



Published in final edited form as:

Circ Res. 2021 April 16; 128(8): 1156–1169. doi:10.1161/CIRCRESAHA.120.316966.

Mechanisms of Congenital Heart Disease Caused by NAA15 Haploinsufficiency

Tarsha Ward¹, Warren Tai¹, Sarah Morton^{1,2}, Francis Impens^{3,4,5}, Petra Van Damme⁶, Delphi Van Haver^{3,4,5}, Evy Timmerman^{3,4,5}, Gabriela Venturini^{1,7}, Kehan Zhang^{8,9}, Min Young Jang¹, Jon A. L. Willcox¹, Alireza Haghighi^{1,10,25}, Bruce D. Gelb¹¹, Wendy K. Chung¹², Elizabeth Goldmuntz¹³, George A. Porter Jr.¹⁴, Richard P. Lifton^{15,16}, Martina Brueckner^{15,17}, H. Joseph Yost¹⁸, Benoit G. Bruneau¹⁹, Joshua Gorham¹, Yuri Kim^{1,20}, Alexandre Pereira¹, Jason Homsey¹, Craig C. Benson¹, Steven R. DePalma¹, Sylvia Varland^{21,22,23}, Christopher S. Chen^{8,9}, Thomas Arnesen^{21,22,24}, Kris Gevaert^{3,5,*}, Christine Seidman^{1,10,25,*}, J.G. Seidman^{1,*}

¹Genetics, Harvard Medical School

²Division of Newborn Medicine, Boston Children's Hospital

³VIB Center for Medical Biotechnology, B-9000 Ghent, Belgium

⁴VIB Proteomics Core, B-9000 Ghent, Belgium

⁵Biomolecular Medicine, Ghent University, B-9000 Ghent, Belgium

⁶Biochemistry and Microbiology, Ghent University

⁷University of Sao Paulo, Sao Paulo, SP 05403-900

⁸Biomedical Engineering, Boston University, MA

⁹The Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston, MA

¹⁰Medicine, Brigham and Women's Hospital

¹¹Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York

¹²Pediatrics and Medicine, Columbia University Medical Center, New York

¹³Cardiology, Children's Hospital of Philadelphia, Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia

¹⁴Pediatrics, University of Rochester Medical Center, Rochester

¹⁵Genetics, Yale University School of Medicine, New Haven

¹⁶Laboratory of Human Genetics and Genomics, Rockefeller University, New York

¹⁷Pediatrics, Yale University School of Medicine, New Haven

Address correspondence to: Dr. Jonathan G. Seidman, New Research Building Room 256, 77 Avenue Louis Pasteur, Boston, MA 02115, Tel: 1-617-432-7830, seidman@genetics.med.harvard.edu.

*These authors contributed equally to this publication.

DISCLOSURES

The authors declare no competing interests.

¹⁸Molecular Medicine Program, University of Utah, Salt Lake City

¹⁹Gladstone Institutes, 1650 Owens St, San Francisco, CA 94158 USA

²⁰Division of Cardiovascular Medicine, Brigham and Women's Hospital

²¹Biomedicine, University of Bergen, N-5020 Bergen, Norway

²²Biological Sciences, University of Bergen, N-5020 Bergen, Norway

²³Donnelly Centre for Cellular and Biomolecular Research, Toronto, ON M5S 3E1, Canada

²⁴Surgery, Haukeland University Hospital, N-5021 Bergen, Norway

²⁵Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA

Abstract

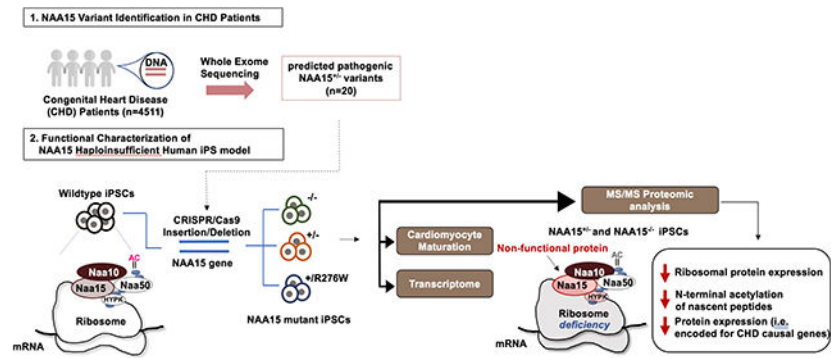
Rationale: NAA15 is a component of the N-terminal (Nt) acetyltransferase complex, NatA. The mechanism by which NAA15 haploinsufficiency causes congenital heart disease (CHD) remains unknown. To better understand molecular processes by which NAA15 haploinsufficiency perturbs cardiac development, we introduced *NAA15* variants into human induced pluripotent stem cells (iPSCs) and assessed the consequences of these mutations on RNA and protein expression.

Objective: We aim to understand the role of NAA15 haploinsufficiency in cardiac development by investigating proteomic effects on NatA complex activity, and identifying proteins dependent upon a full amount of NAA15.

Methods and Results: We introduced heterozygous LoF, compound heterozygous and missense residues (R276W) in iPSC cells using CRISPR/Cas9. Haploinsufficient *NAA15* iPSC cells differentiate into cardiomyocytes, unlike *NAA15*-null iPSC cells, presumably due to altered composition of NatA. Mass spectrometry (MS) analyses reveal ~80% of identified iPSC cell NatA targeted proteins displayed partial or complete Nt-acetylation. Between null and haploinsufficient *NAA15* cells Nt-acetylation levels of 32 and 9 NatA-specific targeted proteins were reduced, respectively. Similar acetylation loss in few proteins occurred in *NAA15*R276W iPSCs. In addition, steady-state protein levels of 562 proteins were altered in both null and haploinsufficient *NAA15* cells; eighteen were ribosomal-associated proteins. At least four proteins were encoded by genes known to cause autosomal dominant CHD.

Conclusions: These studies define a set of human proteins that requires a full NAA15 complement for normal synthesis and development. A 50% reduction in the amount of NAA15 alters levels of at least 562 proteins and Nt-acetylation of only 9 proteins. One or more modulated proteins are likely responsible for NAA15-haploinsufficiency mediated CHD. Additionally, genetically engineered iPSC cells provide a platform for evaluating the consequences of amino acid sequence variants of unknown significance on NAA15 function.

Graphical Abstract



Subject Terms:

Congenital Heart Disease; Genetics; Proteomics; Stem Cells

Keywords

Induced pluripotent stem cells; CRISPR/Cas9 gene editing; stem cell; gene mutation; protein synthesis; NAA15; NatA complex; N-terminal acetylation

INTRODUCTION

Congenital heart disease (CHD), which affects about 1% of newborns, reflects defects in the developmental program of the heart.^{1, 2} Defining the genetic causes of CHD will provide new insights into mechanisms of cardiac development that may eventually benefit CHD patients and their families. Whole exome sequencing (WES) of CHD probands and their parents have enabled the identification of recurrent damaging variants in multiple genes that likely are critical for normal cardiac development.³ Functional studies are necessary to further define these gene functions and the pathogenetic mechanisms of damaging variants. While loss of function (LoF) variants infer that haploinsufficiency of the encoded protein contributes to CHD, the consequence of missense variants on protein function are less readily interpreted, often leading to classification of these as variants of unknown significance (VUS) although some may contribute to the CHD.²

We previously reported two CHD patients with *de novo* heterozygous LoF variants in *NAA15*, which encodes a protein subunit of the N-terminal (Nt) acetyltransferase (NAT) complex, NatA.²⁻⁴ In addition to CHD, these patients had extra-cardiac disorders including neurodevelopmental deficits.^{2, 3} Prior studies have reported damaging *NAA15* variants in patients with other congenital malformations and neurodevelopmental abnormalities.⁵⁻⁷

Acetylation of the N-terminus of proteins is a prevalent modification that occurs in approximately 85% of yeast and human proteins.^{8, 9} The effect of N-terminal acetylation on proteins is diverse, and includes changes to protein stability, complex formation, protein folding, and aggregation.¹⁰ The NatA complex, one of eight NAT types, is essential in most, if not all, eukaryotes and is responsible for the majority of N-terminal acetylation.^{10, 11} This complex binds the ribosome and was shown to acetylate nascent polypeptide chains at

specific N-terminal amino acids (Ser-, Thr-, Ala-, Val-, Gly- and Cys-) after the initiating methionine (iMet) is removed.^{8, 10, 12, 13} The NatA complex is formed by the catalytic subunit NAA10 and the auxiliary unit NAA15. HYPK is a chaperone protein that attaches to the NatA complex along with subunit NAA50.^{14–17} NAA50 also forms part of the NatE complex, which displays a distinct substrate specificity compared to NatA.^{10, 18} NAA15 is a subunit of both the NatA and NatE complexes, and its role is to position the catalytic subunits in close vicinity to the nascent polypeptides; in the case of NAA10, it also modulates its substrate specificity.^{10, 19–21} Additionally, NAA15 interaction with NAA10 and HYPK has been implicated in regulation of protein folding and Nt-acetylation fidelity.^{15, 22} Abnormal NatA complex function has been previously associated with human cancers, and neurological disorders.^{5, 23, 24} To date, the relationship between Nt-acetylation, NAA15, and CHD has not been investigated.

We studied human isogenic iPS cell lines that were engineered to contain *NAA15* variants identified in CHD patients and predicted to be damaging. We evaluated whether iPS cells with *NAA15* variants differentiated into cardiomyocytes. By assessing both N-terminal acetylation and protein levels by mass spectrometry we demonstrate that *NAA15* haploinsufficiency perturbs normal function of undifferentiated iPS cells. We identify proteins that require the full complement of *NAA15* to preserve the integrity of these stem cells for cardiac development.

METHODS

Data Availability.

All data and materials have been made publicly available. Further details are provided in the Major Resources Table located in Online Supplemental Materials.

Study cohort with congenital heart disease (CHD).

CHD subjects (n=4511) were recruited to the Congenital Heart Disease Network Study of the Pediatric Cardiac Genomics Consortium (CHD GENES: [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01196182) identifier [NCT01196182](https://clinicaltrials.gov/ct2/show/study/NCT01196182)) or the DNA biorepository of the Single Ventricle Reconstruction trial after approval from Institutional Review Boards as previously described.^{1,25, 26} All subjects or their parents provided informed consent. Clinical diagnoses, including cardiac and noncardiac congenital anomalies, were obtained from review of patient charts and family interview.

