA10 activity towards EEEI and histone 4 were considered to be inconclusive, and the experiment will have to be repeated. The Xpress-NAA10 and FIH-V5 signals were also found to increase for double transfected cells when using anti-Xpress and anti-V5, whereas only a slight increase was seen in the Xpress-NAA10 signal as well as for endogenous NAA10 when using anti-NAA10. Even though the experiment should be repeated before making any conclusions it may be possible that Xpress-NAA10 and FIH-V5 somehow are stabilizing or upregulating each other either in a direct or indirect manner.

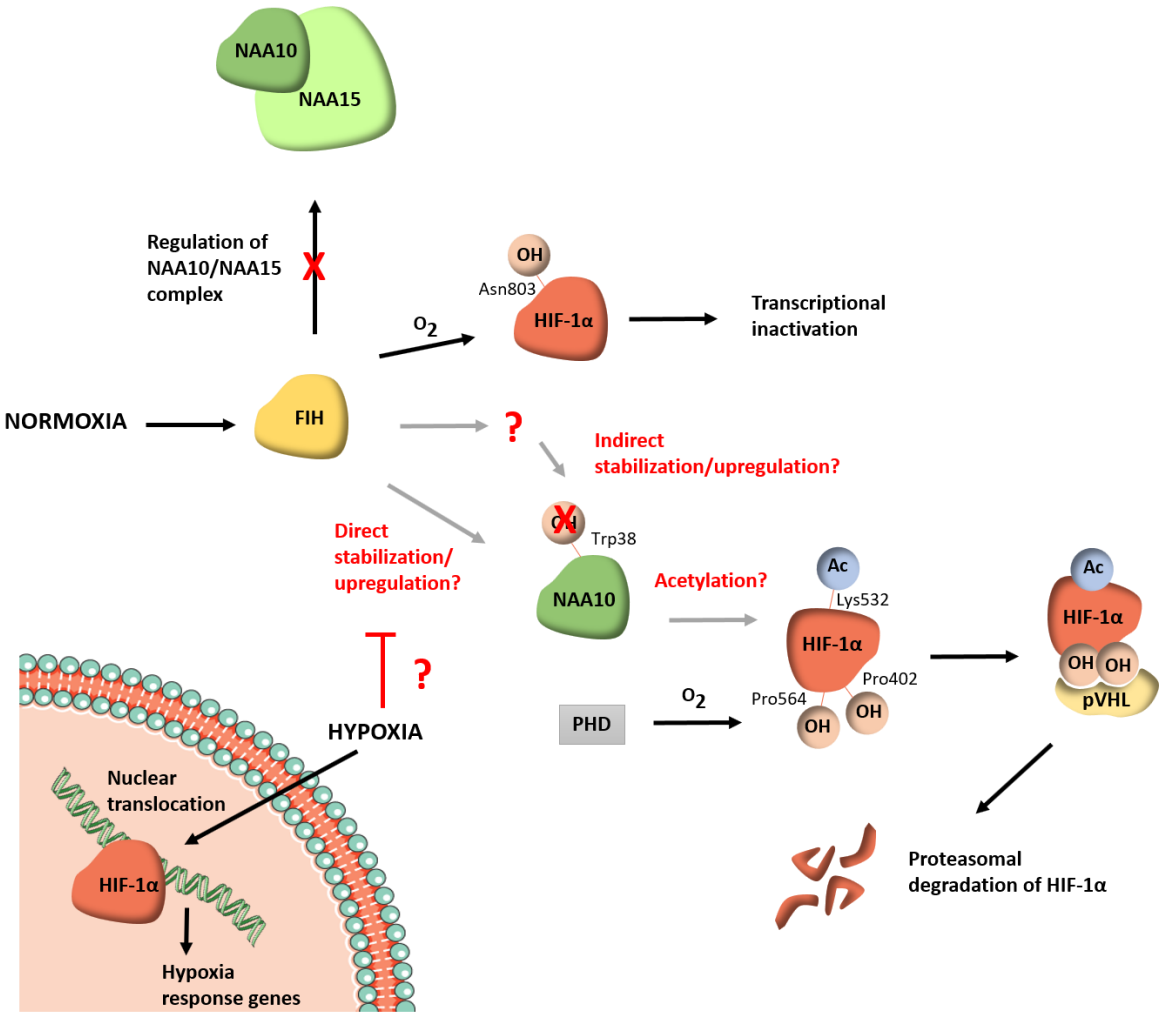


**Figure 6.4: Revised overview of FIH and NAA10 in HIF-1α regulation.** Based on our LC-MS/MS analysis Trp38-OH reported by Kang *et al*. as well as Asn, Asp hydroxylation and Met oxidation events identified in the dataset are likely to be artefacts of in-gel digestion upon MS sample preparation. FIH is still a possible interactor of NAA10 but the link between them remains unknown. It was found that FIH had no impact on the amount of NatA, its complex formation nor its activity towards SESS. WB analyses revealed that the proteins may physically interact, and that FIH-V5 might somehow stabilize or upregulate Xpress-NAA10 either directly or indirectly.