CRISPR gene editing and mutation confirmation.

Isogenic PGP1 iPS cells were modified using CRISPR/Cas9 technology to create *NAA15* LoF or missense mutation cell lines.²⁷ Further details provided in Online METHODS.

Label-free quantitative shotgun proteomics and data analysis.

Two independent iPSC lines of both wild type (WT) cells and *NAA15* mutant cells were grown in 10cm petri dishes until near confluency. iPSC cells were detached and using Accutase (Millipore) and collected in 15cm centrifuge tubes. Cells were centrifuged (Beckman) in a 15 ml tube at 1000rpm for 5 minutes. Cells were washed with PBS, and pelleted by

centrifugation. Eight samples of each genotype was collected and a total of 24 samples were prepared for LC-MS/MS/ analysis. Further details provided in Online METHODS.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifiers PXD017672 and PXD018013.

Bioinformatic analysis and gene ontology analysis.

Gene Ontology (GO) annotation proteome was derived using the R package clusterProfiler.²⁸ Proteins were classified by Gene ontology annotation based on biological process, molecular function and cellular component. Quantified proteins detected by mass spectrometry of iPSCs were used as a background and other parameters with default. The p-values were adjusted by Bonferroni correction for multiple testing.

Statistics

Single comparisons were analyzed by using the Student *t* test, with significance defined as $P < 0.05$. Multiple comparisons between genotypes were analyzed using one-way ANOVA with post hoc Tukey HSD, with significance defined as $P < 0.05$. For proteomics, pairwise SAM t-tests (or pairwise statistical testing using the SAM method²⁹) were performed and significant hits determined using as cutoff values a permutation based FDR of 0.01 (1000 permutations) and a background variance parameter s_0 of 1. For all experiments analyzed for statistical significance, only within-test corrections were made.

RESULTS

NAA15 variants are associated with congenital heart disease and other extra cardiac anomalies.

Whole exome sequencing of 4511 CHD patients^{1, 2, 4} identified four subjects with a rare LoF variant (allele frequency (AF) < 0.00005) in the *NAA15* gene, resulting in NAA15 haploinsufficiency (Figure 1A). Parental analyses indicated that three of these LoF variants (p.Ser761*, p.Lys336Lys fs*6 and p.Arg470*) arose *de novo* in the probands.^{2, 4} The inheritance of the p.Ala718fs variant is uncertain, as parental samples were unavailable. Among ~125,000 subjects in the gnomAD database¹², 14 *NAA15* LoF variants are reported, inferring an 8.9-fold higher frequency of *NAA15* LoF variants in CHD probands ($p=0.002$).

The CHD phenotypes included tetralogy of Fallot, heterotaxy with d-looped ventricles, transposition of the great arteries, and hypoplastic left or right heart syndrome (Figure 1C, Online Table I). In addition to cardiac anomalies, all four CHD patients with a *NAA15* LoF variant had extracardiac anomalies including seizures, neurobehavioral, ophthalmologic, auditory, or orthopedic disorders (Figure 1C and Online Table I).⁵

We also identified 15 very rare (AF $< 1.0 \times 10^{-5}$ or absent from the gnomAD database¹²) inherited *NAA15* missense variants among these 4511 CHD patients (Online Figure IA, Online Table I). In addition, one missense variant, R276W (Figure 1A, Online Figure IA–B), could not be assessed for inheritance. The frequency of rare missense alleles in the PCGC cohort was 0.0035 ($n=16$ of 4511 CHD probands), significantly higher than the

frequency observed among ~115000 Gnomad subjects (frequency=0.002; n=198; p=0.02; odds ratio=1.8). CHD probands with rare *NAA15* missense variants had notably fewer extracardiac anomalies than CHD probands with *NAA15* LoF variants (Online Table I). Hence, despite the observation that these very rare *NAA15* missense variants were transmitted from an unaffected parent, we suspected that some of them contribute to CHD.

Genetically engineered iPSCs model *NAA15* haploinsufficiency.

We introduced *NAA15* variants into the iPSC cell line, PGP1, using CRISPR/Cas9 gene editing (see METHODS) to create human iPSC cells with reduced or no *NAA15* protein. Two independent cell lines were constructed with each genotype *NAA15*^{-/-}, *NAA15*^{+/-}, and *NAA15*^{+R276W} (Figure 1A, Online Figure IA–C, Online Table II) and studied as detailed below. *NAA15* variants in iPSCs were confirmed both by Sanger sequencing and next generation sequencing of PCR amplified products (Online Figure IIA–C).

RNA expression levels for each of the two independent iPSC cell lines in *NAA15*^{+/+} and *NAA15*-mutant iPSCs, were first characterized by RNAseq. *NAA15* mRNA levels were reduced by 56% and 24% in the *NAA15*^{+/-} and *NAA15*^{+R276W} iPSCs, respectively (Figure 2A, Online Table III). There was a 99% decrease in *NAA15* mRNA in *NAA15*^{-/-} iPSCs, suggesting that these variants triggered nonsense-mediated decay (NMD) (Figure 2A). We performed mass spectrometry- (MS-) based shotgun proteomics in at least three replicates for each of the two independent lines. *NAA15* protein was near-normal levels in *NAA15*^{+/-} iPSCs but absent from *NAA15*^{-/-} iPSCs (Figure 2B). The amount of *NAA15* protein detected by mass spectrometry in *NAA15*^{+/-} iPSCs could represent a proportion of truncated *NAA15* polypeptides. To confirm this hypothesis, *NAA15* protein levels were measured in two independent biological replicates of both wildtype and *NAA15* mutant iPSCs by Western blotting (Online Figure IID–E), which supported this conclusion. *NAA15* protein was reduced in *NAA15*^{+/-} iPSCs by about 50%, and full length *NAA15* protein was not detectable in *NAA15*^{-/-} iPSCs (Online Figure IID–E).

Because *NAA15* is normally associated with components *NAA10*, *HYPK*, and *NAA50*^{15, 17, 18} (Figure 1B), we assessed the levels of these proteins in the mutant iPSCs using mass spectrometry. Both *NAA15*^{-/-} and *NAA15*^{+/-} iPSCs showed decreased protein levels of *NAA10* (Figure 2B). *NAA50*, the catalytic unit of NatE and binding partner of *NAA15* protein was reduced in *NAA15*^{-/-} iPSCs. Significant reduction of the *NAA50* protein did not occur in *NAA15*^{+/-} iPSCs (Figure 2B). Noticeably, *HYPK*, a *NAA15* binding partner and subunit of the NatA complex was significantly reduced in both *NAA15*^{+/-} and *NAA15*^{-/-} iPSCs (Figure 2B).

We further explored the functional effects of *NAA15*^{+/-} and *NAA15*^{+R276W} using a yeast assay in which the human NatA complex (hNatA) functionally replaced yeast NatA, as shown by complementation of growth phenotypes with partial rescue of the NatA-specific Nt-acetylome.⁸ HsNatA D335fs and S761* failed to rescue the temperature-sensitive growth phenotype of yNatA (Online Figure IIIA–C), suggesting that *NAA15*^{+/-} results in impaired NatA functionality. However, in this yeast assay, human HsNatA R276W rescued yeast growth suggesting at least partial NatA function in yeast (Online Figure IIIC). Noticeably, *Schizosaccharomyces pombe* *NAA15* has a tryptophan residue at position 276 while all

mammals have an arginine residue, which suggests that this residue is functionally important in human cells. As such, we suggest that rescue of yeast growth experiments might not fully provide functional assessments of mutant human NAA15 proteins.

NAA15^{+/-} and NAA15^{+R276W} iPSCs develop into contractile cardiomyocytes, while NAA15^{-/-} iPSCs have reduced viability.

To test whether *NAA15* variants have an effect on iPS cell maturation, two independent iPS cell lines for each genotype (*NAA15^{-/-}*, *NAA15^{+/-}*, and *NAA15^{+R276W}*) were differentiated into cardiomyocytes using a 13-day differentiation protocol to assess for the development and contractility of sarcomeres^{30, 31}. Both *NAA15^{+/-}* and *NAA15^{+R276W}* cells differentiated into cardiomyocytes (Figure 2C, panel a,b), however, *NAA15^{-/-}* iPSCs grew slowly (data not shown) and failed to differentiate. Cell viability of *NAA15*-mutant iPSCs was assessed (see Online METHODS). Similar to observations reported in *NAA15*/NatA yeast knockout and human knockdown cells^{19, 23, 32–34}, *NAA15^{-/-}* iPSCs had significantly retarded growth and cell death ($p = .001$; average cell death is 12% in *NAA15^{+/+}* iPSCs vs 32% in *NAA15^{-/-}* iPSCs, $n=2$). *NAA15^{+/-}* iPSC-CMs were stained with cardiac troponin T and actinin antibodies to visualize sarcomere structures (Figure 2C, panels a,b). *NAA15* mutant cell sarcomeres were indistinguishable from wildtype cell sarcomeres (Figure 2C, panels a,b, Online Figure IV, Online Table IV).

Contractility of *NAA15^{+/-}* and *NAA15^{+R276W}* iPSC-CMs was monitored by live image analysis (Online Videos I-II). Two biological replicates of *NAA15* mutant iPSC-CMs and one of two biological replicates for wildtype iPSC-CMs were incubated with GFP-actinin lentivirus to enable high fidelity tracking of sarcomere function³⁵ (Online Videos III-V). The lengths of GFP-labelled sarcomeres were measured during the contractile cycle. Contractile measurements of ‘unloaded’ (i.e. sarcomeres that were not working against resistance) demonstrated no difference in rates of sarcomere shortening or contraction between mutant and wildtype cells (Online Figure VA). To better recapitulate native cardiomyocyte architecture and mechanics, the function of ‘loaded’ sarcomeres from these cells grown as 3-dimensional micro-tissue structures^{36, 37}, where sarcomeres work against resistance, were created from wildtype iPSC-CMs and two biological replicates of *NAA15* mutant iPSC-CMs (Online Videos VI-VIII). In contrast to unloaded 2-dimensional tissue sarcomere function, image analysis show loaded *NAA15^{+/-}* iPSC-CM sarcomeres exhibited a smaller percent contraction than wildtype iPSC-CM sarcomeres (Online Figure VB).