## 6.3 FUTURE PERSPECTIVES

NAA10 has been found to be of great importance in human development as well as in tumorigenesis by playing a key role in promoting cell proliferation, cell survival and regulation of cell metabolism (Aksnes, Ree and Arnesen, 2019). Still the mechanisms of how impaired NAA10 activity causes the reported phenotypes are poorly understood, and there are many questions regarding its complex functions that remain to be solved. It is still debated whether NAA10 displays KAT activity towards specific substrates, which will remain a major challenge for future studies.

In order to further investigate whether NAA10 acts on internal lysine residues of HIF-1α it would be interesting to repeat the NAA10-IP and $^{14}$C-Ac-CoA-based *in vitro* acetylation assay with a HIF-1α peptide in order to measure the activity of NatA and NAA10 immunoprecipitated from HAP1 WT and FIH KO cells. It would also be of interest to repeat this experiment with some of the other reported KAT substrates. The concentrations of Ac-CoA used should not be too high to avoid a high degree of chemical acetylation, and to achieve reliable results without artificially pushing the system towards acetylation. Another interesting experiment would be to perform an LC-MS/MS analysis of the Lys acetylome of HAP1 WT and FIH KO cells to look for differences. If the theory of FIH being crucial for NAA10 KAT activity is true, we would expect to see a decrease in Lys-Ac of substrates in FIH KO cells compared to WT cells. The MS analysis could also reveal potential differences in NAA10 hydroxylation status in WT and FIH KO cells.

From experiments done by Arnesen *et al*. it became evident that NAA10 and HIF-1α physically interact, but that neither overexpression nor silencing of NAA10 had an impact on HIF-1α stability (Arnesen *et al*., 2005b). To elucidate the link between NAA10, FIH and HIF-1α the IP experiments using properly functioning FIH, NAA10 and HIF-1α antibodies under normoxic and hypoxic conditions should be repeated both with endogenous and ectopically expressed proteins. In addition to studying physical interactions between the proteins in HEK293 or HeLa cells, it would be of interest to perform HIF-1α-IP from HAP1 WT and FIH KO cells under both normoxic and hypoxic conditions. Based on the suggested participation of FIH in HIF-1α degradation, one would expect a higher expression of HIF-1α in FIH KO cells.

Further investigation of whether NAA10 autoacetylation occurs on Lys136 and whether this acts as a key switch for NAA10 KAT activity may also be an interesting experiment for future studies. By first performing NAA10 IP and subsequently enrich for Ac-Lys one can use LC-MS/MS to look for NAA10 acetylation at the reported autoacetylation sites. Studies based on mutating Lys136 may also be performed to see whether it may impact NAA10 activity.

# 7 REFERENCES


Aksnes, H. *et al.* (2016) 'First Things First: Vital Protein Marks by N-Terminal Acetyltransferases', *Trends in Biochemical Sciences*. doi: 10.1016/j.tibs.2016.07.005.

Aksnes, H. *et al.* (2015) 'An organellar nα-acetyltransferase, naa60, acetylates cytosolic n termini of transmembrane proteins and maintains golgi integrity', *Cell Reports*. doi: 10.1016/j.celrep.2015.01.053.

Aksnes, H., Ree, R. and Arnesen, T. (2019) 'Co-translational, Post-translational, and Non-catalytic Roles of N-Terminal Acetyltransferases', *Molecular Cell*. doi: 10.1016/j.molcel.2019.02.007.

Allfrey, V. G. and Mirsky, A. E. (1964) 'Structural modifications of histones and their possible role in the regulation of RNA synthesis', *Science*. doi: 10.1126/science.144.3618.559.

Arnesen, T. *et al.* (2010) 'The Chaperone-Like Protein HYPK Acts Together with NatA in Cotranslational N-Terminal Acetylation and Prevention of Huntingtin Aggregation', *Molecular and Cellular Biology*. doi: 10.1128/mcb.01199-09.

Arnesen, T. *et al.* (2005a) 'Identification and characterization of the human ARD1-NATH protein acetyltransferase complex', *Biochemical Journal*. doi: 10.1042/BJ20041071.

Arnesen, T. *et al.* (2005b) 'Interaction between HIF-1α (ODD) and hARD1 does not induce acetylation and destabilization of HIF-1α', *FEBS Letters*. doi: 10.1016/j.febslet.2005.10.036.

Arnesen, T. *et al.* (2009) 'Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0901931106.

Baeza, J., Smallegan, M. J. and Denu, J. M. (2015) 'Site-specific reactivity of nonenzymatic lysine acetylation', *ACS Chemical Biology*. doi: 10.1021/cb500848p.

Bartels, T., Choi, J. G. and Selkoe, D. J. (2011) 'α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation', *Nature*. doi: 10.1038/nature10324.