Nt-acetylation in NAA15 haploinsufficient, null and R276W missense iPSCs.

We studied Nt-acetylation of two biological replicates in iPSCs with *NAA15* variants and wildtype using a mass spectrometry-based N-terminal enrichment assay.^{8, 38–40} NatA acetylates N-termini containing Ser-, Thr-, Ala-, Val-, Gly-, or Cys.^{10, 12} We identified a total of 989 previously annotated N termini⁴¹ in one or both of the two independent cell lines for wildtype and *NAA15* mutant iPSCs (Online Table V); approximately 650 N termini peptides were identified in each cell line (Online Table V–VI). NatA targeted sequences represented ~60% of the 989 detected N termini peptides (Online Table V–VI); no

Cys-starting peptides were detected. In *NAA15^{+/+}* iPS cells, we observed 17 N-terminal peptides that were partially acetylated (10%–95%) and 359 with complete (>95%) Nt-acetylation (Figure 3A, Online Table V–VI). Of the completely or partially Nt-acetylated peptides^{8, 10}, 98% have an alanine, threonine or serine at position 2 (Figure 3B, Online Table VI–VII). However, peptides with partial acetylation have a much higher fraction of N-terminal glycine or valine residues (Figure 3C, Online Table VI–VII). The preference for acetylation of proteins with N-terminal alanine and serine residues was preserved in *NAA15*-mutant iPSCs as observed in previous studies (Online Table VI–VII).²³

Protein Nt-acetylation in *NAA15* mutant iPSCs was compared with Nt-acetylation in wildtype iPSCs. Only a limited number of putative NatA-type N termini substrates had >10% difference in acetylation between wildtype and mutant iPSCs (Figure 3, Table 1, Online Figure VI). In *NAA15^{-/-}* iPSCs 32 proteins showed these changes (Figure 3B–D, Table 1). Notably, all proteins that had partial acetylation in WT iPS cells lacked Nt-acetylation in *NAA15^{-/-}*. Nine of these proteins were altered in *NAA15^{+/-}* iPSCs and eight had altered Nt-acetylation in *NAA15^{+/-}R276W* iPSCs (Figure 3E–F, Table 1).

All proteins with reduced acetylation in mutant iPSCs displayed NatA-type substrate specificity (Table 1–2, Online Table V). Greater than 50% of proteins with altered Nt-acetylation had an alanine residue at position 2 and >30% also had an alanine residue at position 3 (Table 2).²³ The distribution of N-terminal residues in proteins with altered Nt-acetylation did not differ significantly from the distribution of N-terminal residues in proteins with unchanged N-terminal acetylation in these cells (not shown). We deduced that *NAA15* variants or deficiency altered N-terminal acetylation in a small number of NatA substrate proteins and with no preference for an extended N-terminal amino acid substrate specificity.

We noticed that some proteins reduced in Nt-acetylation (Figure 3D) have functions in cellular proliferation and survival,^{42–46} perhaps accounting for the deficits in growth and viability of *NAA15^{-/-}* iPSC.^{42–46 42–46 41–45} To understand whether the change in Nt-acetylation status compromised the function of these proteins and contributed to the phenotype of *NAA15^{-/-}* iPSCs, we used a combination of shotgun mass spectrometry-based proteomics and pairwise comparison by SAM t-test (FDR=0.01) (Online Table VIII–X). Nine proteins with reduced acetylation correlated with changes in protein expression in *NAA15^{-/-}* iPSCs (Figure 3D, Figure 4A). Among these, measurable acetylation occurred in three proteins in *NAA15^{+/-}* and in one protein in *NAA15^{+/-}R276W* iPSCs (Figure 3E–F).

Altered protein expression due to NAA15 haploinsufficiency or deficiency.

Using shotgun proteomics we studied two independent undifferentiated iPS cell lines for each *NAA15* mutant genotype and at least three technical replicates for each sample to assess protein levels (Online Table VIII). Among 5196 proteins identified by mass spectrometry, extracts from *NAA15^{+/-}* and/or *NAA15^{-/-}* iPSCs revealed a total of 1209 proteins that were differentially expressed (pairwise comparison by t-test) compared to wildtype iPSCs (Figure 4A–B, Online Figure VII, Online Table VIII–X). Five hundred sixty-two proteins were differentially expressed in both *NAA15^{+/-}* and *NAA15^{-/-}* iPSCs compared to wildtype iPSCs and 505 proteins in *NAA15^{-/-}* iPSCs but not *NAA15^{+/-}* iPSCs

(Figure 4B, Online Figure VII, Online Table VIII–X). More than 60% of differentially expressed proteins had a NatA-target residue at position 2 of the mature protein (Online Table XI). NatA-target residues, alanine and serine, were frequently observed at position 2. Similar frequencies are seen in proteins with no expression changes (Online Table XI). Despite the large numbers of differentially expressed proteins, there were very few, if any, significant differences in mRNA levels, as assessed by RNAseq analysis (Online Table III–IV).

Among differentially expressed proteins identified in *NAA15^{+/-}* iPSCs and/or *NAA15^{-/-}* iPSCs, the levels of HYPK, a component of the NatA complex that interacts with NAA15, were significantly reduced (Figure 1C).^{14–17, 22} A total of 29 out of 54 known interactors of HYPK were detected in the wildtype iPSCs.⁴⁷ Although HYPK was decreased in mass spectrometry measurements, there was no enrichment of HYPK interacting proteins that were differentially expressed. Only 6 of 29 known HYPK interacting proteins were decreased in *NAA15*-mutant iPSCs ($p=0.3$; Online Table IX–X).

To test the hypothesis that differentially expressed proteins shared a common function, the 562 proteins differentially expressed in both *NAA15^{+/-}* and *NAA15^{-/-}* iPSCs were analyzed by gene ontology analysis using clusterProfiler (implemented in R; METHODS)²⁸. Of the three gene ontology enrichment processes that were performed we found no enrichment after Bonferroni correction for proteins with shared biological function or molecular processes; however, 41 ribosomal (-associated) proteins with altered expression levels were observed ($p = 5e-06$), Online Table XII, Figure 4C). These proteins included the catalytic subunit of the NatA complex, NAA10, and ribosomal proteins of the 60S large and 40S small ribosome subunits that function in protein synthesis (Figure 1B, Online Table XII, Figure 4D–E).^{48–52}

Ribosomal protein deficiency due to NAA15 haploinsufficiency.

Twenty-nine ribosomal proteins of the 84 proteins that comprise 60S or 40S ribosomal subunit^{53, 54} were affected in *NAA15^{+/-}* and/or *NAA15^{-/-}* iPSCs; 11 60S large ribosomal proteins and 6 small ribosomal proteins were affected in both *NAA15^{+/-}* and *NAA15^{-/-}* iPSCs (Online Table IX–X, Figure 4D). Twelve ribosomal proteins were differentially expressed in *NAA15^{-/-}* iPSCs, but not in *NAA15^{+/-}* iPSCs (Online Table IX–X).

NAA15 facilitates binding of the NatA complex to the ribosome^{55–58}. The general docking site for NatA resides at the ribosomal exit tunnel near large ribosomal unit RPL25 (L23A), a general docking platform for various factors that are transiently associated with ribosomes.^{55, 57, 59–61} Dysregulated ribosomal proteins observed in both *NAA15^{+/-}* and *NAA15^{-/-}* iPSCs clustered near the NatA docking site (Figure 5A–C). Affected small ribosomal subunits (RPS18, RPS25, RPS16, RPS19, RPS20) were clustered at the head of the small ribosomal subunit⁵³ (Figure 5C). Large ribosomal units⁵³ (RPL5, RPL13, RPL15, RPL19, RPL30, RPL31, RPL35, RPL23A, RPL39) with low protein levels were clustered at the exit tunnel of the ribosome near the NAA15/NatA docking site (Figure 5B).

Fetal heart expression and association with congenital heart disease of proteins perturbed in NAA15-haploinsufficient iPSC cells.

Because extensive protein expression studies of human and mouse fetal heart are not available, we assessed expression of the RNAs encoding these proteins in the mouse and human fetal heart⁶²; approximately 50% of the 562 proteins are expressed in the mouse and human fetal heart (unpublished). Of these 562 proteins, four (DHCR7, MAP2K2, NSD1 and RPL5; Online Table XIII) are encoded by genes known to cause autosomal dominant CHD (Online Table XIII). RNA encoding all four of these proteins are found in multiple cell types in the human fetal heart (Online Figure VIII).

DISCUSSION

Exome sequence analyses of 4511 CHD subjects with extra-cardiac phenotypes identified 20 subjects with rare inherited or *de novo* variants that are predicted to perturb NAA15 (Table S1, Figure S1, Figure 1A). Similar *NAA15* variants occur in patients with neurologic abnormalities^{5, 7}. We explored the consequences of *NAA15* variants using human iPSC cells by both transcriptome (RNAseq) and proteomic analyses (Figure 2; Online Figure II, Online Table III–IV, VIII–X). *NAA15*^{±/±} cells had ~50% of normal levels but like *NAA15*^{±/±}*R276W* iPSCs, these differentiated into cardiomyocytes. *NAA15*^{±/±} iPSC-CMs displayed normal unloaded contractility, but impaired contractility when loaded (i.e. working against resistance; Online Figure V). *NAA15*^{-/-} iPSCs produced no NAA15 and like yeast cells lacking NAA15, grew slowly (Figure 2D).¹⁹ NAA15-deficiency had minimal effect on the transcriptome but significantly altered the proteome consistent with its role in protein modification.

Because NAA15 is a component of the NatA complex, we assessed the level of N-terminal protein acetylation in iPSCs carrying *NAA15* variants using positional proteomics. Approximately 650 proteins with N-terminal sequences that were predicted to be NatA targets could be assessed. Only 32 and 9 proteins had altered Nt-acetylation in NAA15 null and haploinsufficient cells, respectively. By contrast, levels of 562 proteins were altered in both NAA15-null and NAA15-haploinsufficient cells suggesting a possible role for NAA15 and NatA in translation efficiency and/or protein stability in addition to Nt-acetylation. If NAA15 was the only available ribosome anchor for NAA10 and the NatA complex in human cells as it is in yeast cells^{8, 55, 58}, NAA15 null cells should present with a complete lack of Nt-acetylation of all NatA substrates. The explanation for the partial effect observed may be due to the presence of the NAA15 paralogue NAA16 in human cells; however, the low expression of NAA16 in iPSC cells suggests the need for additional studies to measure whether the paralogue to NAA15 had any effect on N-terminal acetylation. It has been reported that NAA16-NAA10 complexes may perform NatA-type Nt-acetylation in human cells and partially act as a backup system for the NAA15-NAA10 NatA complex.⁶³

The levels of 18 of the 81 identified ribosomal proteins were altered in NAA15-haploinsufficient cells. NAA15 anchors the NatA complex to the ribosomal large subunit via RPL25 (L23A).^{55–58, 64} Consistent with this model, some of these affected proteins clustered near the NAA15/NatA docking site. Presumably, NAA15 binding to one or more of these proteins⁵⁶ protects them from subsequent degradation either during ribosome

assembly or subsequent ribosomal activity. Previous studies have demonstrated that ribosomal subunit deficiency leads to a variety of pathologic states.^{48, 50, 51, 65} Damaging human variants in ribosomal subunits RPS19, RPL5⁴, and RPL35A have been identified in CHD patients⁴ (PCGC consortium data not shown) and also contribute to Diamond-Blackfan anemia^{50, 66, 67}, a disorder that is accompanied by CHD in 30% of patients.⁶⁸

Among differentially expressed proteins in NAA15 haploinsufficient cells, four (DHCR7, MAP2K2, NSD1 and RPL5) are encoded by known CHD genes (Online Table XIII, Online Figure VIII), and are most likely responsible for NAA15-haploinsufficient mediated CHD.^{4, 69,70, 71} As the precise mechanisms by which each haploinsufficiency of each of these genes is defined we anticipate the mechanism(s) will be shared by NAA15 haploinsufficient patients.

Exome analyses of CHD probands also identified 16 *NAA15* missense variants, including R276W, a variant of uncertain significance. iPS cells carrying the *NAA15*R276W variant had eight proteins with altered Nt-acetylation, of which four had altered Nt-acetylation in NAA15-haploinsufficient cells. These data suggest that the R276W variant likely impairs NAA15 function, and consequently NatA function, and like heterozygous LoF variants contributes to CHD. Future studies of protein levels in genetically engineered iPS cells carrying other *NAA15* variants found in CHD probands, will help identify those variants contributing to cardiac abnormalities and those variants that are likely benign.

In this study, we use genetically engineered iPS cell models to recapitulate *NAA15* LoF and missense variants discovered in CHD patients. We observe that variants in *NAA15* causing *NAA15* haploinsufficiency or deficiency result in either a reduction or complete depletion of NAA15 and NatA protein levels. We find that aberrant protein expression of NAA15 alters Nt-acetylation of a small number of proteins. Most importantly, a reduction of the NAA15 protein interrupts its interaction with the ribosome. The failed interaction causes ribosomal deficiency and mis-expression of a large number of proteins most likely involved in heart development. Protein expression changes are possibly due to obstructed ribosomal machinery and defects in protein synthesis. One or more proteins affected in both *NAA15*-mutant iPSCs have been documented as a cause of congenital heart defects, and are likely candidates required for normal cardiac developmental processes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors would like to thank Paula Montero Llopis of the MicRoN (Microscopy Resources On the North Quad) core for her support and assistance, specifically for collecting images in Figure 2. We appreciate the study participants and their families, without whom this work would not have been possible.

SOURCES OF FUNDING

We gratefully acknowledge the NHLBI Pediatrics Cardiac Genomics Consortium (PCGC) and Cardiovascular Development Consortium (CVDC) investigators (B.D.G, W.K.C., E.G., G.A.P., R.P.L., M.B., H.J.Y., B.G.B, C.E.S, J.G.S) for their support and expertise in cardiovascular development. Funding support for this study was provided by grants to the PCGC and CVDC by the US National Heart Lung and Blood Institute (UM1 HL098123 (B.D.G.),

U01 HL098153 (E.G.), UM1HL098179 (B.G.B), R01 HL151257(C.E.S), 1UM1HL098166 (J.G.S)), SFARI and the JPB Foundation (W.K.C.), Harvard Medical School Diversity and Inclusion Dean's Postdoctoral Fellowship (T.W.), Stanley J. Sarnoff Cardiovascular Research Foundation (W.T.), HHMI Medical Research Fellowship (M.Y.J.), FAPESP number 19/11921-1 and 18/13706-0 (G.V.), K.Z. acknowledges fellowship support from the American Heart Association, award number 17PRE33660967, John S. Ladue Memorial Fellowship at Harvard Medical School (Y.K.), S.V and T.A were funded by the Research Council of Norway through a FRIPRO mobility grant 261981, which is cofounded by the European Union's Seventh Framework Programme under Marie Curie grant agreement no 608695. Funding was supported by the Engineering Research Centers Program of the National Science Foundation, no. EEC-1647837 (C.S.C.). The Research Council of Norway (project 249843), the Norwegian Health Authorities of Western Norway (project F-12540), and the Norwegian Cancer Society. K.G. acknowledges support from The Research Foundation - Flanders (FWO), project number G008018N. Funding was supported by the Howard Hughes Medical Institute (C.E.S.), and Fondation Leducq 16 CVD 03 (J.G.S.).

Nonstandard Abbreviations and Acronyms:

CHD	Congenital Heart Disease
iPSC	induced pluripotent stem cells
iPSC-CMs	induced pluripotent stem cell-derived cardiomyocytes
LoF	loss of function
PGP1	Personal genome project 1
GFP	green fluorescent protein
Nt	N-terminal

REFERENCES

1. van der Linde D, Konings EE, Slager MA, et al. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol.* 2011;58:2241–7. [PubMed: 22078432]
2. Homsy J, Zaidi S, Shen Y, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science.* 2015;350:1262–6. [PubMed: 26785492]
3. Zaidi S, Choi M, Wakimoto H, et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature.* 2013;498:220–3. [PubMed: 23665959]
4. Jin SC, Homsy J, Zaidi S, et al. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nat Genet.* 2017;49:1593–1601. [PubMed: 28991257]
5. Cheng H, Dharmadhikari AV, Varland S, et al. Truncating Variants in NAA15 Are Associated with Variable Levels of Intellectual Disability, Autism Spectrum Disorder, and Congenital Anomalies. *Am J Hum Genet.* 2018;102:985–994. [PubMed: 29656860]
6. Guo H, Wang T, Wu H, et al. Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. *Mol Autism.* 2018;9:64. [PubMed: 30564305]
7. Zhao JJ, Halvardson J, Zander CS, et al. Exome sequencing reveals NAA15 and PUF60 as candidate genes associated with intellectual disability. *Am J Med Genet B Neuropsychiatr Genet.* 2018;177:10–20. [PubMed: 28990276]
8. Arnesen T, Van Damme P, Polevoda B, et al. Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proceedings of the National Academy of Sciences.* 2009;106:8157–8162.
9. Drazic A, Myklebust LM, Ree R, et al. The world of protein acetylation. *Biochim Biophys Acta.* 2016;1864:1372–401. [PubMed: 27296530]
10. Aksnes H, Drazic A, Marie M, et al. First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. *Trends Biochem Sci.* 2016;41:746–760. [PubMed: 27498224]

11. Aksnes H, Ree R and Arnesen T. Co-translational, Post-translational, and Non-catalytic Roles of N-Terminal Acetyltransferases. *Mol Cell*. 2019;73:1097–1114. [PubMed: 30878283]
12. Frottin F, Martinez A, Peynot P, et al. The proteomics of N-terminal methionine cleavage. *Mol Cell Proteomics*. 2006;5:2336–49. [PubMed: 16963780]
13. Arnold RJ, Polevoda B, Reilly JP, et al. The Action of N-terminal Acetyltransferases on Yeast Ribosomal Proteins. *Journal of Biological Chemistry*. 1999;274:37035–37040.
14. Choudhury KR, Raychaudhuri S and Bhattacharyya NP. Identification of HYPK-interacting proteins reveals involvement of HYPK in regulating cell growth, cell cycle, unfolded protein response and cell death. *PLoS One*. 2012;7:e51415. [PubMed: 23272104]
15. Gottlieb L and Marmorstein R. Structure of Human NatA and Its Regulation by the Huntingtin Interacting Protein HYPK. *Structure*. 2018;26:925–935 e8. [PubMed: 29754825]
16. Raychaudhuri S, Sinha M, Mukhopadhyay D, et al. HYPK, a Huntingtin interacting protein, reduces aggregates and apoptosis induced by N-terminal Huntingtin with 40 glutamines in Neuro2a cells and exhibits chaperone-like activity. *Hum Mol Genet*. 2008;17:240–55. [PubMed: 17947297]
17. Weyer FA, Gumiero A, Lapouge K, et al. Structural basis of HypK regulating N-terminal acetylation by the NatA complex. *Nat Commun*. 2017;8:15726. [PubMed: 28585574]
18. Deng S, Magin RS, Wei X, et al. Structure and Mechanism of Acetylation by the N-Terminal Dual Enzyme NatA/Naa50 Complex. *Structure*. 2019.
19. Mullen JR, Kayne PS, Moerschell RP, et al. Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. *The EMBO Journal*. 1989;8:2067–2075. [PubMed: 2551674]
20. Arnesen T, Anderson D, Baldersheim C, et al. Identification and characterization of the human ARD1–NATH protein acetyltransferase complex. 2005;386:433–443.
21. Liszczak G, Goldberg JM, Foyn H, et al. Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. *Nat Struct Mol Biol*. 2013;20:1098–105. [PubMed: 23912279]
22. Arnesen T, Starheim KK, Van Damme P, et al. The chaperone-like protein HYPK acts together with NatA in cotranslational N-terminal acetylation and prevention of Huntingtin aggregation. *Mol Cell Biol*. 2010;30:1898–909. [PubMed: 20154145]
23. Myklebust LM, Van Damme P, Stove SI, et al. Biochemical and cellular analysis of Ogden syndrome reveals downstream Nt-acetylation defects. *Hum Mol Genet*. 2015;24:1956–76. [PubMed: 25489052]
24. Fluge Ø, Bruland O, Akslen LA, et al. NATH, a novel gene overexpressed in papillary thyroid carcinomas. *Oncogene*. 2002;21:5056–5068. [PubMed: 12140756]
25. Gelb B, Brueckner M, Chung W, et al. The Congenital Heart Disease Genetic Network Study. *Circulation Research*. 2013;112:698–706. [PubMed: 23410879]
26. Ohye RG, Sleeper LA, Mahony L, et al. Comparison of shunt types in the Norwood procedure for single-ventricle lesions. *N Engl J Med*. 2010;362:1980–92. [PubMed: 20505177]
27. Sharma A, Toepfer CN, Ward T, et al. CRISPR/Cas9-Mediated Fluorescent Tagging of Endogenous Proteins in Human Pluripotent Stem Cells. *Current protocols in human genetics*. 2018;96:21 11 1–21 11 20.
28. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16:284–7. [PubMed: 22455463]
29. Tusher VG, Tibshirani R and Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. 2001;98:5116–21. [PubMed: 11309499]
30. Lian X, Hsiao C, Wilson G, et al. Cozzarelli Prize Winner: Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. *Proceedings of the National Academy of Sciences*. 2012;109:E1848–E1857.
31. Lian X, Zhang J, Azarin SM, et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/beta-catenin signaling under fully defined conditions. *Nat Protoc*. 2013;8:162–75. [PubMed: 23257984]
32. Arnesen T, Gromyko D, Pendino F, et al. Induction of apoptosis in human cells by RNAi-mediated knockdown of hARD1 and NATH, components of the protein N-alpha-acetyltransferase complex. *Oncogene*. 2006;25:4350–60. [PubMed: 16518407]