Behnia, R. *et al.* (2004) 'Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p', *Nature Cell Biology*. doi: 10.1038/ncb1120.

Bilton, R. *et al.* (2005) 'Arrest-defective-1 protein, an acetyltransferase, does not alter stability of hypoxia-inducible factor (HIF)-1α and is not induced by hypoxia or HIF', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M504482200.

Blatch, G. L. and Lässle, M. (1999) 'The tetratricopeptide repeat: A structural motif mediating protein-protein interactions', *BioEssays*. doi: 10.1002/(SICI)1521-1878(199911)21:11<932::AID-BIES5>3.0.CO;2-N.

Boutureira, O. and Bernardes, G. J. L. (2015) 'Advances in chemical protein modification', *Chemical Reviews*. doi: 10.1021/cr500399p.

Boys, B. L. *et al.* (2009) 'Protein oxidative modifications during electrospray ionization: Solution phase electrochemistry or corona discharge-induced radical attack?', *Analytical Chemistry*. doi: 10.1021/ac900243p.

Casey, J. P. *et al.* (2015) 'NAA10 mutation causing a novel intellectual disability syndrome with Long QT due to N-terminal acetyltransferase impairment', *Scientific Reports*. doi: 10.1038/srep16022.

Chen, W. W. *et al.* (2016) 'Absolute Quantification of Matrix Metabolites Reveals the Dynamics of Mitochondrial Metabolism', *Cell*. doi: 10.1016/j.cell.2016.07.040.

Clements, A. *et al.* (2003) 'Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase', *Molecular Cell*. doi: 10.1016/S1097-2765(03)00288-0.

Cox, J. *et al.* (2014) 'Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ', *Molecular and Cellular Proteomics*. doi: 10.1074/mcp.M113.031591.

Cox, J. *et al.* (2011) 'Andromeda: A peptide search engine integrated into the MaxQuant environment', *Journal of Proteome Research*. doi: 10.1021/pr101065j.

Cox, J. and Mann, M. (2008) 'MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification', *Nature Biotechnology*. doi: 10.1038/nbt.1511.

Dinh, T. V. *et al.* (2015) 'Molecular identification and functional characterization of the first Nα-acetyltransferase in plastids by global acetylome profiling', *Proteomics*. doi: 10.1002/pmic.201500025.

Drazic, A. *et al.* (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: 10.1073/pnas.1718336115.

Drazic, A. and Arnesen, T. (2017) '[14C]-acetyl-coenzyme A-based in vitro N-terminal acetylation assay', in *Methods in Molecular Biology*. doi: 10.1007/978-1-4939-6850-3_1.

Drazic, A. *et al.* (2016) 'The world of protein acetylation', *Biochimica et Biophysica Acta - Proteins and Proteomics*. doi: 10.1016/j.bbapap.2016.06.007.

Evjenth, R. *et al.* (2009) 'Human Naa50p (Nat5/San) displays both protein Nα- and Nε-acetyltransferase activity', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M109.001347.

Fisher, T. S. *et al.* (2005) 'Analysis of ARD1 function in hypoxia response using retroviral RNA interference', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M412055200.

Fluge, Ø. *et al.* (2002) 'NATH, a novel gene overexpressed in papillary thyroid carcinomas', *Oncogene*. doi: 10.1038/sj.onc.1205687.

Forte, G. M. A., Pool, M. R. and Stirling, C. J. (2011) 'N-terminal acetylation inhibits protein targeting to the endoplasmic reticulum', *PLoS Biology*. doi: 10.1371/journal.pbio.1001073.

Froelich, J. M. and Reid, G. E. (2008) 'The origin and control of ex vivo oxidative peptide modifications prior to mass spectrometry analysis', *Proteomics*. doi: 10.1002/pmic.200700792.

Glozak, M. A. *et al.* (2005) 'Acetylation and deacetylation of non-histone proteins', *Gene*. doi: 10.1016/j.gene.2005.09.010.

Goris, M. *et al.* (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: 10.1073/pnas.1719251115.

Gottlieb, L. and Marmorstein, R. (2018) 'Structure of Human NatA and Its Regulation by the Huntingtin Interacting Protein HYPK', *Structure*. doi: 10.1016/j.str.2018.04.003.

Hein, M. Y. *et al.* (2015) 'A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances', *Cell*. doi: 10.1016/j.cell.2015.09.053.

Hole, K. *et al.* (2011) 'The human N-Alpha-acetyltransferase 40 (hNaa40p/hNatD) is conserved from yeast and N-terminally acetylates histones H2A and H4', *PLoS ONE*. doi: 10.1371/journal.pone.0024713.

Holmes, W. M. *et al.* (2014) 'Loss of amino-terminal acetylation suppresses a prion phenotype by modulating global protein folding', *Nature Communications*. doi: 10.1038/ncomms5383.