33. Gromyko D, Arnesen T, Rynningen A, et al. Depletion of the human Nalpha-terminal acetyltransferase A induces p53-dependent apoptosis and p53-independent growth inhibition. *Int J Cancer*. 2010;127:2777–89. [PubMed: 21351257]
34. Kalvik TV and Arnesen T. Protein N-terminal acetyltransferases in cancer. *Oncogene*. 2013;32:269–76. [PubMed: 22391571]
35. Toepfer CN, Sharma A, Cicconet M, et al. SarcTrack. *Circ Res*. 2019;124:1172–1183. [PubMed: 30700234]
36. Hinson JT, Chopra A, Nafissi N, et al. HEART DISEASE. Titin mutations in iPS cells define sarcomere insufficiency as a cause of dilated cardiomyopathy. *Science*. 2015;349:982–6. [PubMed: 26315439]
37. Legant WR, Pathak A, Yang MT, et al. Microfabricated tissue gauges to measure and manipulate forces from 3D microtissues. *Proc Natl Acad Sci U S A*. 2009;106:10097–102. [PubMed: 19541627]
38. Gevaert K, Goethals M, Martens L, et al. Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat Biotechnol*. 2003;21:566–9. [PubMed: 12665801]
39. Staes A, Impens F, Van Damme P, et al. Selecting protein N-terminal peptides by combined fractional diagonal chromatography. *Nat Protoc*. 2011;6:1130–41. [PubMed: 21799483]
40. Van Damme P, Van Damme J, Demol H, et al. A review of COFRADIC techniques targeting protein N-terminal acetylation. *BMC Proc*. 2009;3 Suppl 6:S6.
41. UniProt C UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019;47:D506–D515. [PubMed: 30395287]
42. Yan J, Walz K, Nakamura H, et al. COP9 signalosome subunit 3 is essential for maintenance of cell proliferation in the mouse embryonic epiblast. *Mol Cell Biol*. 2003;23:6798–808. [PubMed: 12972600]
43. Zhang Y, Cooke M, Panjwani S, et al. Histone h1 depletion impairs embryonic stem cell differentiation. *PLoS Genet*. 2012;8:e1002691. [PubMed: 22589736]
44. Mange A, Coyaud E, Desmetz C, et al. FKBP4 connects mTORC2 and PI3K to activate the PDK1/Akt-dependent cell proliferation signaling in breast cancer. *Theranostics*. 2019;9:7003–7015. [PubMed: 31660083]
45. Teng T, Mercer CA, Hexley P, et al. Loss of tumor suppressor RPL5/RPL11 does not induce cell cycle arrest but impedes proliferation due to reduced ribosome content and translation capacity. *Mol Cell Biol*. 2013;33:4660–71. [PubMed: 24061479]
46. Bhattacharjee RB and Bag J. Depletion of nuclear poly(A) binding protein PABPN1 produces a compensatory response by cytoplasmic PABP4 and PABP5 in cultured human cells. *PLoS One*. 2012;7:e53036. [PubMed: 23300856]
47. Choudhury KR, Bucha S, Baksi S, et al. Chaperone-like protein HYPK and its interacting partners augment autophagy. *Eur J Cell Biol*. 2016;95:182–94. [PubMed: 27067261]
48. Cheng Z, Mugler CF, Keskin A, et al. Small and Large Ribosomal Subunit Deficiencies Lead to Distinct Gene Expression Signatures that Reflect Cellular Growth Rate. *Mol Cell*. 2019;73:36–47 e10. [PubMed: 30503772]
49. Gregory B, Rahman N, Bommakanti A, et al. The small and large ribosomal subunits depend on each other for stability and accumulation. *Life Sci Alliance*. 2019;2.
50. Narla A and Ebert BL. Ribosomopathies: human disorders of ribosome dysfunction. *Blood*. 2010;115:3196–205. [PubMed: 20194897]
51. Peisker K, Braun D, Wolfle T, et al. Ribosome-associated complex binds to ribosomes in close proximity of Rpl31 at the exit of the polypeptide tunnel in yeast. *Mol Biol Cell*. 2008;19:5279–88. [PubMed: 18829863]
52. Steffen KK, McCormick MA, Pham KM, et al. Ribosome deficiency protects against ER stress in *Saccharomyces cerevisiae*. *Genetics*. 2012;191:107–18. [PubMed: 22377630]
53. Khatter H, Myasnikov AG, Natchiar SK, et al. Structure of the human 80S ribosome. *Nature*. 2015;520:640–5. [PubMed: 25901680]
54. Humphrey W, Dalke A and Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14:33–8, 27–8. [PubMed: 8744570]

55. Gautschi M, Just S, Mun A, et al. The yeast N(alpha)-acetyltransferase NatA is quantitatively anchored to the ribosome and interacts with nascent polypeptides. *Mol Cell Biol.* 2003;23:7403–14. [PubMed: 14517307]
56. Knorr AG, Schmidt C, Tesina P, et al. Ribosome-NatA architecture reveals that rRNA expansion segments coordinate N-terminal acetylation. *Nat Struct Mol Biol.* 2019;26:35–39. [PubMed: 30559462]
57. Plevoda B, Brown S, Cardillo TS, et al. Yeast N(alpha)-terminal acetyltransferases are associated with ribosomes. *J Cell Biochem.* 2008;103:492–508. [PubMed: 17541948]
58. Varland S and Arnesen T. Investigating the functionality of a ribosome-binding mutant of NAA15 using *Saccharomyces cerevisiae*. *BMC Res Notes.* 2018;11:404. [PubMed: 29929531]
59. Kramer G, Boehringer D, Ban N, et al. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat Struct Mol Biol.* 2009;16:589–97. [PubMed: 19491936]
60. Kramer G, Rauch T, Rist W, et al. L23 protein functions as a chaperone docking site on the ribosome. *Nature.* 2002;419:171–4. [PubMed: 12226666]
61. Wilson DN and Nierhaus KH. Ribosomal proteins in the spotlight. *Crit Rev Biochem Mol Biol.* 2005;40:243–67. [PubMed: 16257826]
62. Cui Y, Zheng Y, Liu X, et al. Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell Rep.* 2019;26:1934–1950 e5. [PubMed: 30759401]
63. Arnesen T, Gromyko D, Kagabo D, et al. A novel human NatA Nalpha-terminal acetyltransferase complex: hNaa16p-hNaa10p (hNat2-hArd1). *BMC Biochem.* 2009;10:15. [PubMed: 19480662]
64. Magin RS, Deng S, Zhang H, et al. Probing the interaction between NatA and the ribosome for co-translational protein acetylation. *PLoS One.* 2017;12:e0186278. [PubMed: 29016658]
65. Warren AJ. Molecular basis of the human ribosomopathy Shwachman-Diamond syndrome. *Adv Biol Regul.* 2018;67:109–127. [PubMed: 28942353]
66. Farrar JE, Nater M, Caywood E, et al. Abnormalities of the large ribosomal subunit protein, Rpl35a, in Diamond-Blackfan anemia. *Blood.* 2008;112:1582–92. [PubMed: 18535205]
67. Gazda HT, Sheen MR, Vlachos A, et al. Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *Am J Hum Genet.* 2008;83:769–80. [PubMed: 19061985]
68. Vlachos A, Osorio DS, Atsidaftos E, et al. Increased Prevalence of Congenital Heart Disease in Children With Diamond Blackfan Anemia Suggests Unrecognized Diamond Blackfan Anemia as a Cause of Congenital Heart Disease in the General Population: A Report of the Diamond Blackfan Anemia Registry. *Circ Genom Precis Med.* 2018;11:e002044. [PubMed: 29748317]
69. Machet L [Genetic diseases on the Internet: OMIM]. *Ann Dermatol Venereol.* 1998;125:645. [PubMed: 9805561]
70. Online Mendelian Inheritance in Man O. 2014.
71. Oyston J Online Mendelian Inheritance in Man. *Anesthesiology.* 1998;89:811–2. [PubMed: 9743436]
72. Genomes Project, C., Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. [PubMed: 26432245]
73. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20, 1297–1303. [PubMed: 20644199]
74. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43, 11 10 11–33. [PubMed: 25431634]
75. Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, and Liu X (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* 24, 2125–2137. [PubMed: 25552646]

76. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, and Mann M (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 13, 2513–2526. [PubMed: 24942700]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

NOVELTY AND SIGNIFICANCE

What is Known?

- NAA15 is an 866 amino acid subunit of the N-terminal (Nt) acetyltransferase complex, NatA, which acetylates the N-terminal residue of many eukaryotic proteins
- Congenital Heart Disease (CHD) patients have genetic variants in the NAA15 gene predicted to be harmful, yet there is limited knowledge of how NAA15 variants cause CHD.
- Human induced pluripotent stem cells (iPSCs) can be used as a human model to study biological consequences of genetic variants that may cause disease

What New Information Does this Article Contribute?

- Exome sequence analysis of 4511 CHD patients identified 20 subjects with rare inherited or de novo variants predicted to perturb NAA15, which may be responsible for their condition.
- NAA15 haploinsufficiency in iPS cells has little effect on the transcriptome. However, a 50% reduction of NAA15 protein does lead to impaired cardiomyocyte contractility and proteome integrity.
- Expression levels of ribosome-associated proteins and four proteins that encode genes known to cause autosomal dominant CHD are altered in NAA15 haploinsufficient and deficient iPS cells..

Congenital heart disease (CHD), which affects about 1% of newborns, reflects defects in heart development during fetal life. Defining genetic causes of CHD provides new insights into mechanisms of cardiac development that may eventually benefit CHD patients and their families. We have used induced pluripotent stem cells to model NAA15 haploinsufficiency. In iPS cells, NAA15 haploinsufficiency and deficiency cause loss of N-terminal acetylation of 9 and 32 proteins. In addition, we discovered 562 proteins with altered expression levels; eighteen proteins are ribosome-associated. Among affected proteins, four proteins are known CHD genes, and are likely candidates that require a full complement of NAA15. We conclude that CHD patients with NAA15 haploinsufficiency likely have altered ribosomal activity, which effects steady state levels of proteins essential for early cardiac development. Genetically engineered iPS cells can be used as a platform for defining the potential pathogenicity of NAA15 variants of unknown significance.

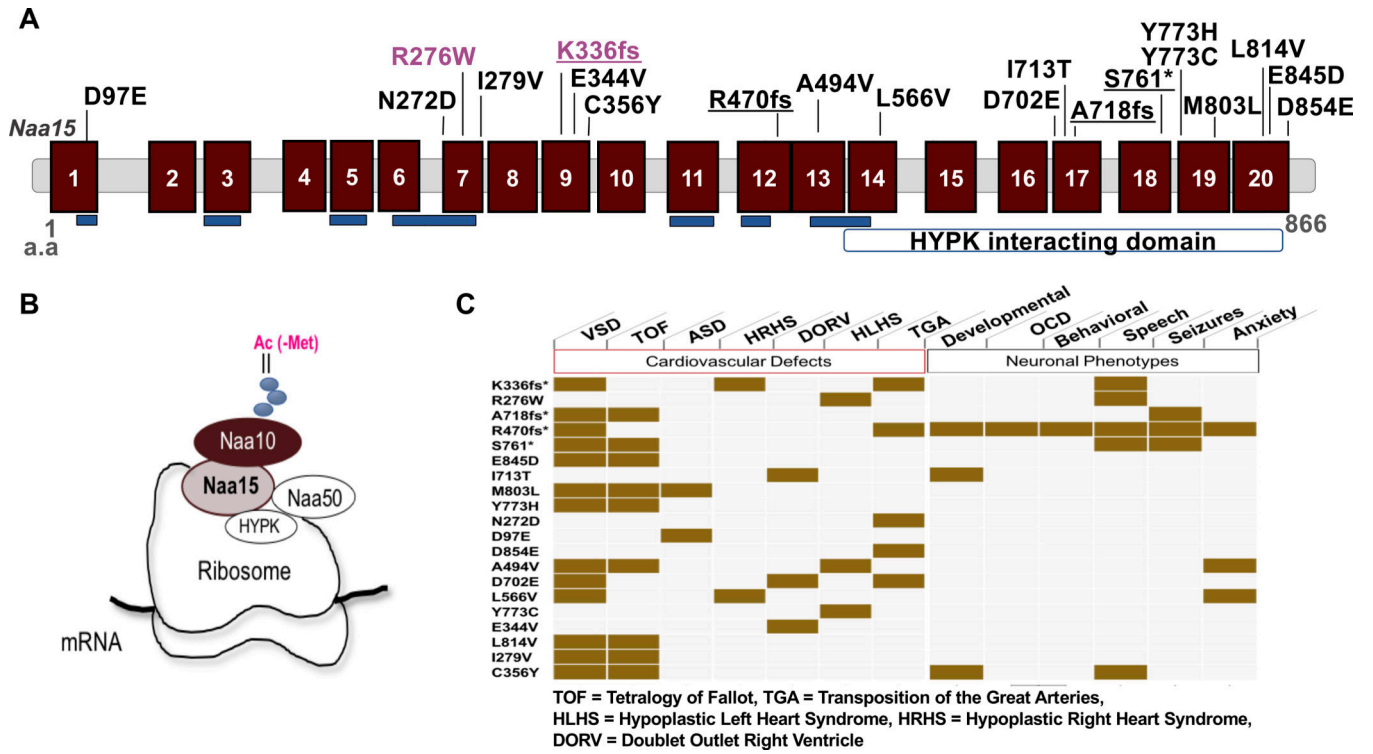


Figure 1. NAA15 variants discovered in patients with congenital heart disease patients.

A) Schematic diagram of the *NAA15* gene. The *NAA15* gene consists of tetracopeptide repeats (blue) essential for interaction with the NatA complex subunit NAA10 and a HYPK interacting domain at the C terminus. Location of variants identified in CHD patients (LoF variants are underlined) and CRISPR/Cas9-derived variants in iPSCs (magenta) are shown. B) The subunit composition of NatA and NatE complexes. NAA15 is the auxiliary unit for both complexes. C) Table of cardiovascular and neuronal clinical phenotypes of CHD patients. * indicates LoF variants.

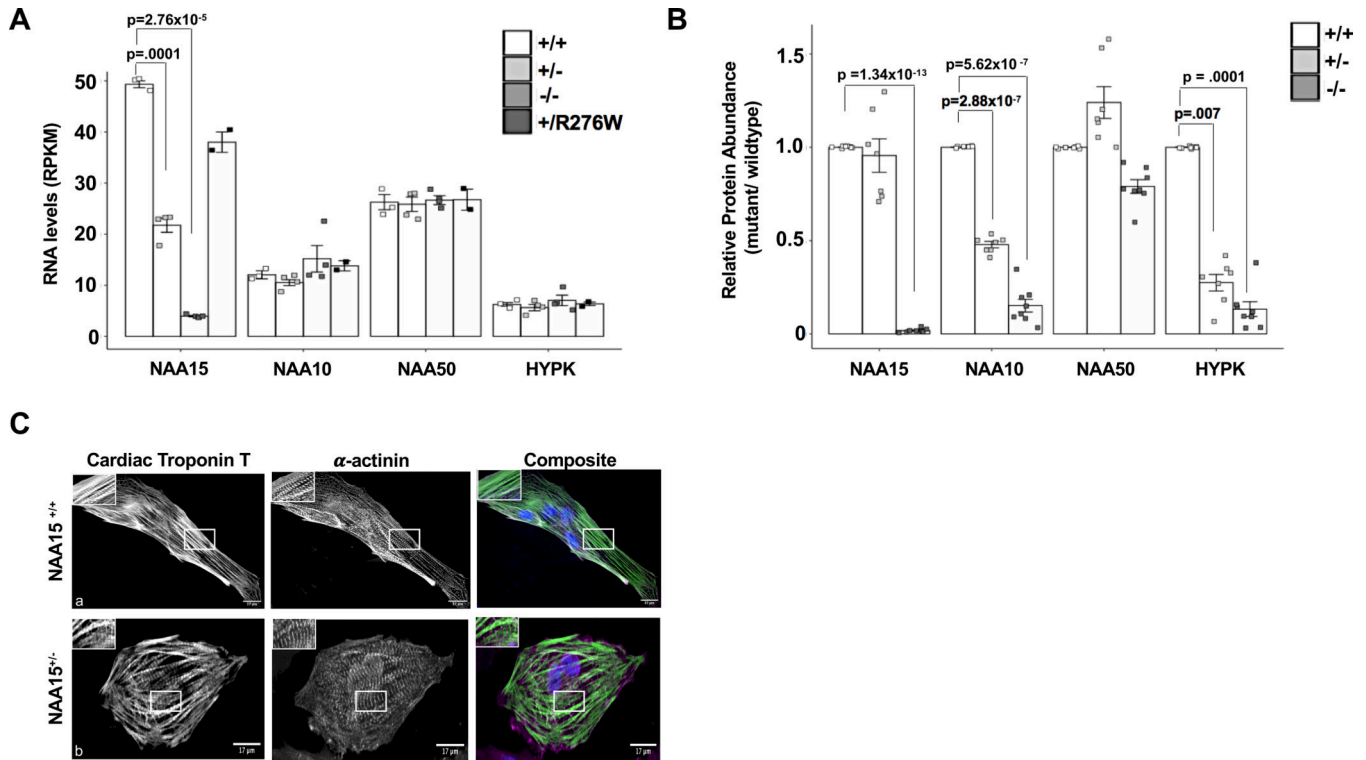


Figure 2. *NAA15*^{+/-} and *NAA15*^{+R276W} iPSCs develop into cardiomyocytes, while *NAA15* null iPSCs have reduced viability.

A) Graph shows RNA expression of the NatA components in *NAA15*-mutant iPSCs. *NAA15* RNA levels are significantly lowered in iPS cells with *NAA15*^{+/-} and *NAA15*^{-/-} variants compared to wildtype iPSCs. Data was collected from two independent cell lines for each genotype and as technical replicates for selected lines. Total cell lines analyzed: *NAA15*^{+/+} (n=3), *NAA15*^{-/-} (n=4), *NAA15*^{+/-} (n=4), and *NAA15*^{+R276W} (n=2). Significance of differences between *NAA15* mutant iPSCs and wildtype iPSCs were evaluated by Student T-tests, $p < 0.05$ (p-values were adjusted by Bonferroni correction; only significant adjusted p-values are displayed). All data points are presented and plotted as mean \pm SEM. B) Graph represents fold change ratios of relative protein levels in *NAA15* mutant iPSCs compared to wildtype iPSCs for Nat complex subunits. Relative protein abundance was quantified by MaxLFQ algorithms integrated in the MaxQuant software. Data was collected from two independent undifferentiated iPS cell lines for each *NAA15* mutant genotype and as technical replicates for each cell line. Total cell lines analyzed: *NAA15*^{+/-} (n=8), *NAA15*^{-/-} (n=8), *NAA15*^{+/-} (n=7), and *NAA15*^{+R276W} (n=8)). Significant changes indicated by p-values were calculated using pairwise comparison SAM t-test method and a permutation based FDR of 0.01 (1000 permutations) as cutoff values with a background variance parameter s_0 of 1. All data points are presented and plotted as mean \pm SEM. C) Representative images of *NAA15*^{+/+} iPSCs and *NAA15*^{+/-} iPSCs stained with cardiac troponin T antibody (green), α -actinin (magenta), and DAPI for nuclei (blue). A representative image for each cell type is presented. Normal sarcomeres were observed in both *NAA15* mutant and wildtype iPSC-CMs. Magnification, 60X; scale bar 25 μ m.