Hong, H. *et al.* (2017) 'Molecular Basis of Substrate Specific Acetylation by N-Terminal Acetyltransferase NatB', *Structure*. doi: 10.1016/j.str.2017.03.003.

Huttlin, E. L. *et al.* (2017) 'Architecture of the human interactome defines protein communities and disease networks', *Nature*. doi: 10.1038/nature22366.

Hwang, C. S., Shemorry, A. and Varshavsky, A. (2010) 'N-terminal acetylation of cellular proteins creates specific degradation signals', *Science*. doi: 10.1126/science.1183147.

Imhof, A. *et al.* (1997) 'Acetylation of general transcription factors by histone acetyltransferases', *Current Biology*. doi: 10.1016/S0960-9822(06)00296-X.

Ingram, A. K., Cross, G. A. M. and Horn, D. (2000) 'Genetic manipulation indicates that ARD1 is an essential Nα-acetyltransferase in Trypanosoma brucei', *Molecular and Biochemical Parasitology*. doi: 10.1016/S0166-6851(00)00322-4.

Ivan, M. *et al.* (2001) 'HIFα targeted for VHL-mediated destruction by proline hydroxylation: Implications for O2 sensing', *Science*. doi: 10.1126/science.1059817.

Jaakkola, P. *et al.* (2001) 'Targeting of HIF-α to the von Hippel-Lindau ubiquitylation complex by O 2 -regulated prolyl hydroxylation', *Science*. doi: 10.1126/science.1059796.

Jeong, J. W. *et al.* (2002) 'Regulation and destabilization of HIF-1α by ARD1-mediated acetylation', *Cell*. doi: 10.1016/S0092-8674(02)01085-1.

Kalvik, T. V. and Arnesen, T. (2013) 'Protein N-terminal acetyltransferases in cancer', *Oncogene*. doi: 10.1038/onc.2012.82.

Kang, J. *et al.* (2018) 'FIH permits NAA10 to catalyze the oxygen-dependent lysyl-acetylation of HIF-1α', *Redox Biology*. doi: 10.1016/j.redox.2018.09.002.

Keilhauer, E. C., Hein, M. Y. and Mann, M. (2015) 'Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS)', *Molecular and Cellular Proteomics*. doi: 10.1074/mcp.M114.041012.

Kelstrup, C. D. *et al.* (2018) 'Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics', *Journal of Proteome Research*. doi: 10.1021/acs.jproteome.7b00602.

Kim, K. *et al.* (2015) 'Reducing protein oxidation in low-flow electrospray enables deeper investigation of proteoforms by top down proteomics', *EuPA Open Proteomics*. doi: 10.1016/j.euprot.2015.05.005.

Klont, F. *et al.* (2018) 'Assessment of Sample Preparation Bias in Mass Spectrometry-Based Proteomics', *Analytical Chemistry*. doi: 10.1021/acs.analchem.8b00600.

Lando, D. *et al.* (2002) 'Asparagine hydroxylation of the HIF transactivation domain: A hypoxic switch', *Science*. doi: 10.1126/science.1068592.

Lee, C. F. *et al.* (2010) 'hNaa10p contributes to tumorigenesis by facilitating DNMT1-mediated tumor suppressor gene silencing', *Journal of Clinical Investigation*. doi: 10.1172/JCI42275.

Lee, E. J. *et al.* (2017) 'SAMHD1 acetylation enhances its deoxynucleotide triphosphohydrolase activity and promotes cancer cell proliferation', *Oncotarget*. doi: 10.18632/oncotarget.19704.

Lim, J. H., Park, J. W. and Chun, Y. S. (2006) 'Human arrest defective 1 acetylates and activates β-catenin, promoting lung cancer cell proliferation', *Cancer Research*. doi: 10.1158/0008-5472.CAN-06-3171.

Linē, A. *et al.* (2002) ‘Serological identification and expression analysis of gastric cancer-associated genes’, *British Journal of Cancer*. doi: 10.1038/sj.bjc.6600321.

Liszczak, G., Arnesen, T. and Marmorsteins, R. (2011) ‘Structure of a ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation’, *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111.282863.

Liszczak, G. *et al.* (2013) ‘Molecular basis for N-terminal acetylation by the heterodimeric NatA complex’, *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.2636.

Lu Vo, T. T. *et al.* (2017) ‘ARD1-mediated aurora kinase A acetylation promotes cell proliferation and migration’, *Oncotarget*. doi: 10.18632/oncotarget.19332.

Magin, R. S., Liszczak, G. P. and Marmorstein, R. (2015) ‘The molecular basis for Histone H4- and H2A-specific amino-terminal acetylation by NatD’, *Structure*. doi: 10.1016/j.str.2014.10.025.

Magin, R. S., March, Z. M. and Marmorstein, R. (2016) ‘The N-terminal acetyltransferase Naa10/ARD1 does not acetylate lysine residues’, *Journal of Biological Chemistry*. doi: 10.1074/jbc.M115.709428.

Martin, D. T. *et al.* (2007) ‘Tubedown expression correlates with the differentiation status and aggressiveness of neuroblastic tumors’, *Clinical Cancer Research*. doi: 10.1158/1078-0432.CCR-06-1716.

Midorikawa, Y. *et al.* (2002) ‘Identification of genes associated with dedifferentiation of hepatocellular carcinoma with expression profiling analysis’, *Japanese Journal of Cancer Research*. doi: 10.1111/j.1349-7006.2002.tb01301.x.

Monda, J. K. *et al.* (2013) ‘Structural conservation of distinctive N-terminal acetylation-dependent interactions across a family of mammalian NEDD8 ligation enzymes’, *Structure*. doi: 10.1016/j.str.2012.10.013.

Murray-Rust, T. A. *et al.* (2006) ‘Purified recombinant hARD1 does not catalyse acetylation of Lys 532 of HIF-1α fragments in vitro’, *FEBS Letters*. doi: 10.1016/j.febslet.2006.02.012.

Myklebust, L. M., Støve, S. I. and Arnesen, T. (2015) ‘Naa10 in development and disease’, *Oncotarget*. doi: 10.18632/oncotarget.5867.

Ponomarenko, E. A. *et al.* (2016) ‘The Size of the Human Proteome: The Width and Depth’, *International Journal of Analytical Chemistry*. doi: 10.1155/2016/7436849.

Popp, B. *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*. doi: 10.1038/ejhg.2014.150.

Qian, X. *et al.* (2017) ‘Phosphoglycerate Kinase 1 Phosphorylates Beclin1 to Induce Autophagy’, *Molecular Cell*. doi: 10.1016/j.molcel.2017.01.027.

Ree, R. *et al.* (2015) ‘The N-terminal acetyltransferase Naa10 is essential for zebrafish development’, *Bioscience Reports*. doi: 10.1042/BSR20150168.

Ree, R., Varland, S. and Arnesen, T. (2018) ‘Spotlight on protein N-terminal acetylation’, *Experimental and Molecular Medicine*. doi: 10.1038/s12276-018-0116-z.

Ren, T. *et al.* (2008) ‘Generation of novel monoclonal antibodies and their application for detecting ARD1 expression in colorectal cancer’, *Cancer Letters*. doi: 10.1016/j.canlet.2008.01.028.

Roberts, K. M. and Fitzpatrick, P. F. (2013) ‘Mechanisms of tryptophan and tyrosine hydroxylase’, *IUBMB Life*. doi: 10.1002/iub.1144.

Rodriguez, J. *et al.* (2016) 'Substrate-Trapped Interactors of PHD3 and FIH Cluster in Distinct Signaling Pathways', *Cell Reports*. doi: 10.1016/j.celrep.2016.02.043.

Rope, A. 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*. doi: 10.1016/j.ajhg.2011.05.017.

Salceda, S. and Caro, J. (1997) 'Hypoxia-inducible Factor 1α (HIF-1α) Protein Is Rapidly Degraded by the Ubiquitin-Proteasome System under Normoxic Conditions', *Journal of Biological Chemistry*. doi: 10.1074/jbc.272.36.22642.

Saunier, C. *et al.* (2016) 'Expanding the Phenotype Associated with NAA10-Related N-Terminal Acetylation Deficiency', *Human Mutation*. doi: 10.1002/humu.23001.

Scholz, C. C. *et al.* (2016) 'FIH Regulates Cellular Metabolism through Hydroxylation of the Deubiquitinase OTUB1', *PLoS Biology*. doi: 10.1371/journal.pbio.1002347.

Seo, J. H. *et al.* (2010) 'Arrest defective 1 autoacetylation is a critical step in its ability to stimulate cancer cell proliferation', *Cancer Research*. doi: 10.1158/0008-5472.CAN-09-3258.

Seo, J. H. *et al.* (2016) 'ARD1-mediated Hsp70 acetylation balances stress-induced protein refolding and degradation', *Nature Communications*. doi: 10.1038/ncomms12882.

Setty, S. R. G. *et al.* (2004) 'Golgi targeting of Arf-like GTPase Arl3p requires its Nα-acetylation and the integral membrane protein Sys1p', *Nature Cell Biology*. doi: 10.1038/ncb1121.

Shacter, E. (2000) 'Quantification and significance of protein oxidation in biological samples', in *Drug Metabolism Reviews*. doi: 10.1081/DMR-100102336.

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

Shin, D. H. *et al.* (2009) 'Arrest defective-1 controls tumor cell behavior by acetylating myosin light chain kinase', *PLoS ONE*. doi: 10.1371/journal.pone.0007451.

Shin, S. H. *et al.* (2014) 'Arrest defective 1 regulates the oxidative stress response in human cells and mice by acetylating methionine sulfoxide reductase A', *Cell Death and Disease*. doi: 10.1038/cddis.2014.456.

Starheim, K. K., Gevaert, K. and Arnesen, T. (2012) 'Protein N-terminal acetyltransferases: When the start matters', *Trends in Biochemical Sciences*. doi: 10.1016/j.tibs.2012.02.003.

Støve, S. I. I. *et al.* (2016) 'Crystal Structure of the Golgi-Associated Human Nα-Acetyltransferase 60 Reveals the Molecular Determinants for Substrate-Specific Acetylation', *Structure*. doi: 10.1016/j.str.2016.04.020.

Sönnichsen, B. *et al.* (2005) 'Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans', *Nature*. doi: 10.1038/nature03353.

Tanka, S. *et al.* (1989) 'Cloning and molecular characterization of the gene rimL wich encodes an enzyme acetylating ribosomal protein L12 of Escherichia coki K12', *MGG Molecular & General Genetics*. doi: 10.1007/BF02464895.

Towbin, H. *et al.* (1979) 'Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.76.9.4350.

Trexler, A. J. and Rhoades, E. (2012) 'N-terminal acetylation is critical for forming α-helical oligomer of α-synuclein', *Protein Science*. doi: 10.1002/pro.2056.

Tyanova, S. *et al.* (2016) 'The Perseus computational platform for comprehensive analysis of (prote)omics data', *Nature Methods*. doi: 10.1038/nmeth.3901.

Van Damme, P. *et al.* (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: 10.1074/mcp.M110.004580.

Van Damme, P. *et al.* (2011b) 'NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation', *PLoS Genetics*. doi: 10.1371/journal.pgen.1002169.

Varland, S., Osberg, C. and Arnesen, T. (2015) 'N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects', *Proteomics*. doi: 10.1002/pmic.201400619.

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

Vetting, M. W. *et al.* (2005) 'Structure and functions of the GNAT superfamily of acetyltransferases', *Archives of Biochemistry and Biophysics*. doi: 10.1016/j.abb.2004.09.003.

Vo, T. T. L. *et al.* (2018) 'Versatility of ARD1/NAA10-mediated protein lysine acetylation', *Experimental and Molecular Medicine*. doi: 10.1038/s12276-018-0100-7.

Wagner, G. R. and Payne, R. M. (2013) 'Widespread and enzyme-independent Nε-acetylation and Nε-succinylation of proteins in the chemical conditions of the mitochondrial matrix', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M113.486753.

Wang, G. L. *et al.* (1995) 'Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O 2 tension', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.92.12.5510.

Wang, Y. *et al.* (2010) 'Drosophila variable nurse cells encodes arrest defective 1 (ARD1), the catalytic subunit of the major N-terminal acetyltransferase complex', *Developmental Dynamics*. doi: 10.1002/dvdy.22418.

Wang, Zehua *et al.* (2012) 'Inactivation of androgen-induced regulator ARD1 inhibits androgen receptor acetylation and prostate tumorigenesis', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1113356109.

Yang, M. *et al.* (2011a) 'Asparagine and aspartate hydroxylation of the cytoskeletal ankyrin family is catalyzed by factor-inhibiting hypoxia-inducible factor', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M110.193540.

Yang, M. *et al.* (2011b) 'Factor-inhibiting hypoxia-inducible factor (FIH) catalyses the post-translational hydroxylation of histidinyl residues within ankyrin repeat domains', *FEBS Journal*. doi: 10.1111/j.1742-4658.2011.08022.x.

Yoon, H. *et al.* (2014) 'NAA10 controls osteoblast differentiation and bone formation as a feedback regulator of Runx2', *Nature Communications*. doi: 10.1038/ncomms6176.

Yoshikawa, A. *et al.* (1987) 'Cloning and nucleotide sequencing of the genes rimI and rimJ which encode enzymes acetylating ribosomal proteins S18 and S5 of Escherichia coli K12', *MGG Molecular & General Genetics*. doi: 10.1007/BF00331153.

Yu, F. *et al.* (2001) 'HIF-1α binding to VHL is regulated by stimulus-sensitive proline hydroxylation', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.181341498.

Yu, M. *et al.* (2009a) ‘Immunohistochemical analysis of human arrest-defective-1 expressed in cancers in vivo’, *Oncology Reports*. doi: 10.3892/or_00000303.

Yu, M. *et al.* (2009b) ‘Correlation of expression of human arrest-defective-1 (hard1) protein with breast cancer’, *Cancer Investigation*. doi: 10.3109/07357900902769723.

Zeng, Y. *et al.* (2014) ‘Inhibition of STAT5a by Naa10p contributes to decreased breast cancer metastasis’, *Carcinogenesis*. doi: 10.1093/carcin/bgu132.

Zhang, Y. *et al.* (2013) ‘Protein analysis by shotgun/bottom-up proteomics’, *Chemical Reviews*. doi: 10.1021/cr3003533.

# 8 SUPPLEMENTARY

**Table 8.1:** List of NAA10 peptides identified in MS analysis by Kang *et al*.

| Peptide | | MS/MS counts | |
|---|---|---|---|
| | **Modification *** | **DMOG** | **Control** |
| AALHLYSnTLnFQISEVEPK | Asn8-OH, Asn11-OH | 4 | 11 |
| AALHLYSnTLNFQISEVEPK | Asn8-OH | 2 | 10 |
| AALHLYSNTLnFQISEVEPK | Asn11-OH | 2 | 10 |
| AALHLYSNTLNFQISEVEPK | Unmodified | 68 | 90 |
| AALHLYSNTLNFQISEVEPKYYADGEDAYAMK | Unmodified | 2 | 1 |
| AALHLYSNTLNFQISEVEPKYYADGEdAYAMKR | Asp27-OH | | 1 |
| AMIENFNAKYVSLHVR | Unmodified | 2 | |
| AMIEnFNAKYVSLHVR | Asn5-OH | | 1 |
| AMIENFnAKYVSLHVR | Asn7-OH | | 1 |
| AMIENFNAKYVSLHVRK | Unmodified | 1 | |
| DLSEVSETTESTDVK | Unmodified | 2 | |
| DLSEVSETTESTDVKDSSEASDSAS | Unmodified | 57 | 53 |
| DLTQMAdELR | Asp7-OH | 3 | 3 |
| DLTQMAdELR | Unmodified | 2 | 2 |
| dLTQMAdELRR | Asp1-OH, Asp7-OH | 2 | 2 |
| dLTQMADELRR | Asp1-OH | 1 | 1 |
| DLTQMAdELRR | Asp7-OH | 4 | 3 |
| DLTQMADELRR | Unmodified | 9 | 16 |
| DLTQMADELRRHLELK | Unmodified | 1 | |
| EEKGLAAEDSGGDSKDLSEVSETTESTDVK | Unmodified | 1 | |
| GLAAEDSGGDSK | Unmodified | 1 | 1 |
| GLAAEDSGGDSKDLSEVSETTESTDVK | Unmodified | 5 | 4 |
| GLAAEDSGGDSKDLSEVSETTESTDVKDSSEASDSAS | Unmodified | 126 | 135 |
| GLAAEdSGGdSKDLSEVSETTESTDVKDSSEASDSAS | Asp6-OH, Asp10-OH | | 2 |
| GLAAEdSGGDSKdLSEVSETTESTDVKDSSEASDSAS | Asp6-OH, Asp13-OH | | 2 |
| GLAAEdSGGDSKDLSEVSETTESTDVKDSSEASDSAS | Asp6-OH | | 2 |
| GLAAEDSGGdSKdLSEVSETTESTDVKDSSEASDSAS | Asp10-OH, Asp13-OH | | 1 |
| GLAAEDSGGdSKDLSEVSETTESTdVKDSSEASDSAS | Asp10-OH, Asp25-OH | | 1 |
| GLAAEDSGGdSKDLSEVSETTESTDVKdSSEASDSAS | Asp10-OH, Asp28-OH | | 1 |
| GLAAEDSGGdSKDLSEVSETTESTDVKDSSEASDSAS | Asp10-OH | | 2 |
| GLAAEDSGGDSKdLSEVSETTESTdVKDSSEASDSAS | Asp13-OH, Asp25-OH | | 1 |
| GLAAEDSGGDSKdLSEVSETTESTDVKdSSEASDSAS | Asp13-OH, Asp28-OH | | 1 |
| GLAAEDSGGDSKdLSEVSETTESTDVKDSSEASDSAS | Asp13-OH | | 2 |
| GLAAEDSGGDSKDLSEVSETTESTdVKdSSEASDSAS | Asp25-OH, Asp28-OH | | 1 |
| GLAAEDSGGDSKDLSEVSETTESTdVKDSSEASdSAS | Asp25-OH, Asp34-OH | | 1 |
| GLAAEDSGGDSKDLSEVSETTESTdVKDSSEASDSAS | Asp25-OH | | 2 |
| GLAAEDSGGDSKDLSEVSETTESTDVKdSSEASdSAS | Asp28-OH, Asp34-OH | | 1 |
| GLAAEDSGGDSKDLSEVSETTESTDVKdSSEASDSAS | Asp28-OH | | 2 |
| GLAAEDSGGDSKDLSEVSETTESTDVKDSSEASdSAS | Asp34-OH | | 2 |
| GRHVVLGAIENK | Unmodified | 2 | 2 |
| GRHVVLGAIENKVESK | Unmodified | 23 | 8 |
| HVVLGAIENK | Unmodified | | 2 |
| HVVLGAIENKVESK | Unmodified | 83 | 80 |
| IVGYVLAK | Unmodified | | 1 |
| IVGYVLAKMEEdPDDVPHGHITSLAVKR | Asp12-OH | 1 | |
| IVGYVLAKMEEDPdDVPHGHITSLAVKR | Asp14-OH | 1 | |
| IVGYVLAKMEEDPDdVPHGHITSLAVKR | Asp15-OH | 1 | |
| LGLAQKLMdQASR | Asp9-OH | 5 | |
| LGLAQKLMDQASR | Unmodified | 4 | 4 |
| LGLAQKLMdQASR | Asp9-OH | | 6 |
| LMDQASR | Unmodified | 3 | 3 |

| | | | |
|---|---|---|---|
| MEEdPdDVPHGHITSLAVK | Asp4-OH, Asp6-OH | 1 | 1 |
| MEEdPDdVPHGHITSLAVK | Asp4-OH, Asp7-OH | 1 | 1 |
| MEEdPDDVPHGHITSLAVK | Asp4-OH | 7 | 5 |
| MEEDPddVPHGHITSLAVK | Asp6-OH, Asp7-OH | 1 | |
| MEEDPdDVPHGHITSLAVK | Asp6-OH | 7 | 5 |
| MEEDPDdVPHGHITSLAVK | Asp7-OH | 5 | 3 |
| MEEDPDDVPHGHITSLAVK | Unmodified | 9 | 5 |
| MEEdPdDVPHGHITSLAVKR | Asp4-OH, Asp6-OH | 1 | |
| MEEdPDdVPHGHITSLAVKR | Asp4-OH, Asp7-OH | 1 | 1 |
| MEEdPDDVPHGHITSLAVKR | Asp4-OH | 5 | 3 |
| MEEDPddVPHGHITSLAVKR | Asp6-OH, Asp7-OH | 1 | 1 |
| MEEDPdDVPHGHITSLAVKR | Asp6-OH | 4 | 2 |
| MEEDPDdVPHGHITSLAVKR | Asp7-OH | 3 | 1 |
| MEEDPDDVPHGHITSLAVKR | Unmodified | 11 | 8 |
| RDLTQMADELRR | Unmodified | 2 | 2 |
| RHLELKEK | Unmodified | | 1 |
| RLGLAQK | Unmodified | | 1 |
| RLGLAQKLMdQASR | Asp10-OH | 2 | 3 |
| RLGLAQKLMDQASR | Unmodified | 3 | 3 |
| YVSLHVR | Unmodified | | 1 |
| YYAdGEdAYAMK | Asp4-OH, Asp7-OH | 2 | 4 |
| YYAdGEDAYAMK | Asp4-OH | 2 | 3 |
| YYADGEdAYAMK | Asp7-OH | 2 | 4 |
| YYADGEDAYAMK | Unmodified | 6 | 12 |
| YYAdGEdAYAMKR | Asp4-OH, Asp7-OH | 1 | 3 |
| YYAdGEDAYAMKR | Asp4-OH | 3 | 5 |
| YYADGEdAYAMKR | Asp7-OH | 4 | 7 |
| YYADGEDAYAMKR | Unmodified | 10 | 20 |
| YYFYHGLSwPQLSYIAEDEnGK | Trp9-OH, Asn20-OH | | 1 |
| YYFYHGLSwPQLSYIAEDENGK | Trp9-OH | 1 | 2 |
| YYFYHGLSWPQLSYIAEdEnGK | Asp18-OH, Asn20-OH | 1 | 1 |
| YYFYHGLSWPQLSYIAEdENGK | Asp18-OH | 2 | 1 |
| YYFYHGLSWPQLSYIAEDENGK | Unmodified | 2 | 4 |
| YYFYHGLSwPQLSYIAEdENGKIVGYVLAK | Trp9-OH, Asp18-OH | 1 | 5 |
| YYFYHGLSWPQLSYIAEdEnGKIVGYVLAK | Asp18-OH, Asn20-OH | 2 | 4 |
| YYFYHGLSwPQLSYIAEDEnGKIVGYVLAK | Trp9-OH, Asn20-OH | | 5 |
| YYFYHGLSwPQLSYIAEDENGKIVGYVLAK | Trp9-OH | | 15 |
| YYFYHGLSWPQLSYIAEdENGKIVGYVLAK | Asp18-OH | 3 | 8 |
| YYFYHGLSWPQLSYIAEDEnGKIVGYVLAK | Asn20-OH | 3 | 9 |
| YYFYHGLSWPQLSYIAEDENGKIVGYVLAK | Unmodified | 36 | 51 |

* The numbering of modified residues is relative to the specific peptide, and not to the NAA10 sequence.

All data retrieved from Perseus for the primary MS analysis and the deep PTM analysis can be retrieved from the link below. The excel document contains six sheets of information about all peptides identified, NAA10 peptides and protein groups for both analyses. For all the peptides identified information about their modifications, intensity in each sample, MS/MS events etc. can be found. For the NAA10 peptides their position relative to the NAA10 sequence is also included. All the different protein groups are presented with the LFQ intensity for each protein groups in each sample, the sequence coverage etc.

http://vedlegg.uib.no/?id=69b2de70a71e5b03b6fcca87e8dabee7