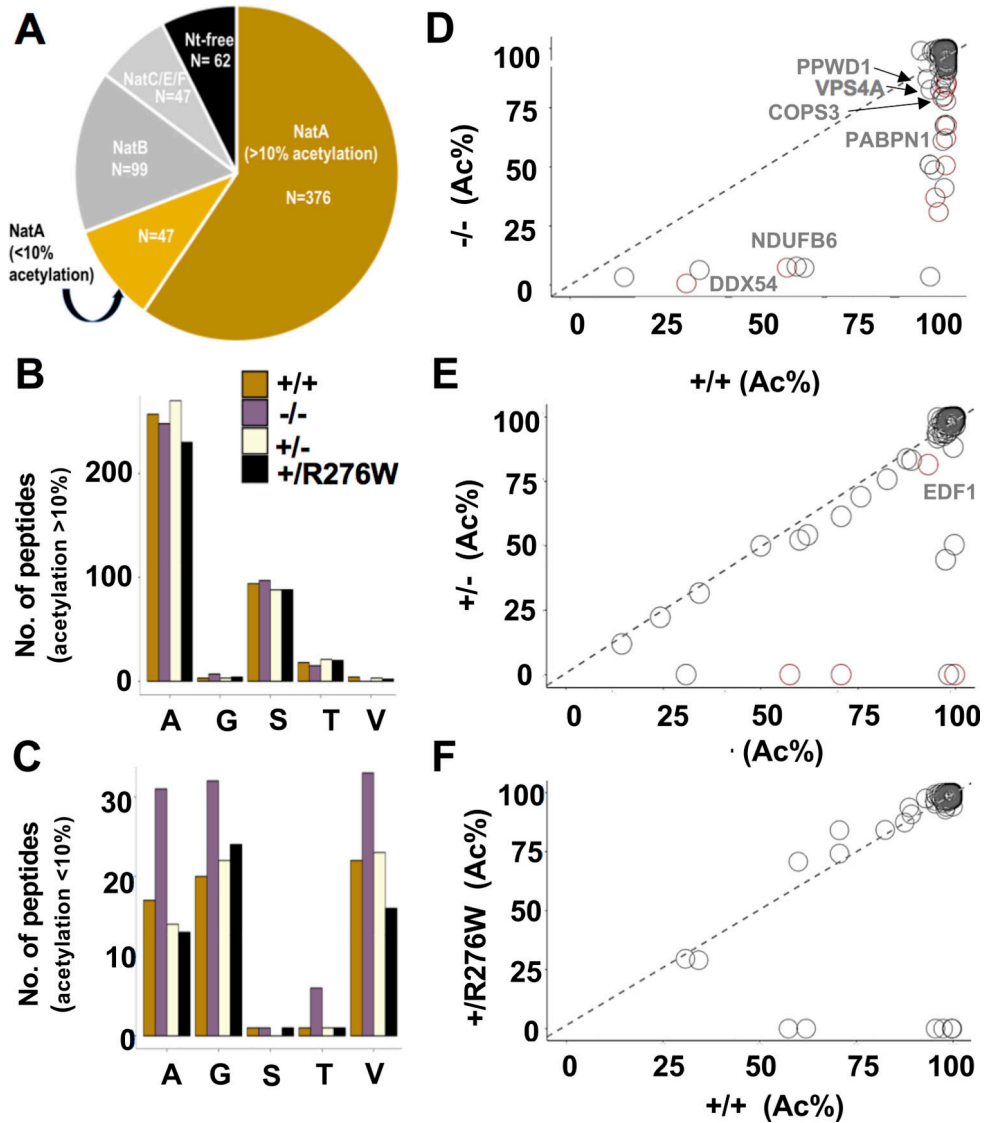


Figure 3. N-terminal acetylation in *NAA15*^{+/-} and *NAA15*^{-/-} iPSCs.

A) N-terminal peptides acetylated in *NAA15*^{+/+} iPSCs are categorized by NAT-type substrate specificity. NatA targets are displayed according to percent Nt-acetylation; Nt-acetylation >10% highlighted in gold and Nt-acetylation <10% highlighted in light gold. B-C) The number of Nt-acetylated NatA-type N-terminal peptides represented by residues A, G, S, T, or V in *NAA15* mutant and wildtype iPSCs. Data is displayed as the absolute number of peptides measured from the sum of two biological replicates for each genotype. Alanine and serine residues are preferentially acetylated in peptides with >10% Nt-acetylation; glycine and valine residues are representative of N-termini with <10% Nt-acetylation. Primary peptide summary provided in Online Table V. There were no statistically significant differences between genotypes (hypergeometric test; data not shown). D-F) Representative scatter plots display the correlation of the degrees of Nt-acetylation of all N termini identified as NatA targets in *NAA15*^{-/-}, *NAA15*^{+/-}, *NAA15*^{+/-}R276W and *NAA15*^{+/+} iPSCs. Each plot compares *NAA15* mutant iPSCs to wildtype iPSCs.

Differentially acetylated N-termini are found below the line of identity (y intercept = 0, slope =1). Peptides lose acetylation moieties in *NAA15*-mutant iPSCs. *NAA15*^{-/-} iPSCs have the largest number of peptides (n=32) with reduced Nt-acetylation. There are N-termini with reduced N-terminal acetylation in *NAA15*^{+/-} iPSCs (n=9) and *NAA15*^{+/-R276W} iPSCs (n=8). Decreased protein expression changes are highlighted in red.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

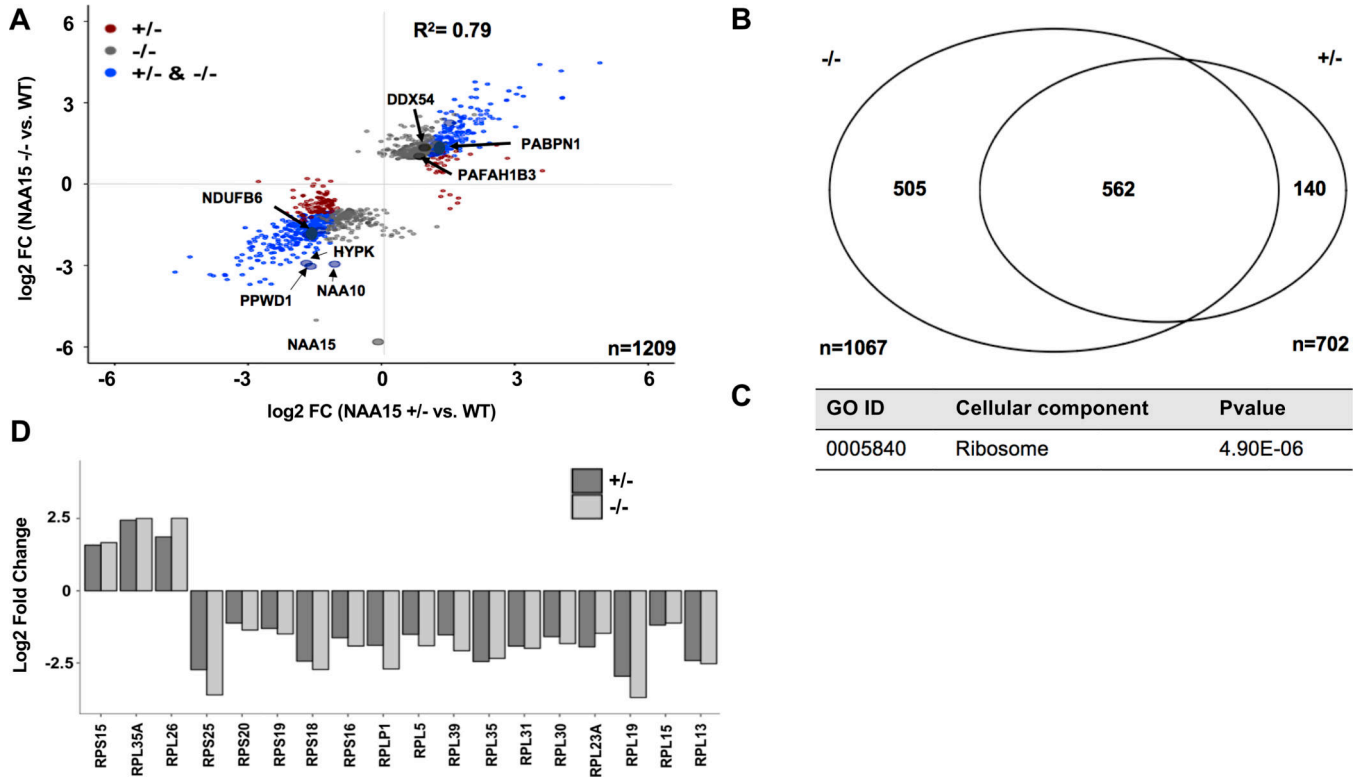


Figure 4. Proteins differentially expressed in *NAA15*^{+/-} and *NAA15*^{-/-} iPSC cells.
 A) Scatterplot displays log₂ fold changes of differentially expressed proteins in *NAA15*^{+/-} and/or *NAA15*^{-/-} iPSCs compared to wildtype iPSCs. Approximately 1200 proteins are differentially expressed, nine and three proteins of which have reduced Nt-acetylation in *NAA15*^{-/-} iPSCs and *NAA15*^{+/-} iPSCs, respectively. Data was collected from two independent undifferentiated iPS cell lines for each *NAA15* mutant genotype and at least three technical replicates for each sample. B) Comparison of differentially expressed proteins in *NAA15*^{-/-} and *NAA15*^{+/-} iPS cells. Differential expression was considered for log₂ fold change |1|. Identified proteins: 5196, Quantified proteins: 3911 proteins detected by mass spectrometry C) Gene ontology enrichment of differentially expressed proteins in both *NAA15*^{-/-} and *NAA15*^{+/-} iPSCs. A total of forty-one proteins are localized to the ribosome. Raw p-value presented in data table. D) Bar graph represents log₂ fold change of ribosomal proteins that are differentially expressed in both *NAA15*^{-/-} and *NAA15*^{+/-} iPSCs.

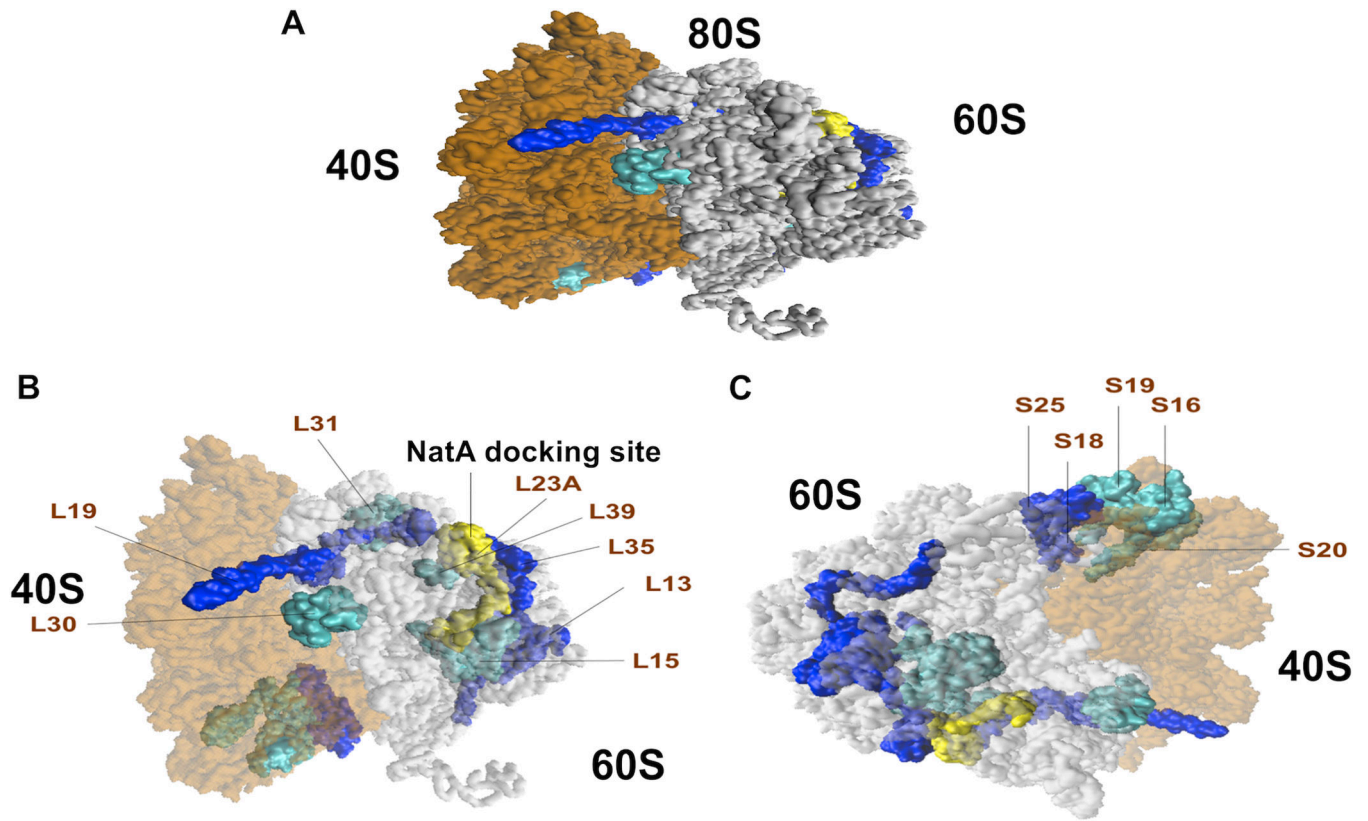


Figure 5. Ribosomal proteins of the 40S small and 60S large subunits affected in *NAA15*-mutant iPSCs.

A) Graphical representation of the human 80S ribosomes with affected 40S small and 60S large subunits in both *NAA15*^{-/-} and *NAA15*^{+/-} iPSCs. Ribosomal proteins down in expression highlighted in dark blue (\log_2 fold change < -2) and light blue ($-2 < \log_2$ fold change > -1). NatA Docking site is highlighted in yellow. B) Diagram showing location of affected 60S large subunits and c) 40S small subunits. Large ribosomal proteins are clustered and located near the NatA docking site. Small ribosomal proteins affected in their expression clustered together.

Table 1.

Proteins with altered N-terminal acetylation in wildtype and NAA15 mutant iPSCs*

Accession	Description	Nf-Acetylation percentage				Protein levels (fold change)		
		+/+	+/-	-/-	+R276W	+/- / +/+	-/- / +/+	-/- / +/+
Q9UNS2	COP9 signalosome complex subunit 3 (COP93)	99.4	99.9	83.6	99.6	1.87	2.00	2.00
Q8TDD1	ATP-dependent RNA helicase (DDX54)	30.8	0	0.7	29.6	2.00	2.46	2.46
O60869	Endothelial differentiation-related factor 1 (EDF1)	93	81.6	99	97.5	0.38	0.31	0.31
Q9NZM5	Ribosome biogenesis protein (NOP53)	70.6	0	0	84	2.83	4.92	4.92
P16401	Histone H1.5	96.7	92.8	37.1	98.2	0.20	0.22	0.22
P16403	Histone H1.2	99.5	98.4	50.6	99.5	0.27	0.29	0.29
P16402	Histone H1.3	99.6	99.5	62	99.6	0.27	0.29	0.29
P10412	Histone H1.4	97.6	93.7	30.9	98.4	0.27	0.29	0.29
O95139	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 6 (NDUFB6)	57.5	0	7.4	0	0.33	0.27	0.27
Q86U42	Polyadenylate-binding protein 2(PABPN1)	99.8	99.3	67.6	0	2.30	2.30	2.30
Q15102	Platelet-activating factor acetylhydrolase 1B subunit gamma (PAFAH1B3)	99.7	99.6	84.7	99.4	3.48	4.29	4.29
Q96BP3	Peptidylprolyl isomerase domain and WD repeat-containing protein 1 (PPWD1)	99.8	0	85.6	98.8	0.33	0.13	0.13
Q9UN37	Vacuolar protein sorting-associated protein 4 (VPS4A)	98.6	98.6	79.4	99.4	1.00	0.27	0.27
Q9HB71	Calcyclin-binding protein (CACYPBP)	99.4	99.6	67.9	100	.81	.93	.93
P14854	Cytochrome c oxidase subunit 6B1 (COX6B1)	99.2	99.2	67.3	99.2	1.62	1.62	1.62
Q02790	Peptidyl-prolyl cis-trans isomerase (FKBP4)	14.3	11.9	3.3	14.3	0.81	3.43	3.43
Q92616	eIF-2-alpha kinase activator GCN1	62	54.2	7.3	0	1.41	1.15	1.15
O76003	Glutaredoxin-3 (GLRX3)	98.2	0	85.3	94	1.15	1.07	1.07
Q00839	Heterogeneous nuclear ribonucleoprotein U (HNRNPU)	99	98.7	79.9	96.4	1.52	1.62	1.62
Q81XQ5	Kelch-like protein 7 (KLHL7)	96.5	94	48.7	98.3	0.76	0.81	0.81
Q96AG4	Leucine-rich repeat-containing protein 59 (LRRC59)	60	52.3	7.9	70.7	1.41	1.62	1.62
P62937	Peptidyl-prolyl cis-trans isomerase A (PPIA)	34.3	31.7	6.4	29	1.15	0.93	0.93
P62913	60S ribosomal protein L11 (RPL11)	99.6	99.3	77.9	99.6	1.00	0.93	0.93
P40429	60S ribosomal protein L13a (RPL13A)	95.4	92	3.5	95.3	0.76	0.76	0.76
P25398	40S ribosomal protein S12 (RPS12)	98	99	82.4	92.9	1.23	1.23	1.23

Accession	Description	Nε-Acetylation percentage				Protein levels (fold change)	
		+/+	+/-	-/-	+R276W	+/- / +/+	-/- / +/+
Q8NI27	THO complex subunit 2 (THOC2)	99.7	50.4	0	0	1.00	1.07
Q9Y5J9	Mitochondrial import inner membrane translocase subunit (TIMM8B)	95.5	96	82.7	0	1.62	1.52
O94826	Mitochondrial import receptor subunit (TOMM70A)	98.7	98	61	98.4	1.41	1.62
Q04323	UBX domain-containing protein 1 (UBXN1)	99.1	96.7	41	99.3	0.81	0.81
P31946	14-3-3 protein beta/alpha (YWHAB)	95.3	93.7	50.9	96.6	1.32	1.41
Q9NXW9	Alpha-ketoglutarate-dependent dioxygenase (ALKBH4)	97.5	44.6	0	0	N.D. [‡]	N.D. [‡]
Q9NRG0	Chromatin accessibility complex protein 1 (CHRC1)	99.4	88.3	99.7	99.1	N.D. [‡]	N.D. [‡]
Q9Y241	HIG1 domain family member 1A, mitochondrial (HIG1A)	98.9	99.4	87.3	99.3	N.D. [‡]	N.D. [‡]
P17096	HIG1 domain family member 1A, mitochondrial (HMGAI)	99.5	99.5	88.8	99.6	N.D. [‡]	N.D. [‡]

* Altered acetylation levels = >10% deviation

[‡] Protein levels could not be determined or quantified

Table 2.

Second and third amino acid identities of peptides with altered Nt-acetylation in NAA15 mutant iPSCs

Amino Acid Residue	Number of Peptides		
	-/-	+/-	+R276W
AA	10	6	5
AD	0	1	0
AE	5	1	1
AQ	2	0	1
AS	2	0	0
SE	5	0	0
SG	1	0	0
SS	1	0	0
ST	1	0	0
TA	1	0	0
TG	1	1	1
TK	1	0	1
TM	1	0	0
TT	1	0	0
VN	1	0	0

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript