Amino-terminal acetylation of proteins: role in human disease and biology

Gholson J. Lyon, M.D. Ph.D.

CSH Cold Spring Harbor Laboratory
Stanley Institute for Cognitive Genomics

Acknowledgments

STANLEY INSTITUTE FOR COGNITIVE GENOMICS
COLD SPRING HARBOR LABORATORY

Michael Klingener
Jason O'Rawe
Yiyang Wu
Michael Schatz
Giuseppe Narzisi

Max Doerfel

Thomas Arnesen
Rune Evjenth
Johan R. Lillehaug

our study families

THE UNIVERSITY OF UTAH

UTAH Foundation for BIOMEDICAL research

I moved to Utah in July 2009 to find at least one new human disease, thus revealing new biology.

- July 2009-December 2009: Attended weekly genetics case conference in which 10-30 genetic cases are presented weekly, led by Dr. Alan Rope and attended by Drs. John Carey and John Opitz.
- There are indeed MANY idiopathic disorders not described in the literature, many of which have neuropsychiatric manifestations. I thought about hundreds of such cases, looking for the ideal first family to sequence.

Discovering a new syndrome and its genetic basis.

ARTICLE

Using VAAST to Identify an X-Linked Disorder Resulting in Lethality in Male Infants Due to N-Terminal Acetyltransferase Deficiency

Alan F. Rope,[1] Kai Wang,[2,19] Rune Evjenth,[3] Jinchuan Xing,[4] Jennifer J. Johnston,[5] Jeffrey J. Swensen,[6,7] W. Evan Johnson,[8] Barry Moore,[4] Chad D. Huff,[4] Lynne M. Bird,[9] John C. Carey,[1] John M. Opitz,[1,4,6,10,11] Cathy A. Stevens,[12] Tao Jiang,[13,14] Christa Schank,[8] Heidi Deborah Fain,[15] Reid Robison,[15] Brian Dalley,[16] Steven Chin,[6] Sarah T. South,[1,7] Theodore J. Pysher,[6] Lynn B. Jorde,[4] Hakon Hakonarson,[2] Johan R. Lillehaug,[3] Leslie G. Biesecker,[5] Mark Yandell,[4] Thomas Arnesen,[3,17] and Gholson J. Lyon[15,18,20,*]

The American Journal of Human Genetics *89*, 1–16, July 15, 2011

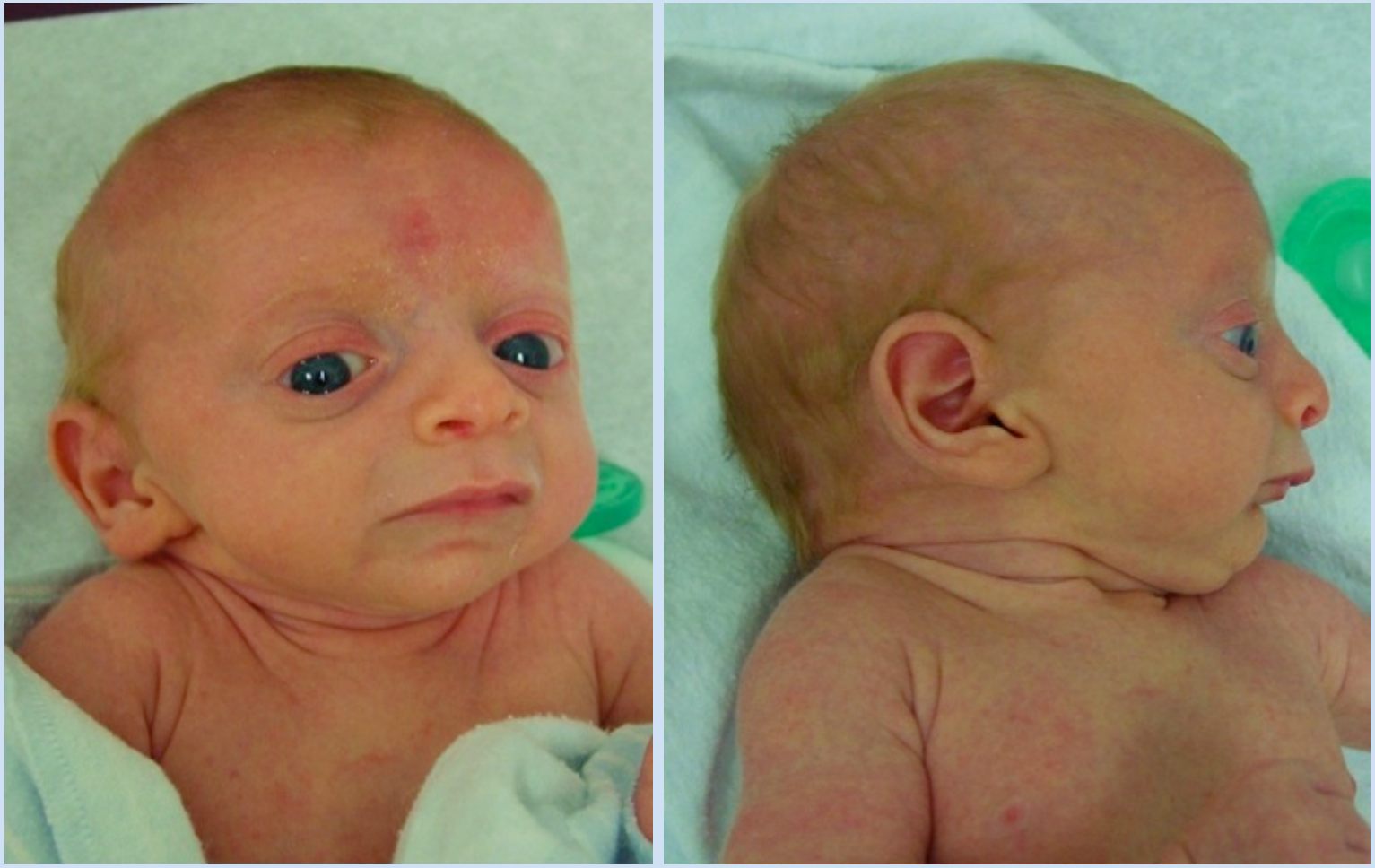
I met the entire family on March 29, 2010

Photo of mother
with son in late 1970's

This is the first boy in the late 1970's.

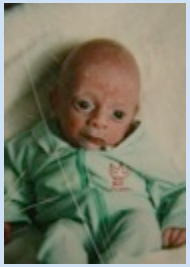
First boy. Called "a little old man" by the family. Died around ~1 year of age, from cardiac arrhythmias.

This is the "Proband" photograph presented at Case Conference.

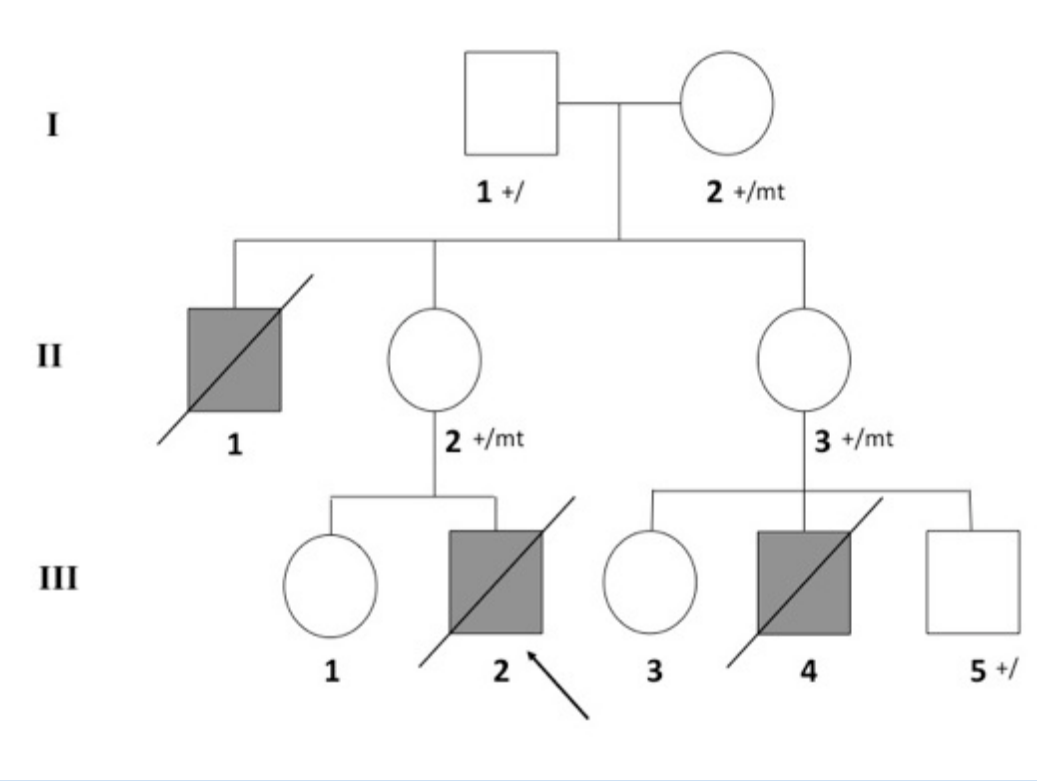


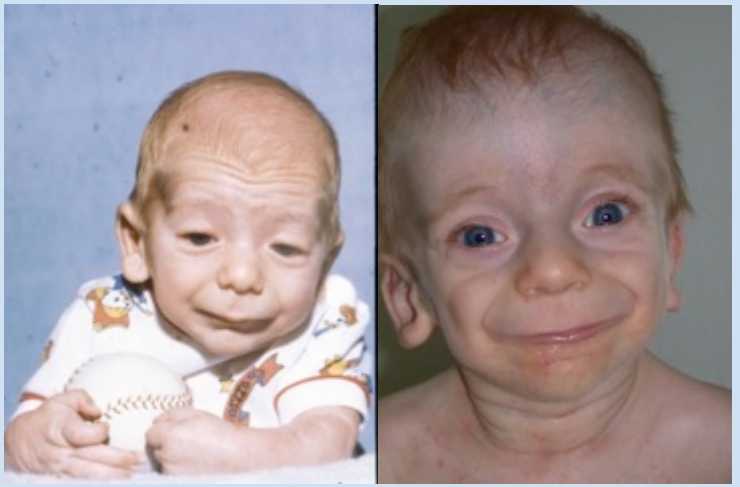
prominence of eyes, down-sloping palpebral fissures, thickened eyelids, large ears, beaking of nose, flared nares, hypoplastic nasal alae, short columella, protruding upper lip, micro-retrognathia

Family now in October 2011, with five mutation-positive boys dying from the disease.



An unrelated second family was also identified, due to sharing the same genotype, i.e. the same mutation.

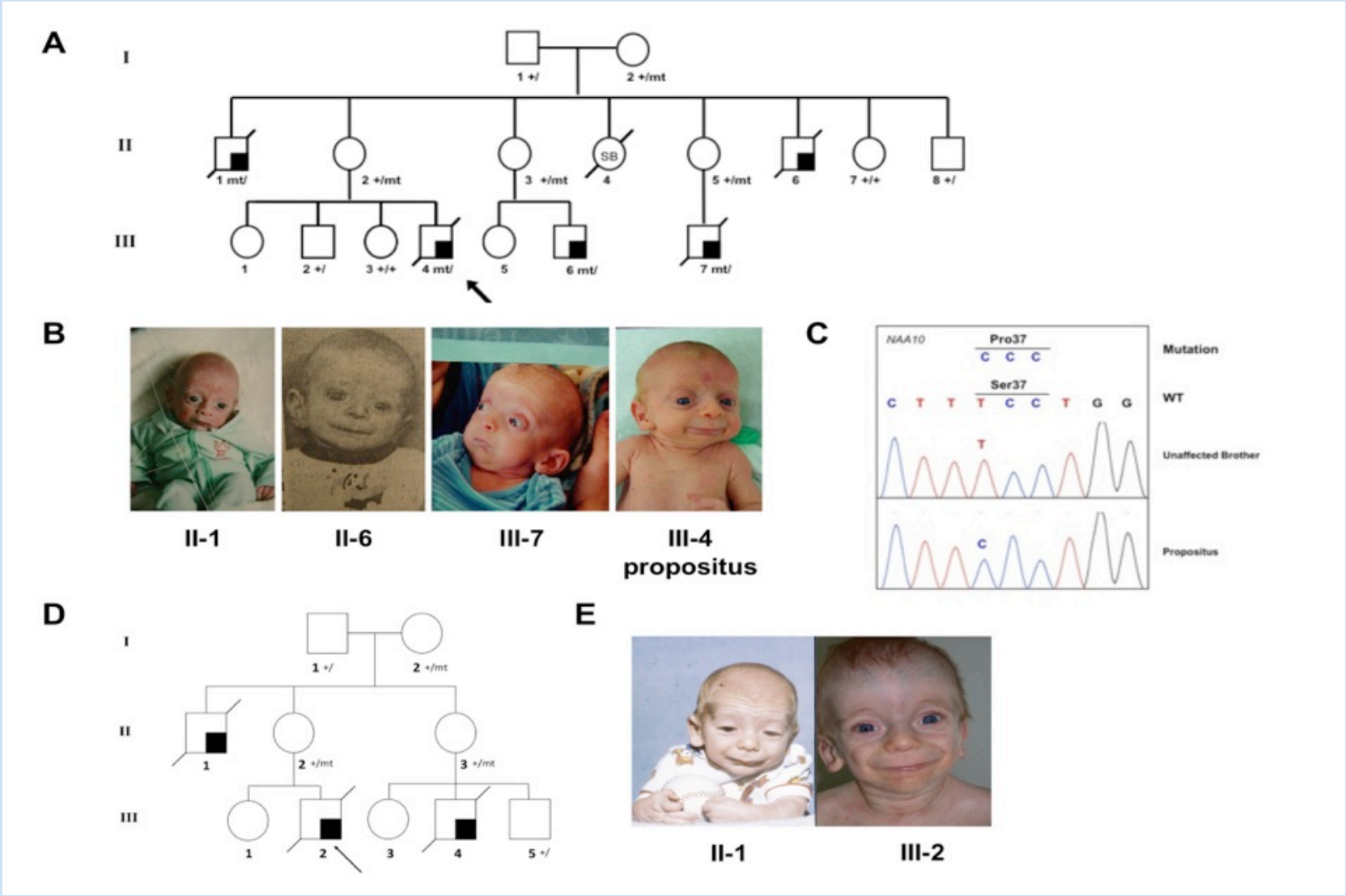




II-1 III-2

Contributed by Les Biesecker and colleagues at NIH

Ogden Syndrome – in 2011

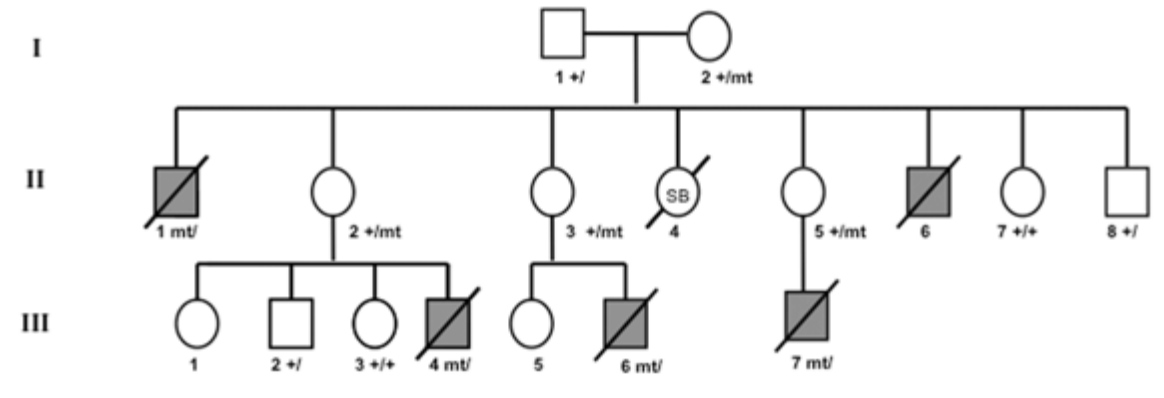


We found the SAME mutation in two unrelated families, with a very similar phenotype in both families, helping prove that this genotype contributes to the phenotype observed.

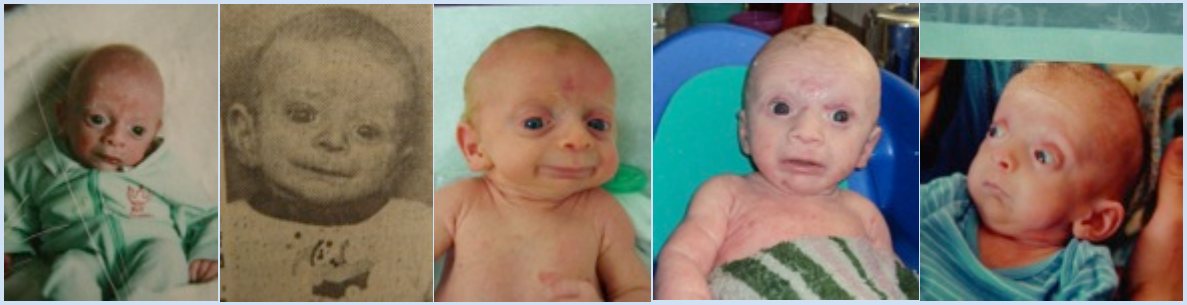
These are the Major Features of the Syndrome.

Table 1. Features of the syndrome	
Growth	post-natal growth failure
Development	global, severe delays
Facial	prominence of eyes, down-sloping palpebral fissures, thickened lids large ears beaking of nose, flared nares, hypoplastic alae, short columella protruding upper lip micro-retrognathia
Skeletal	delayed closure of fontanels broad great toes
Integument	redundancy / laxity of skin minimal subcutaneous fat cutaneous capillary malformations
Cardiac	structural anomalies (ventricular septal defect, atrial level defect, pulmonary artery stenoses) arrhythmias (Torsade de points, PVCs, PACs, SVtach, Vtach) death usually associated with cardiogenic shock preceded by arrythmia.
Genital	inguinal hernia hypo- or cryptorchidism
Neurologic	hypotonia progressing to hypertonia cerebral atrophy neurogenic scoliosis
Shaded regions include features of the syndrome demonstrating variability. Though variable findings of the cardiac, genital and neurologic systems were observed, all affected individuals manifested some pathologic finding of each.	

A

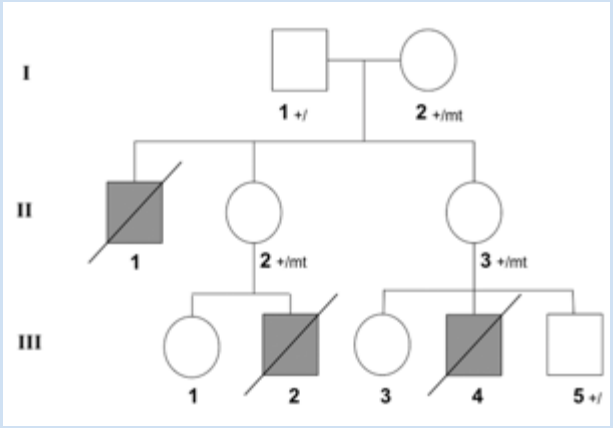


B



II-1 II-6 III-4 III-6 III-7

C



D

II-1 III-2

- **We performed X-chromosome exon capture with Agilent, followed by Next Gen Sequencing with Illumina.**

- **We analyzed the data with ANNOVAR and VAAST (Variant Annotation, Analysis and Search Tool). New computational tools for identifying disease-causing mutations by individual genome sequencing.**

Yandell, M. *et al.* 2011. "A probabilistic disease-gene finder for personal genomes." *Genome Res*. 21 (2011). doi:10.1101/gr.123158.111.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38, e164.

VAAST integrates AAS & Variant frequencies in a single probabilistic framework

- non-coding variants scored using allele frequency differences

- n_i : frequency of variant type among all variants observed in Background and Target genomes

- a_i: frequency of variant type among disease causing mutations in OMIM

- This approach means that *every* variant can be scored, non-synonymous, synonymous, coding, and non-coding. Phylogenetic conservation not required.

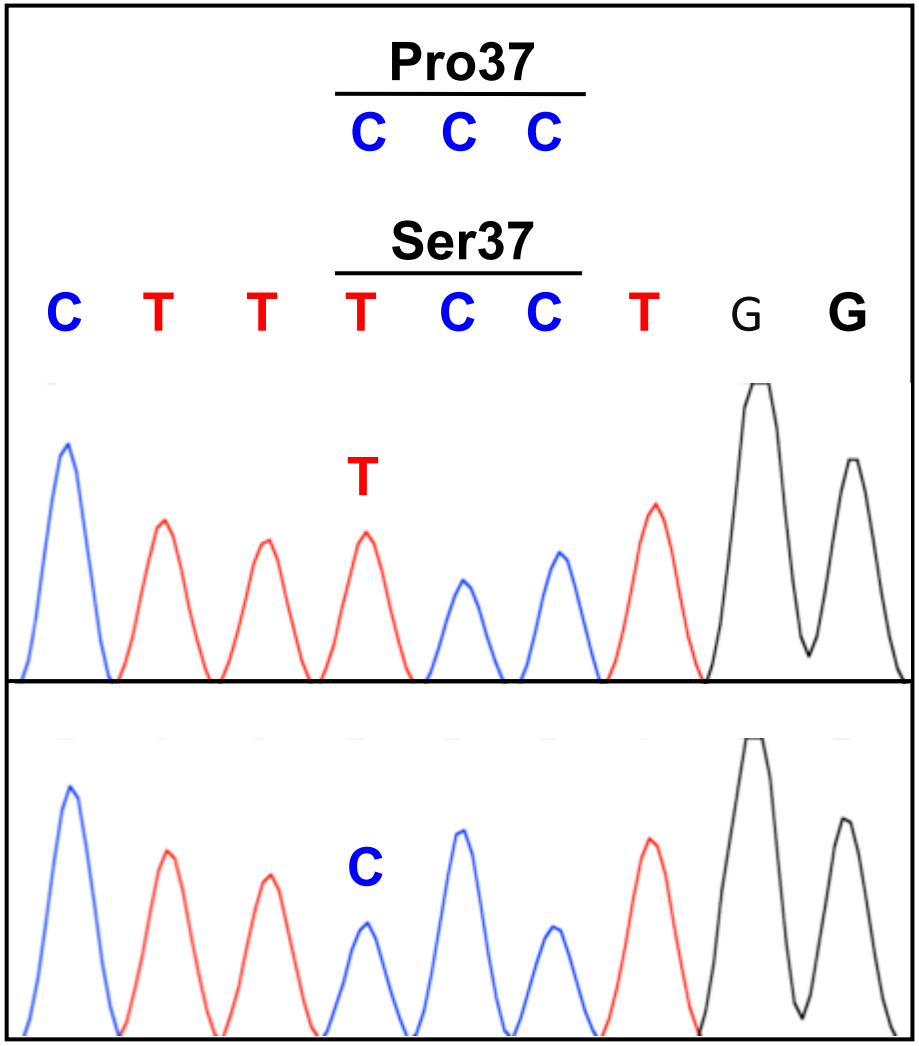
Omicia™ Unlocking Individualized Medicine

THE UNIVERSITY OF UTAH

Analysis with VAAST readily identified a few likely candidates.

Table 3. Summary of the filtering procedure and candidate genes using VAAST			
SNV calling pipeline	GATK	Samtools	GNUMAP
III-4 (total SNVs)	1546	1499	2168
III-4 (nsSNVs)	146	114	155
VAAST candidate genes (NAA10 ranking)	4 (3)	3 (2)	5 (2)
Present in III-4 and mother II-2 (nsSNVs)	122	107	116
VAAST candidate genes (NAA10 ranking)	3 (2)	2 (1)	2 (2)
Present in III-4, mother II-2, and grandmother I-2 (nsSNVs)	115	95	104
VAAST candidate genes (NAA10 ranking)	2 (1)	2 (1)	1 (1)
Present in III-4, II-2, and I-2, absent in brother III-2 and uncle II-8 (nsSNVs)	8	6	8
VAAST candidate genes (NAA10 ranking)	1 (1)	1 (1)	2 (1)

This is the mutation we found… one nucleotide change out of 6 billion nucleotides in a diploid genome…



Mutation

WT

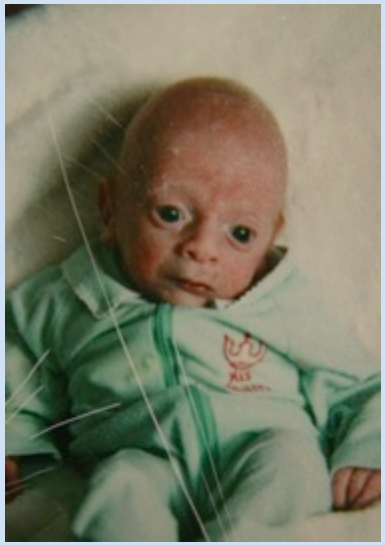
Unaffected Brother

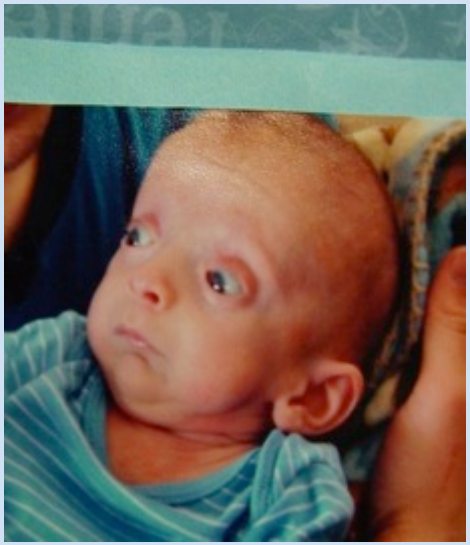
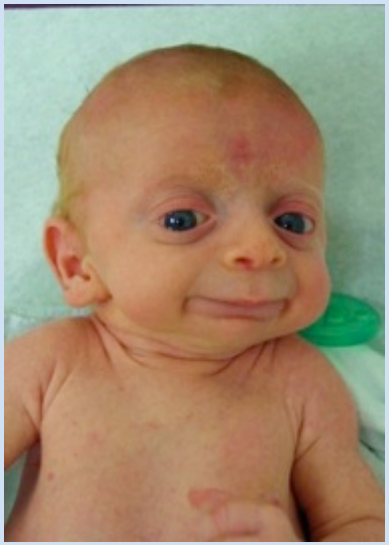
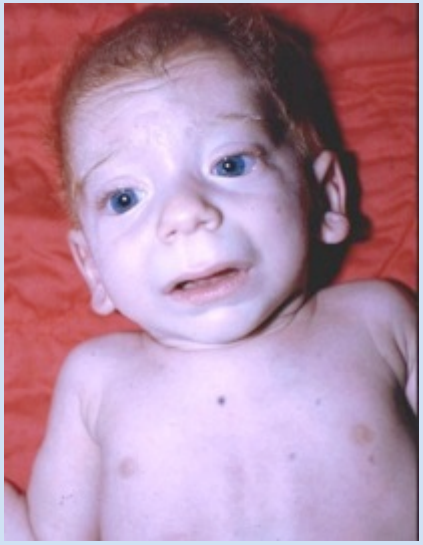
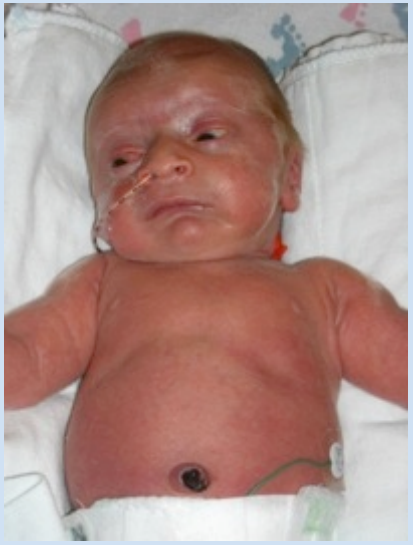
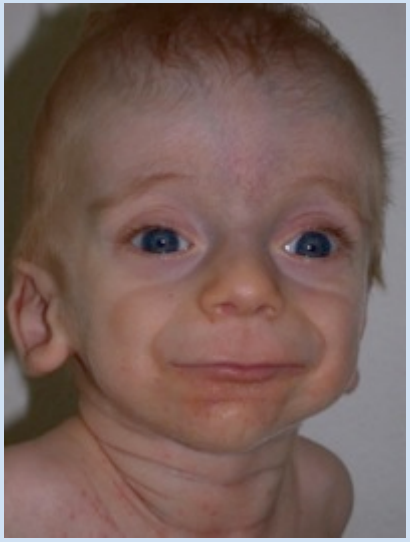
Proband

Proving Relevance of the mutation

- Present in two unrelated families with very similar phenotype of affected boys.
- Blinded Sanger sequencing showed perfect segregation of the mutation with the disease. Mutation present in Proband, Carrier Mother, Carrier Grandmother and other carrier mothers. Absent in unaffected brother and unaffected uncle.
- Also present in DNA from formalin-fixed paraffin-embedded tissue from two other deceased affected boys, found in pathology department, saved in one case for 30 years.
- Mutation NOT present in ~6000 exomes or genomes sequenced at BGI, CHOP and Utah for other projects.

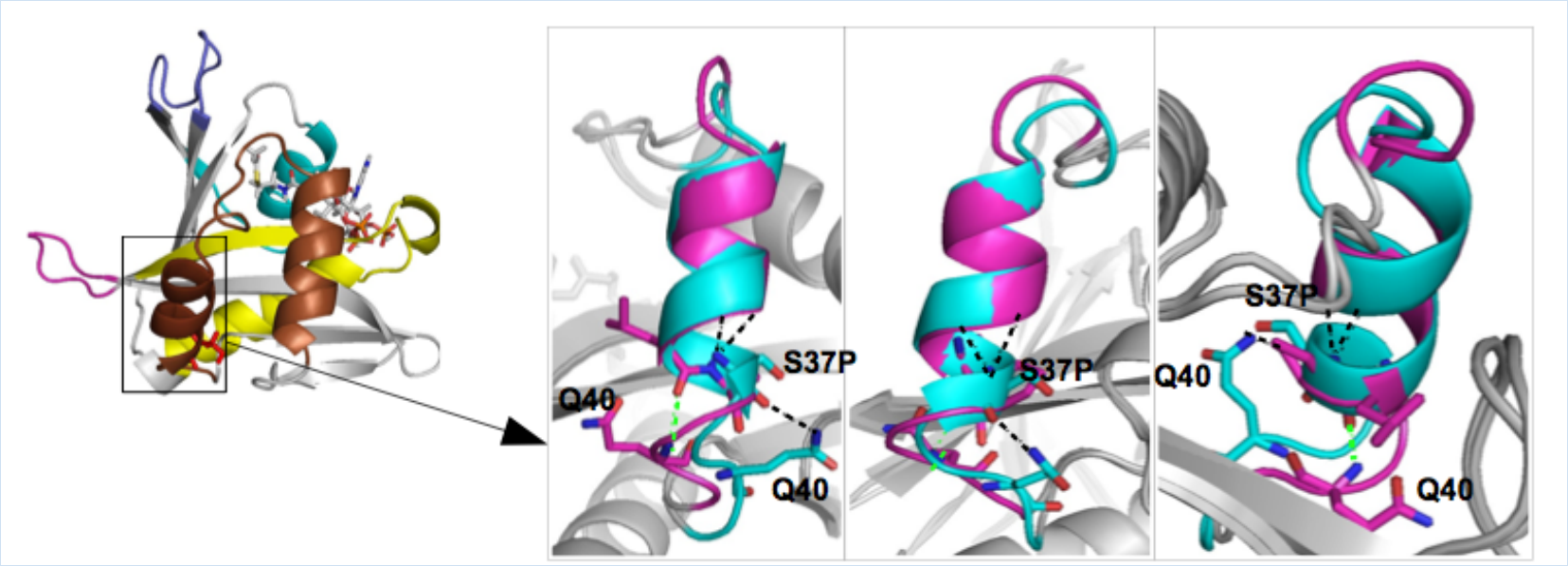
Ogden Syndrome, in honor of where the first family lives, in Ogden, Utah



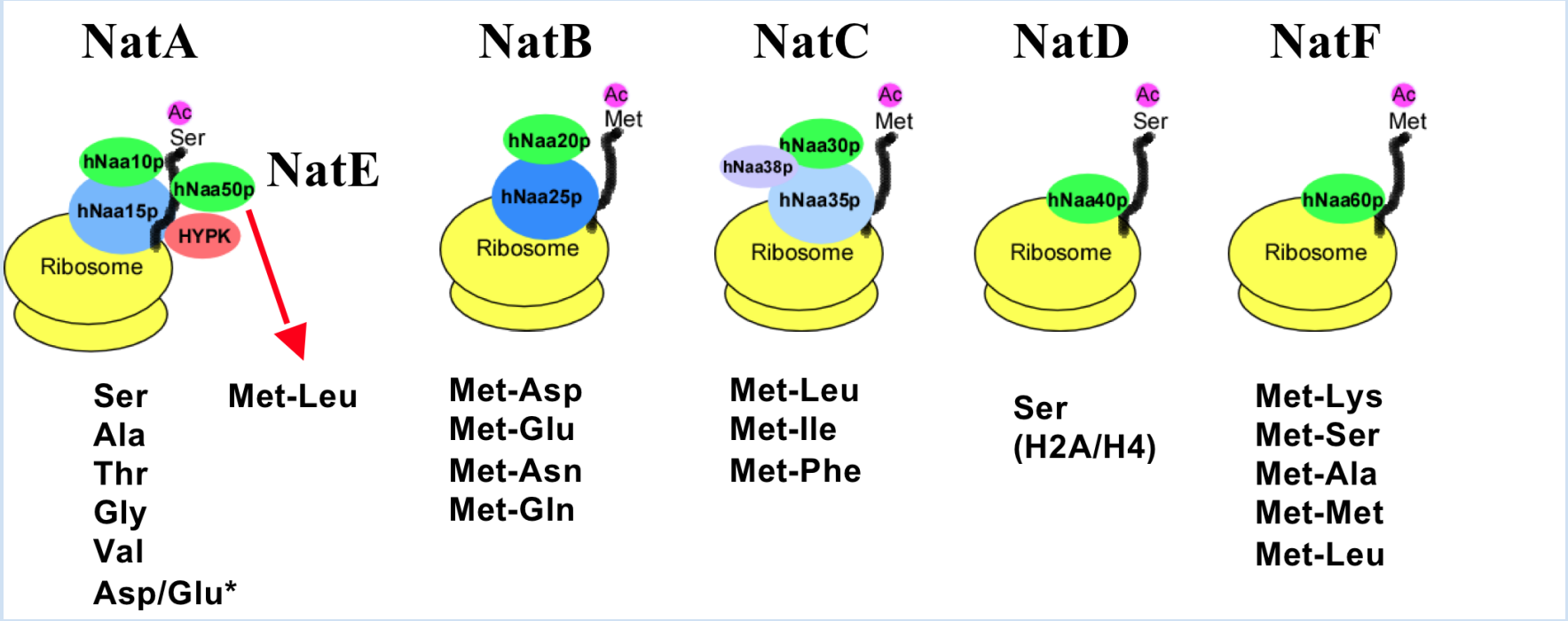






The mutation is a missense resulting in Serine to Proline change in Naa10p

- Ser 37 is conserved from yeast to human
- Ser37Pro is predicted to affect functionality (SIFT and other prediction programs)
- Structural modelling of hNaa10p wt (cyan) and S37P (pink)



The mutation disrupts the N-terminal acetylation machinery (NatA) in human cells.



Slide courtesy of Thomas Arnesen

N^{ε}-acetylation: Lysine Acetylation (KATs, HATs)

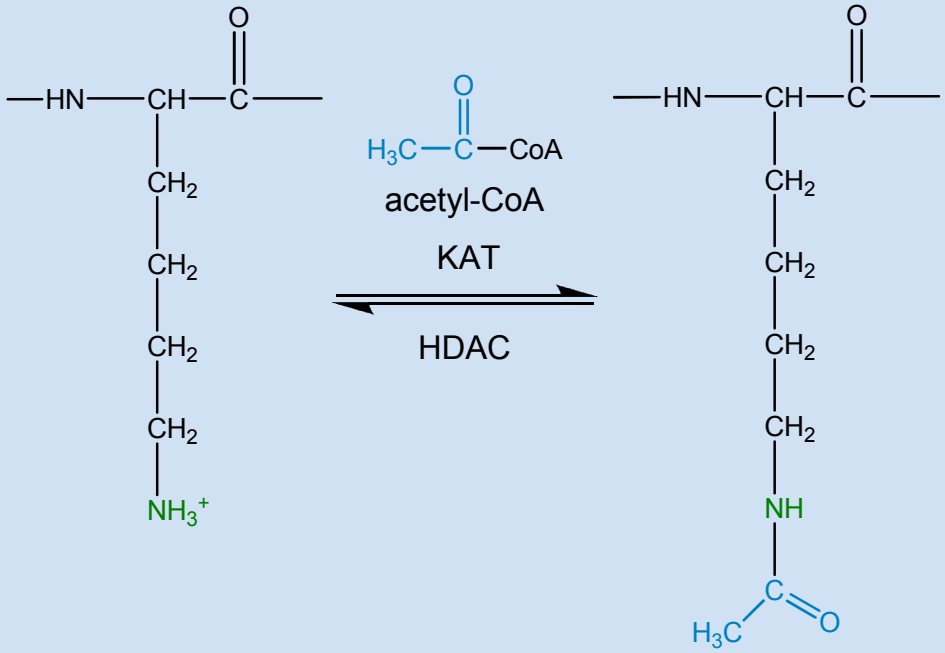


Figure 1.3: N^{ε}-acetylation. Lysine acetyltransferases (KATs) transfer the acetyl group of acetyl-CoA to the side chain of lysine, and thus remove the positive charge of the amino acid. Deacetylases (HDACs) catalyze the reverse reaction.

Annette Katharina Brenner

Dissertation for the degree philosophiae doctor (PhD) 2012

at the University of Bergen

N^{α}-acetylation: Amino-terminal Acetylation

In N^{α}-acetylation, the acetyl group of acetyl-CoA is transferred to the backbone of the N-terminal amino acid of the target protein (figure 1.4).

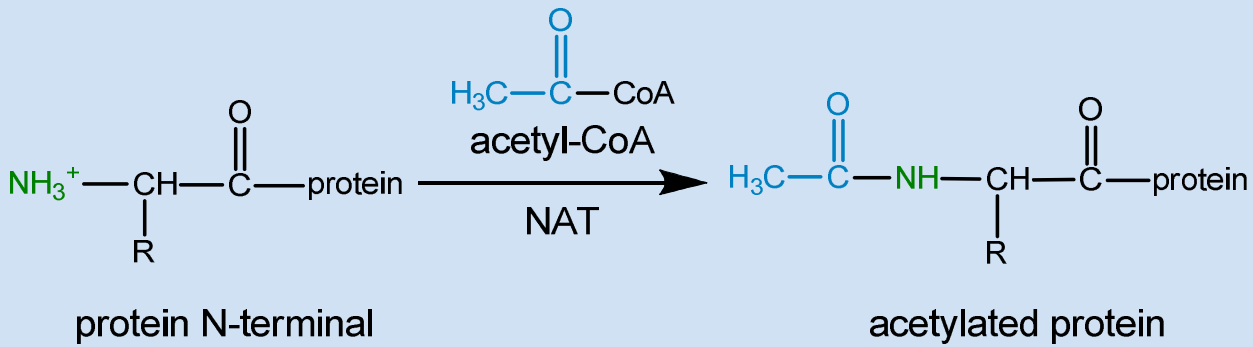


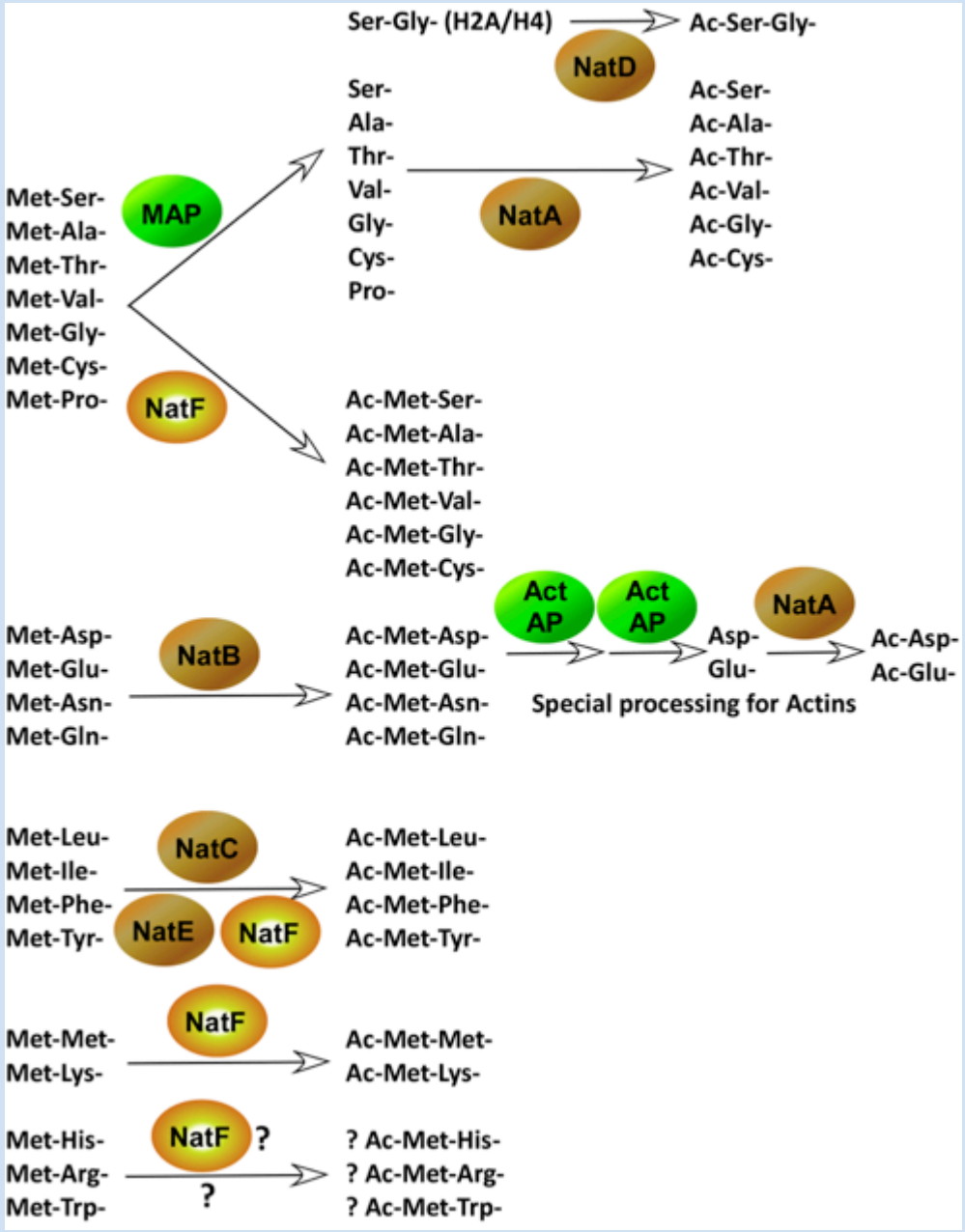
Figure 1.4: N^{α}-acetylation. The acetyl group of acetyl-CoA is irreversibly transferred to the N-terminal amino acid of the target protein by an N-terminal acetyltransferase (NAT).

Annette Katharina Brenner

Dissertation for the degree philosophiae doctor (PhD) 2012

at the University of Bergen

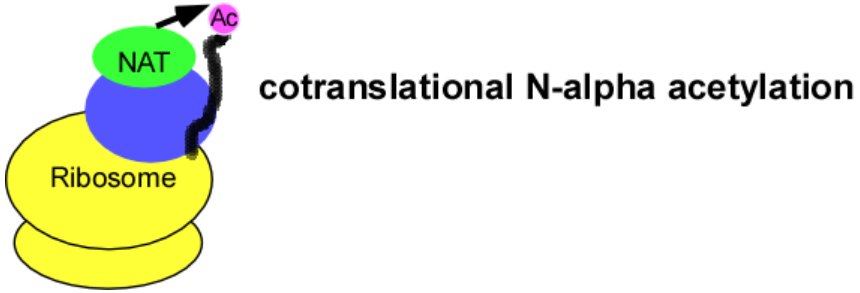
N-terminal processing in human cells



Van Damme P *et al.*, *PLoS Genet*, 2011

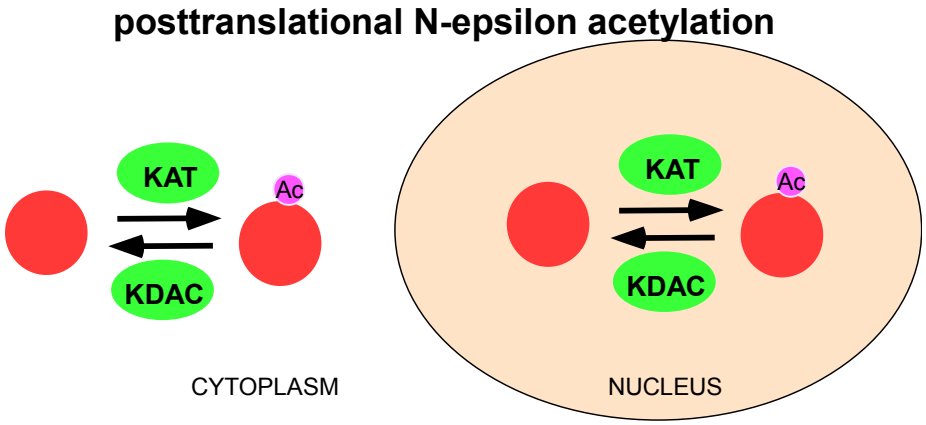
Protein acetylation in higher eukaryotes

1.



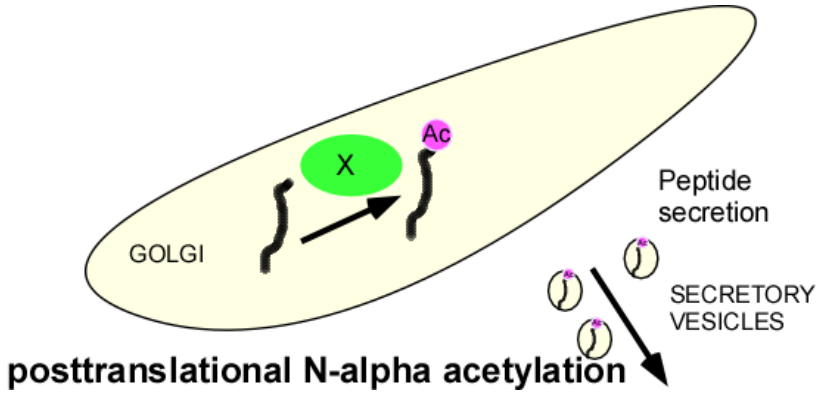
-Irreversible
-N-terminus
-Cytoplasm

2.



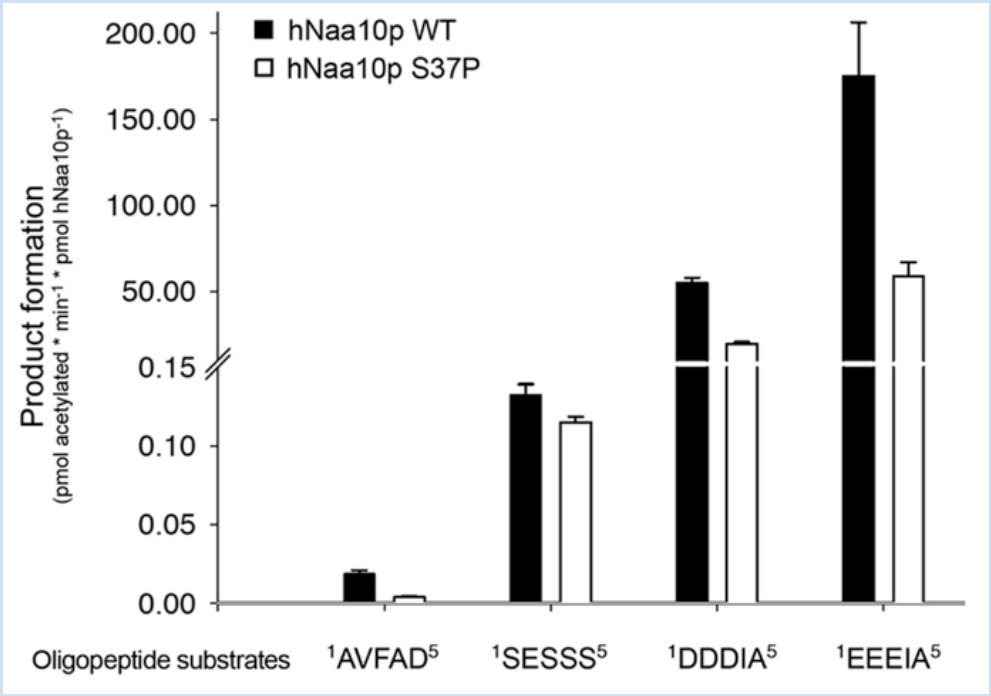
-Reversible
-Lysine
-Nucleus-cytoplasm

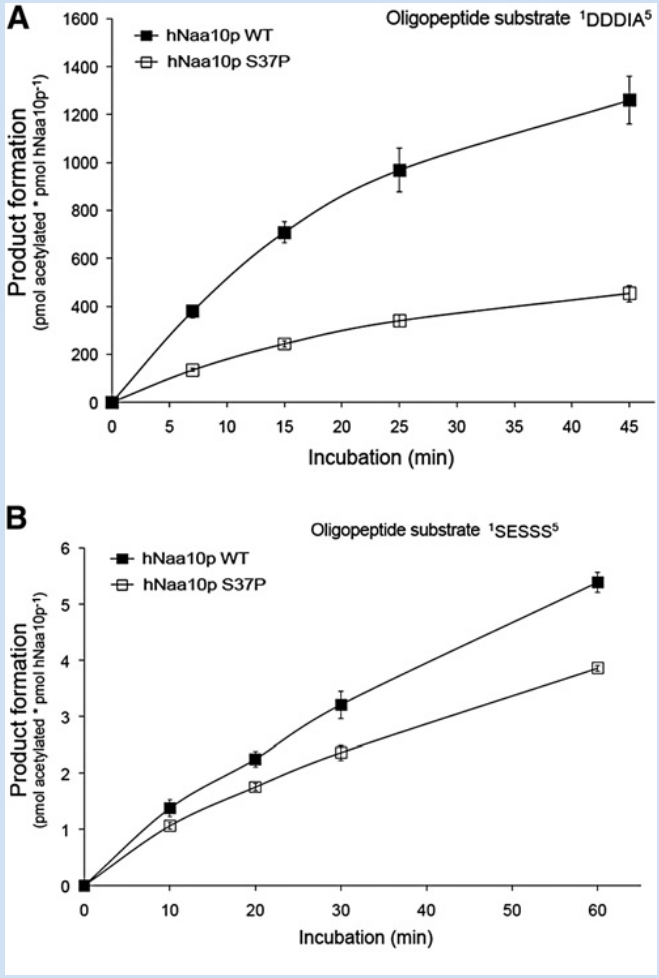
3.



-Irreversible?
-N-terminus
-Vesicles etc.

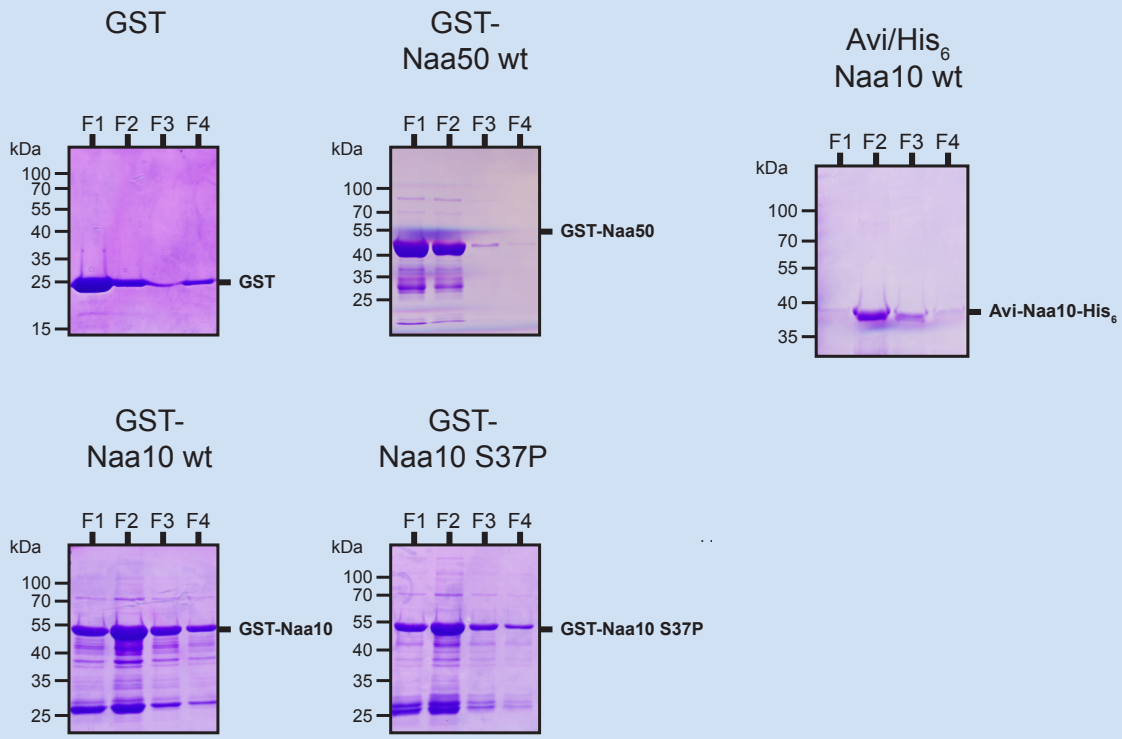
NAT activity of recombinant hNaa10p WT or p.Ser37Pro towards synthetic N-terminal peptides

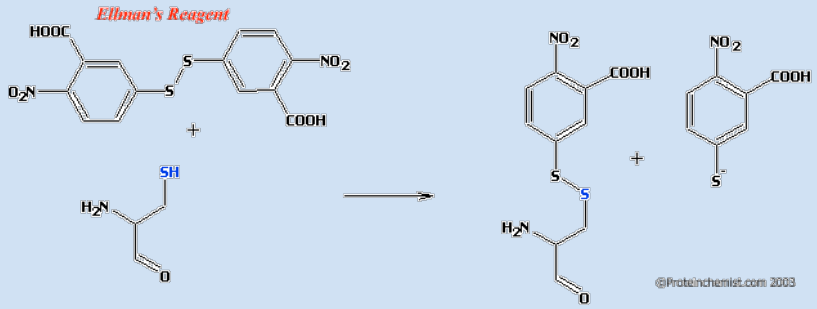


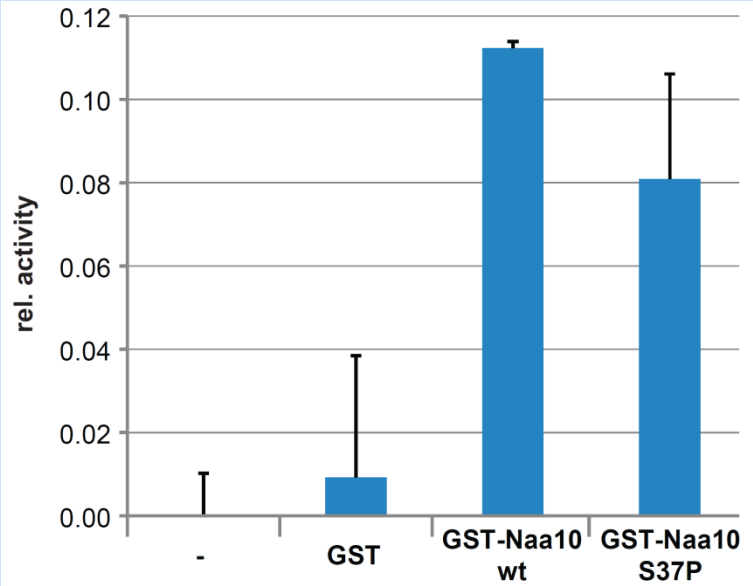


Assay performed in Thomas Arnesen lab

Protein Expression and purification





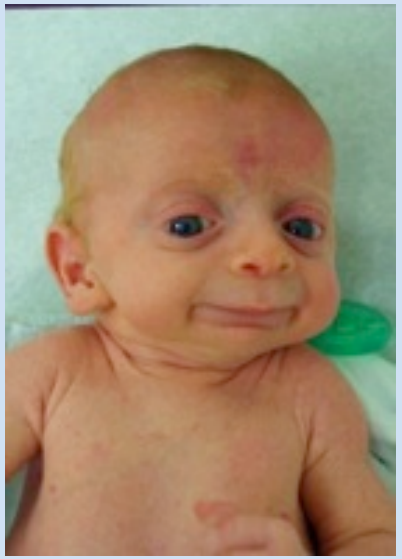


Peptide substrate: SYSMEHFRWGKPVGKKRRPVKVYP and corresponds to amino acids 1-24 of the human adrenocorticotropic hormone (ACTH

Big Questions though:

Simulated structure of S37P mutant

?



What is the molecular basis of Ogden syndrome?

- Naa10/Naa15 complex
- Naa10 localisation
- Naa10 function

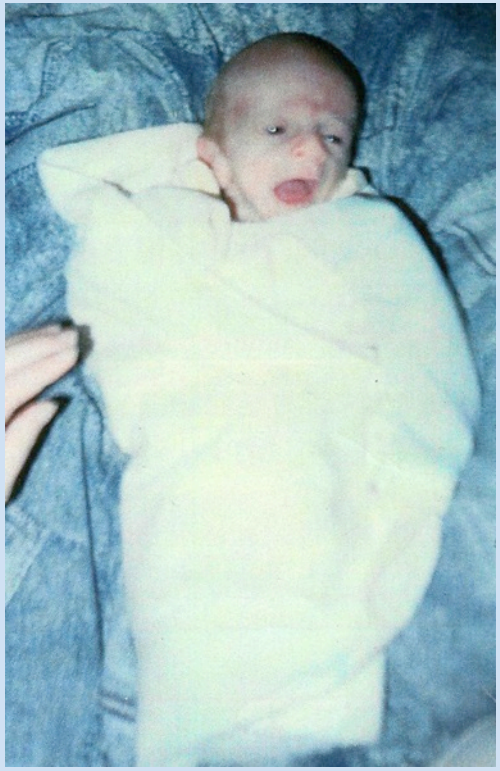
what can we learn from ogden syndrome?

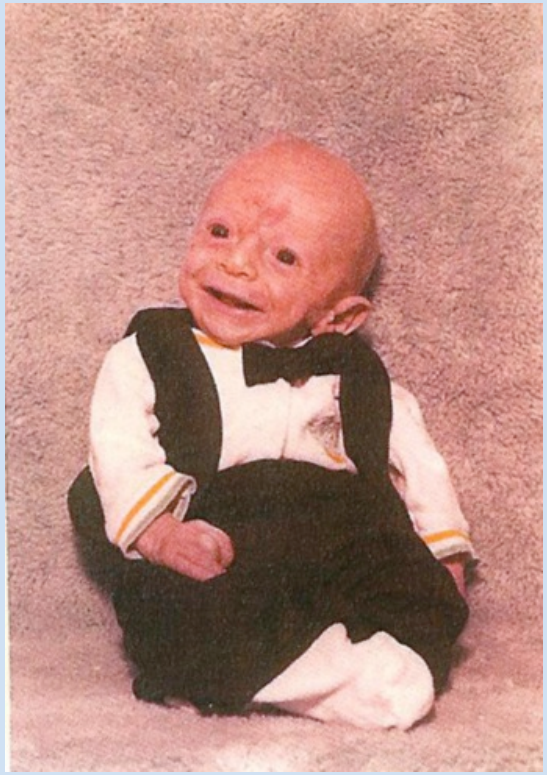
- characterizing different model systems (fibroblasts, yeast, *C. elegans)*

These are the Major Features of the Syndrome.

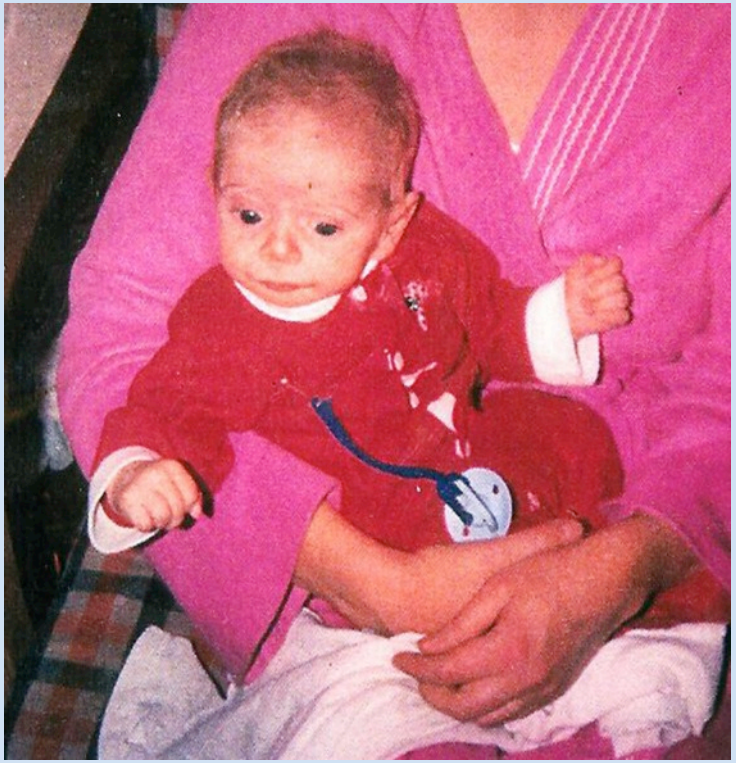
Table 1. Features of the syndrome	
Growth	post-natal growth failure
Development	global, severe delays
Facial	prominence of eyes, down-sloping palpebral fissures, thickened lids large ears beaking of nose, flared nares, hypoplastic alae, short columella protruding upper lip micro-retrognathia
Skeletal	delayed closure of fontanels broad great toes
Integument	redundancy / laxity of skin minimal subcutaneous fat cutaneous capillary malformations
Cardiac	structural anomalies (ventricular septal defect, atrial level defect, pulmonary artery stenoses) arrhythmias (Torsade de points, PVCs, PACs, SVtach, Vtach) death usually associated with cardiogenic shock preceded by arrythmia.
Genital	inguinal hernia hypo- or cryptorchidism
Neurologic	hypotonia progressing to hypertonia cerebral atrophy neurogenic scoliosis
Shaded regions include features of the syndrome demonstrating variability. Though variable findings of the cardiac, genital and neurologic systems were observed, all affected individuals manifested some pathologic finding of each.	

Family 1: II-1

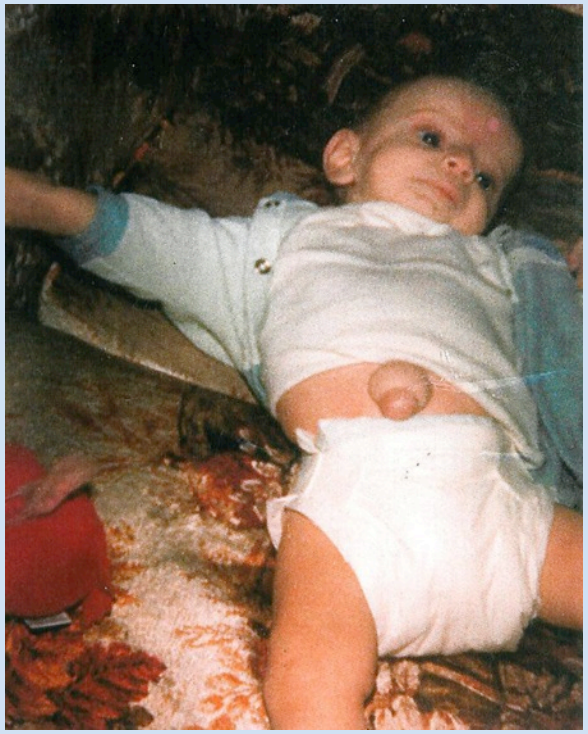


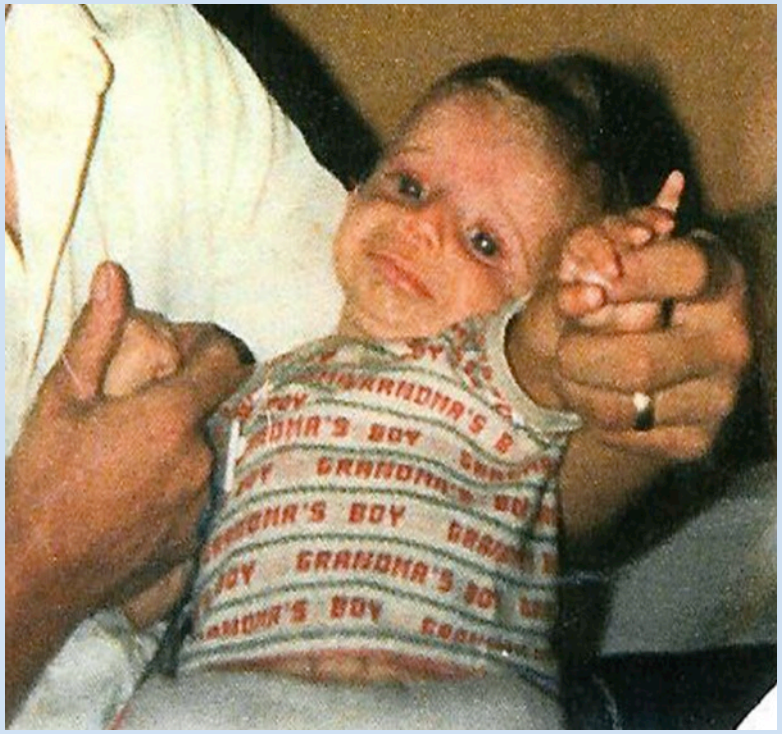


Family 1: II-1

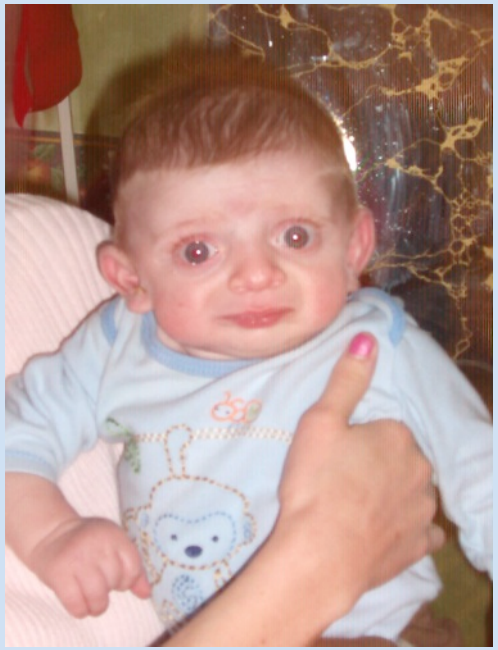


Family 1: II-6

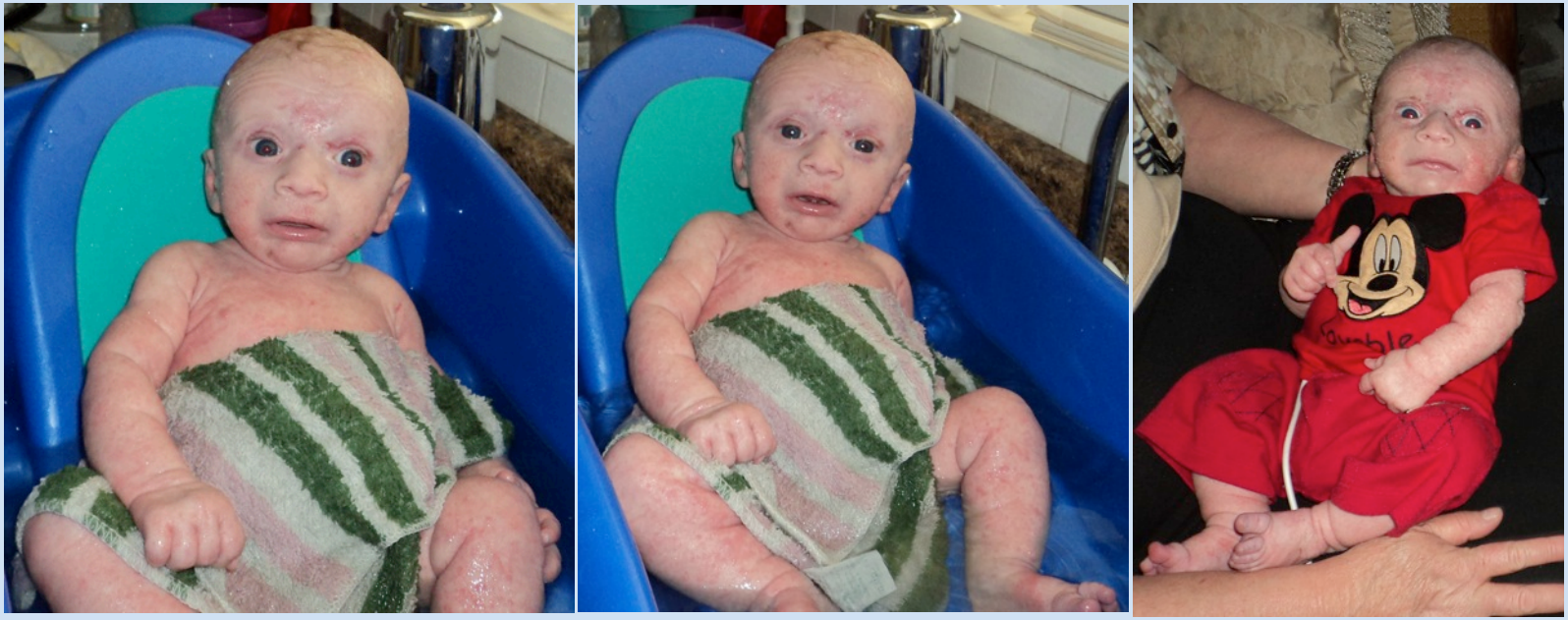




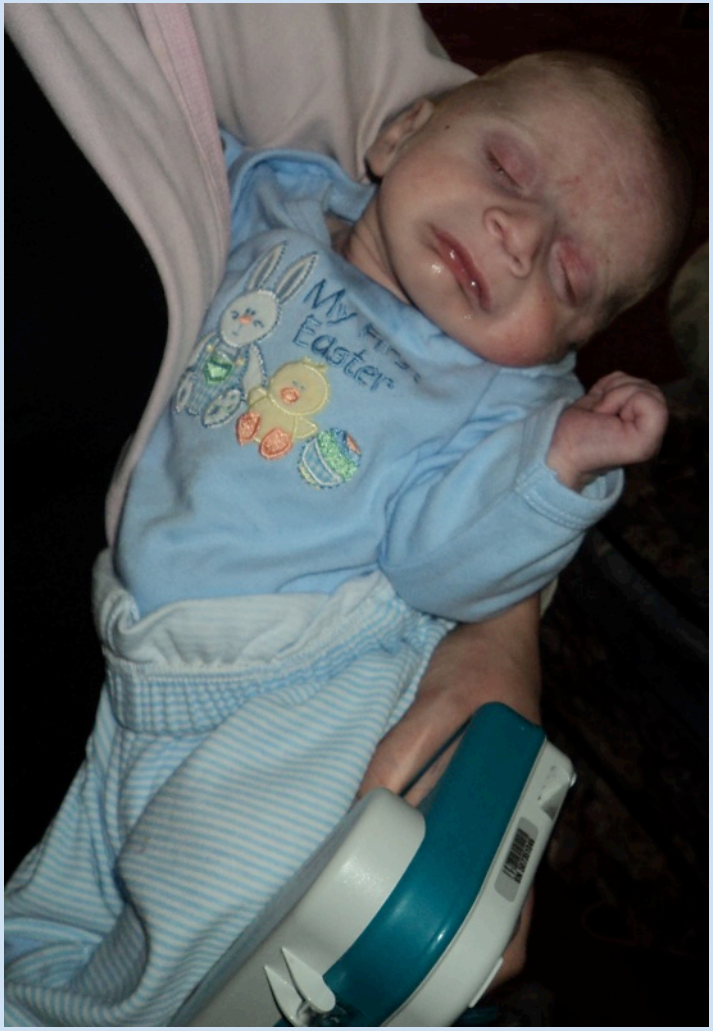
Family 1: III-4



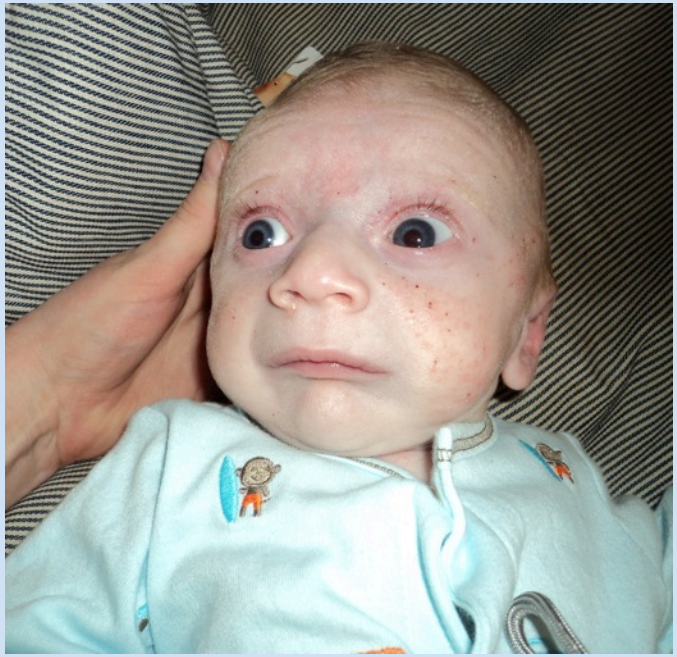
Family 1: III-6



Family 1: III-6



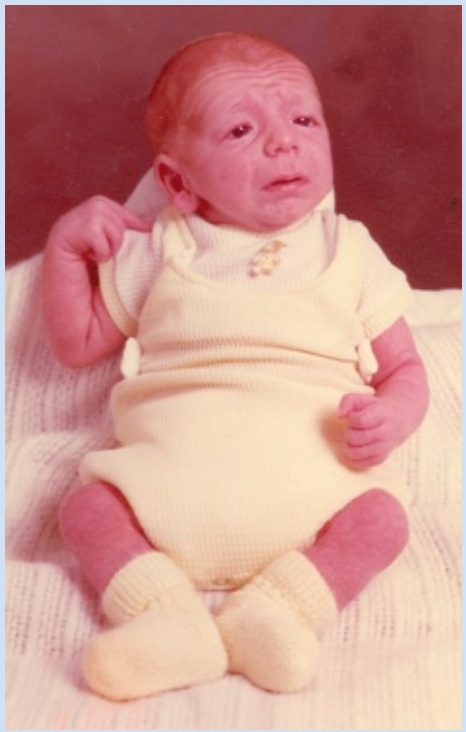
Family 1: III-6

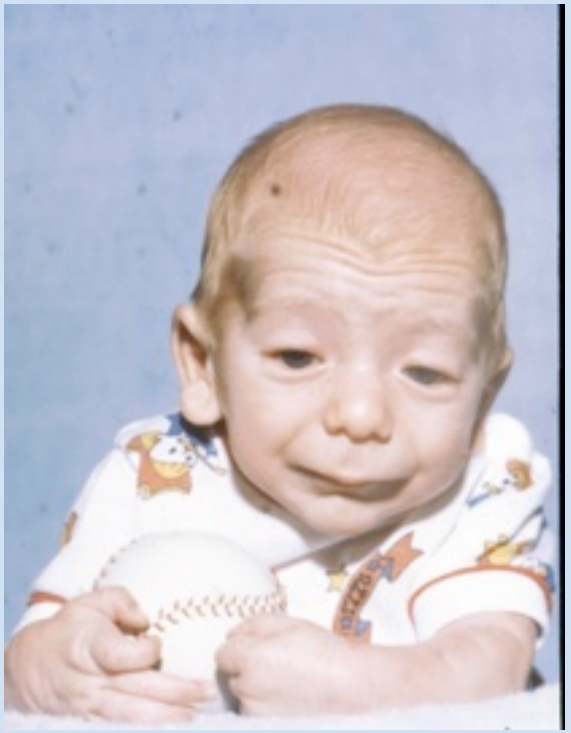


These are the Major Features of the Syndrome.

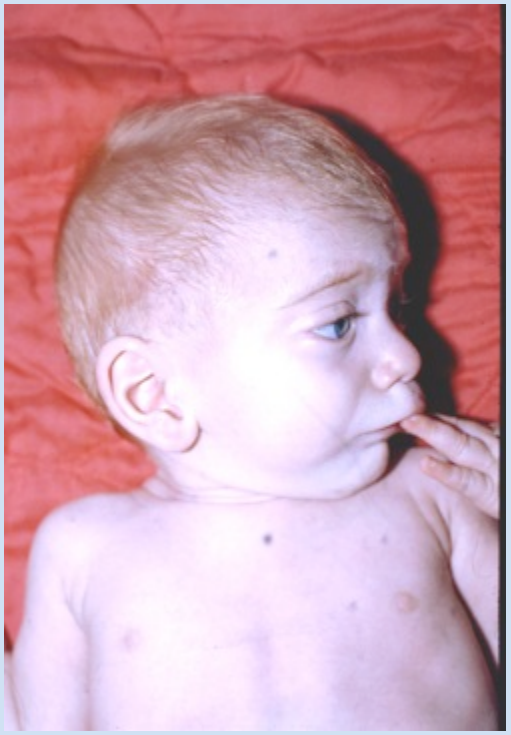
Table 1. Features of the syndrome	
Growth	post-natal growth failure
Development	global, severe delays
Facial	prominence of eyes, down-sloping palpebral fissures, thickened lids large ears beaking of nose, flared nares, hypoplastic alae, short columella protruding upper lip micro-retrognathia
Skeletal	delayed closure of fontanels broad great toes
Integument	redundancy / laxity of skin minimal subcutaneous fat cutaneous capillary malformations
Cardiac	structural anomalies (ventricular septal defect, atrial level defect, pulmonary artery stenoses) arrhythmias (Torsade de points, PVCs, PACs, SVtach, Vtach) death usually associated with cardiogenic shock preceded by arrythmia.
Genital	inguinal hernia hypo- or cryptorchidism
Neurologic	hypotonia progressing to hypertonia cerebral atrophy neurogenic scoliosis
Shaded regions include features of the syndrome demonstrating variability. Though variable findings of the cardiac, genital and neurologic systems were observed, all affected individuals manifested some pathologic finding of each.	

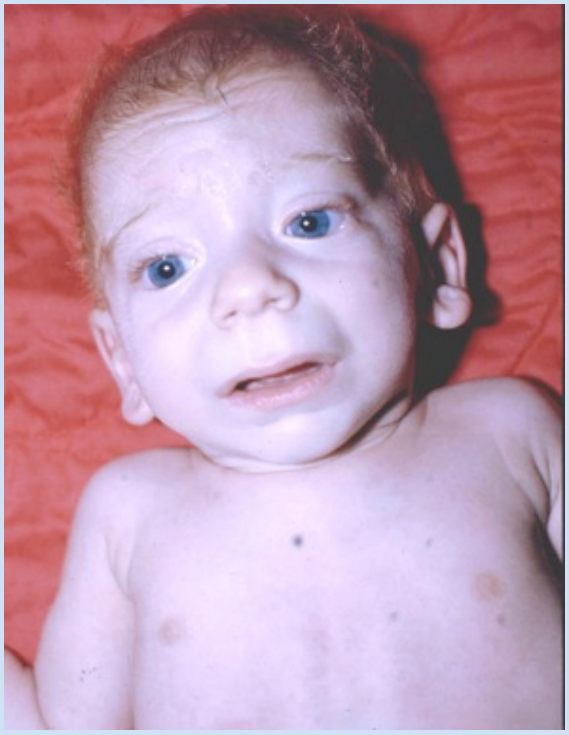
Family 2: II-1



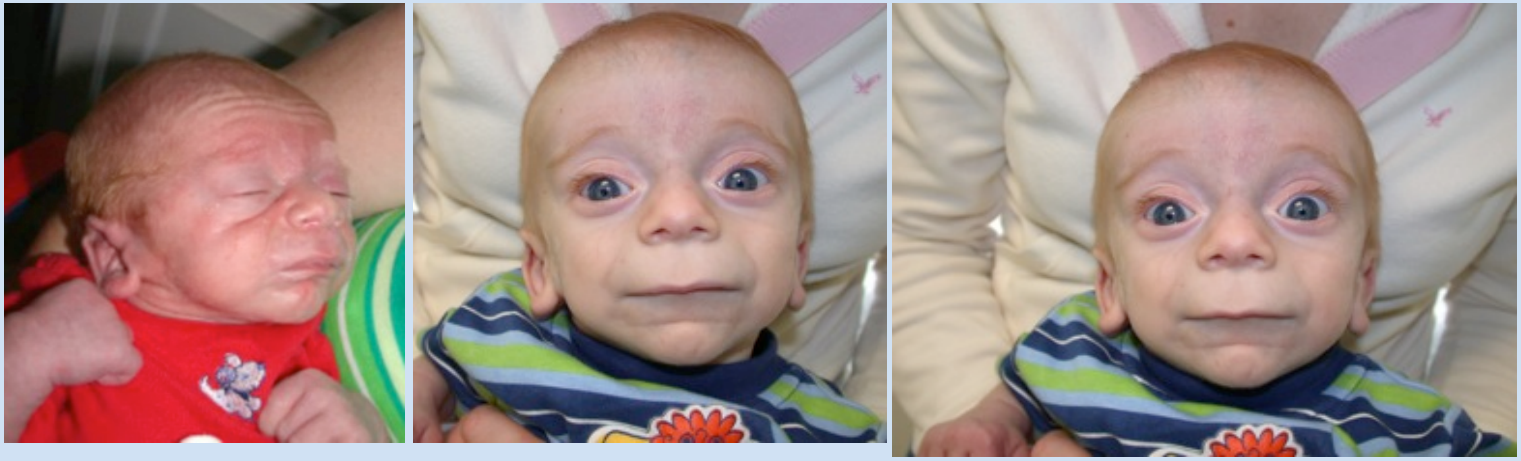


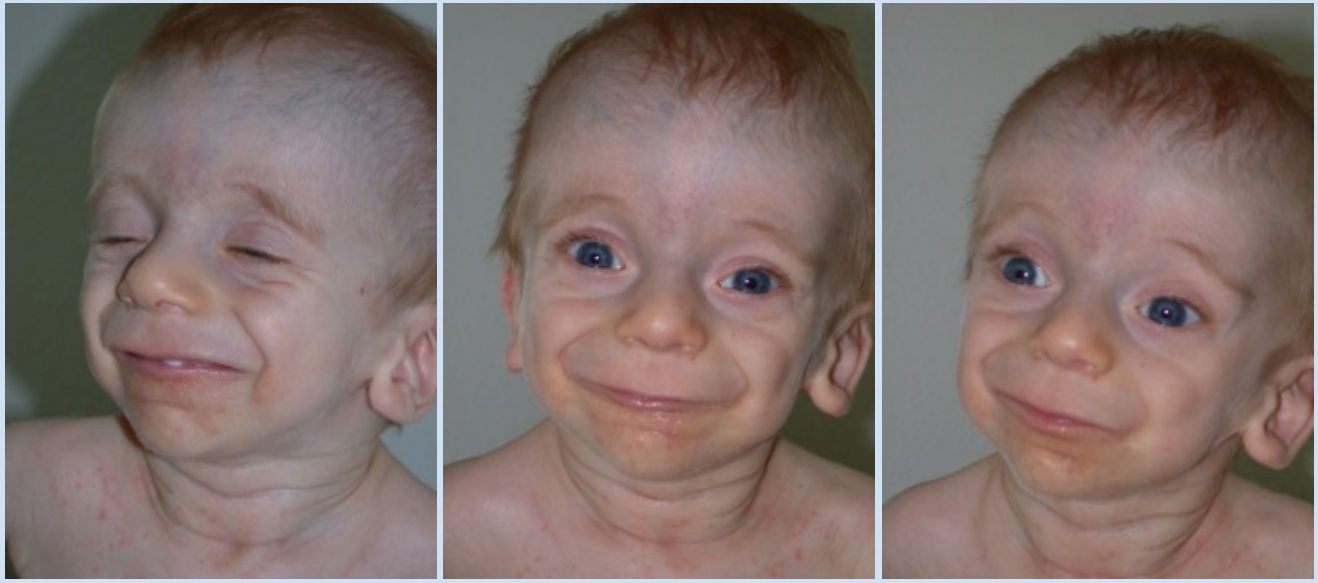
Family 2: II-1



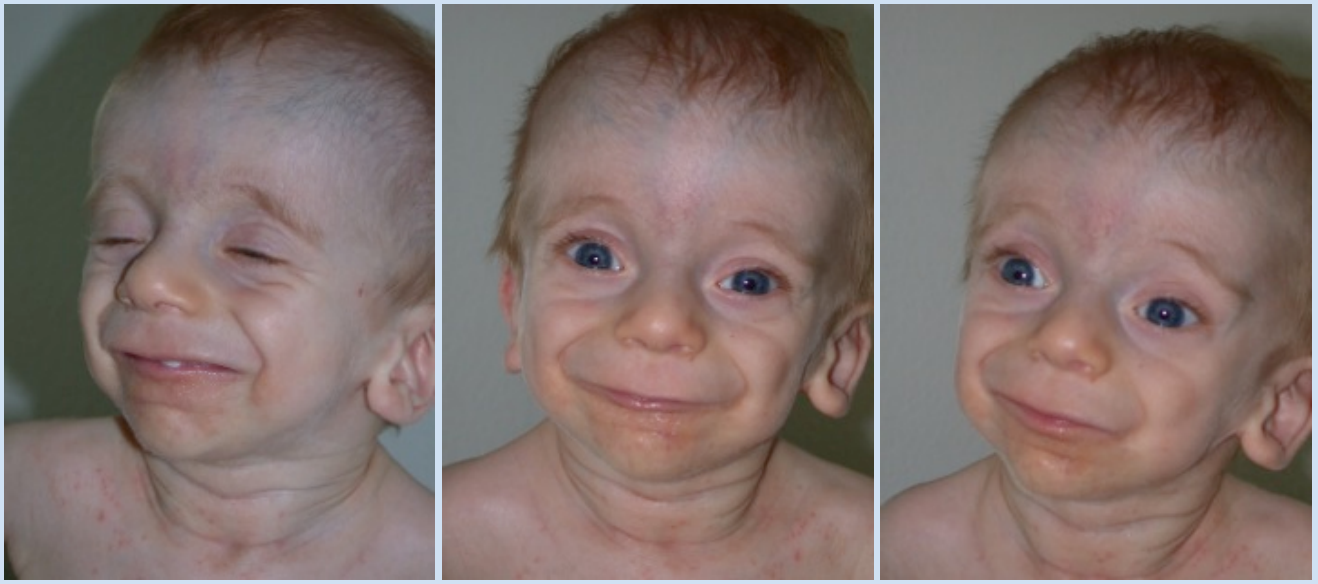


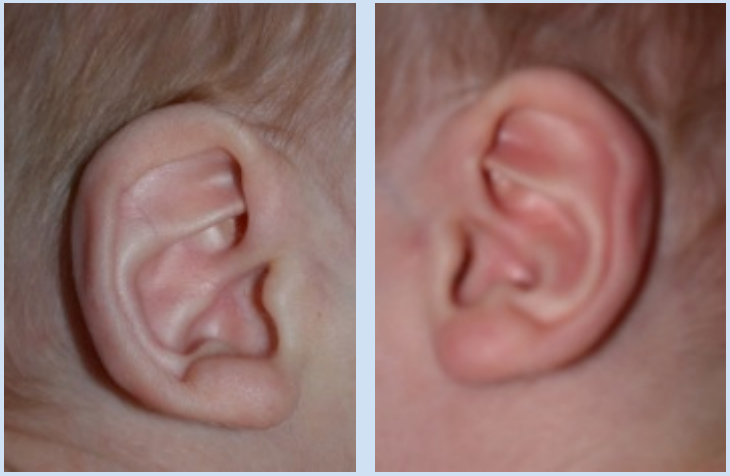
Family 2: III-2



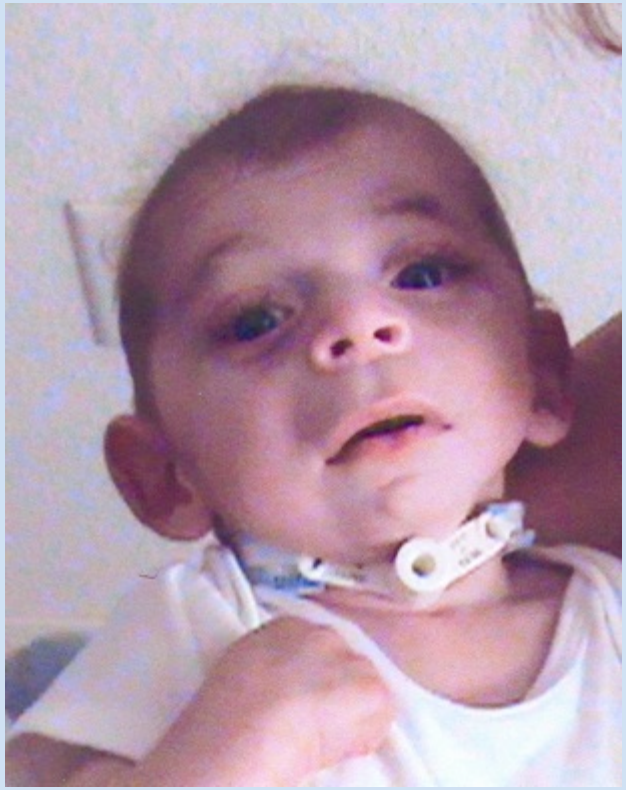


Family 2: III-2





Family 2: III-4



NAA10/NatA is essential for life

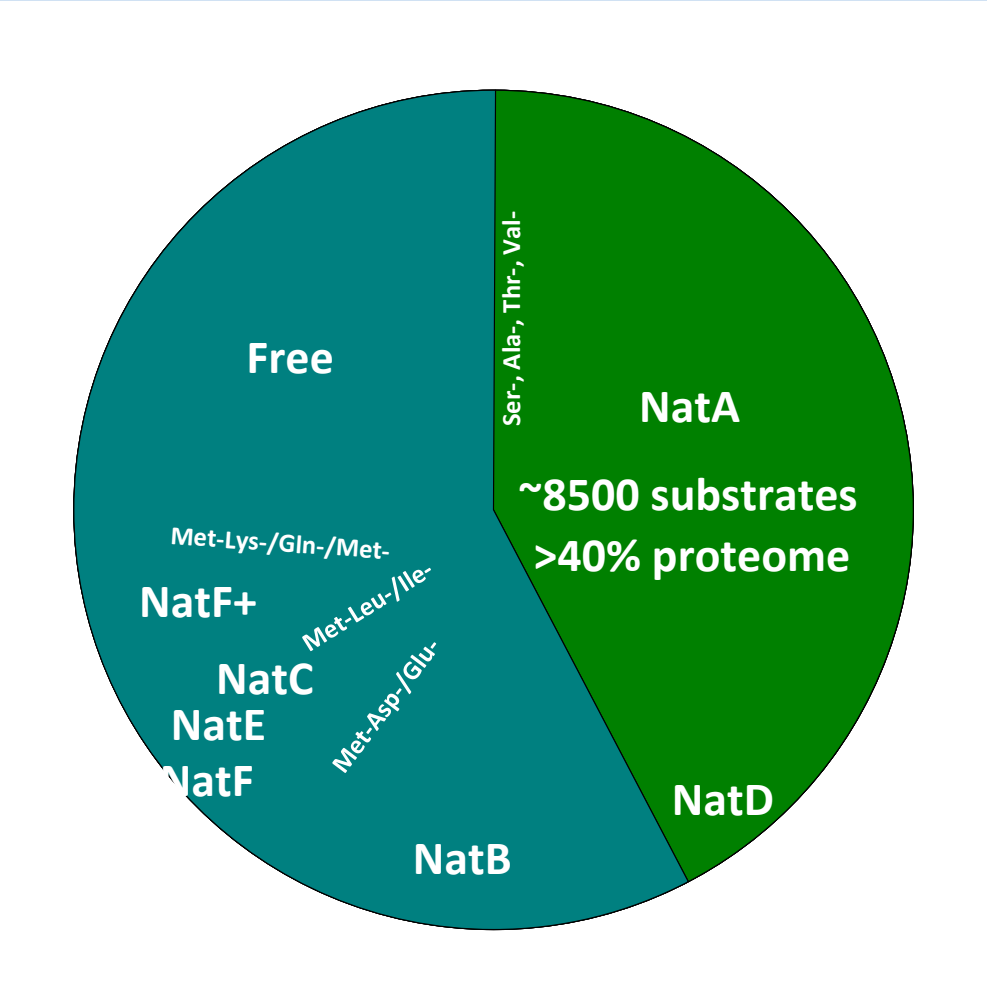
C. elegans

T. Brucei

D. melanogaster

Wang *et al.*, *Dev Dyn,* 2010; Ingram *et al*., *Mol Biochem Parasitol*, 2000; Sonnichsen *et al*., *Nature*, 2005

The human N-Acetylome



- A majority of soluble human proteins are N-terminally acetylated
- NatA is a major protein modifying enzyme of the human proteome

Arnesen T *et al., Proc Natl Acad Sci USA*, 2009; Van Damme P *et al., PLoS Genet*, 2011

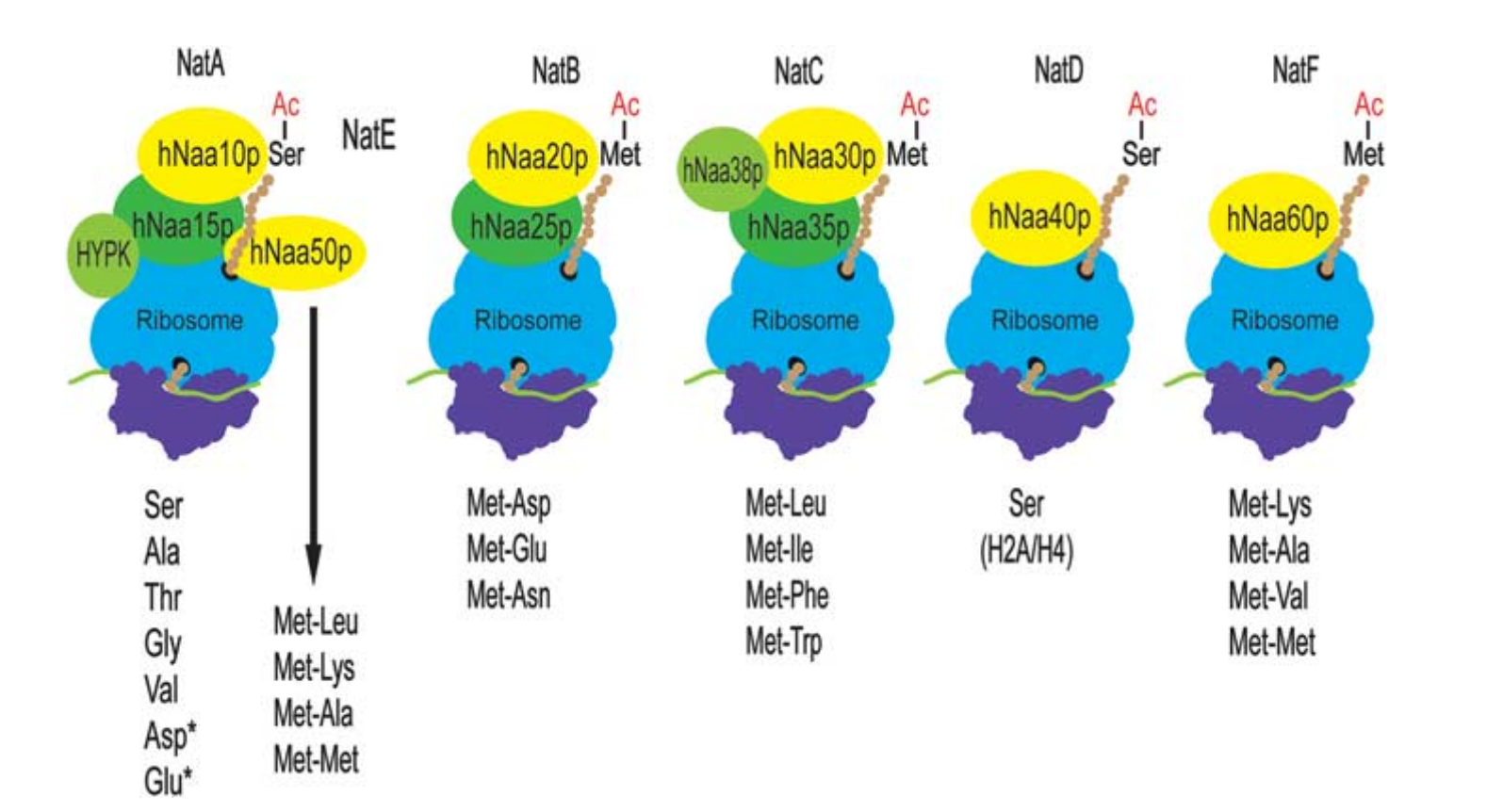


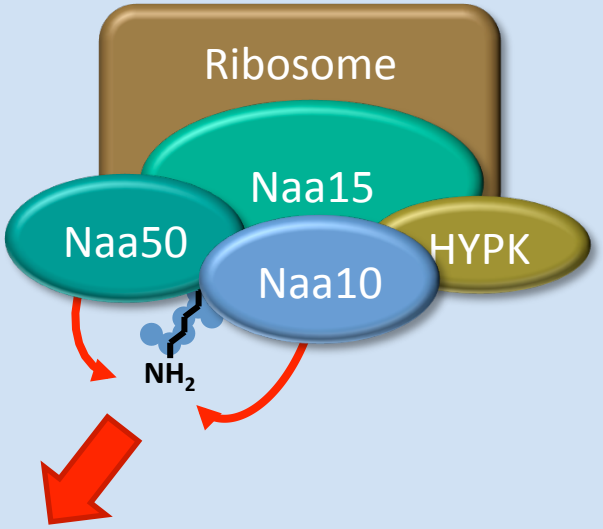
Figure 1. Overview of human NATs and their substrate specificities. The human NATs are composed of catalytic subunits (yellow) and auxiliary subunits (green), and all are associated with ribosomes (blue). NatA potentially acetylates Ser-, Ala-, Thr-, Gly- and Val- N-termini after the iMet has been removed by methionine aminopeptidases (MetAPs). NatB potentially acetylates Met-Asp-, Met-Glu- and Met-Asn-, whereas NatC may target Met-Leu-, Met-Ile-, Met-Phe- and Met-Trp-. NatD apparently only acetylates the Ser- N-termini of histones H2A and H4. NatE and NatF demonstrate some specificity towards NatC-type substrates as well as Met-Lys-, Met-Ala-, Met-Met- and Met-Val-. *NatA or hNaa10p may also mediate post-translational acetylation of mature actins harboring Asp- and Glu- N-termini.

REVIEW

Protein N-terminal acetyltransferases in cancer

TV Kalvik[1] and T Arnesen[1,2]

Naa10 function

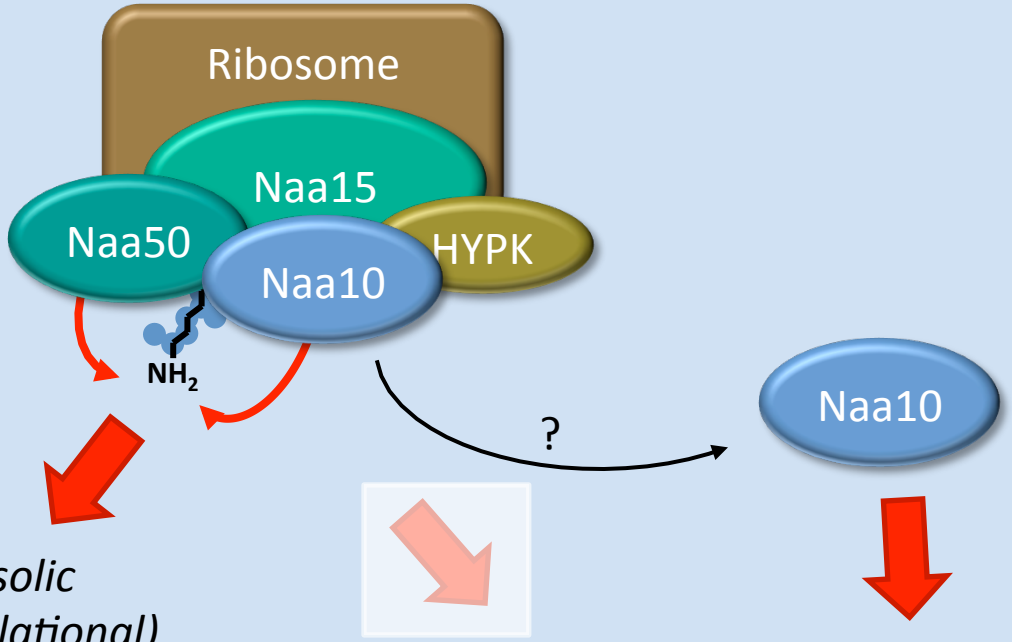


cytosolic
(co-translational)

- Rho (Rac/Cdc42 cell migration)
- Caspase2 (apoptosis)
- TSC2 (mTOR cell proliferation)
- HIF-1α (angiogenesis)
- α -Tubulin

protein stability, activity and sorting

Naa10 function



cytosolic
(co-translational)

- Rho (Rac/Cdc42 cell migration)
- Caspase2 (apoptosis)
- TSC2 (mTOR cell proliferation)
- HIF-1α (angiogenesis)
- α -Tubulin

protein stability, activity and sorting

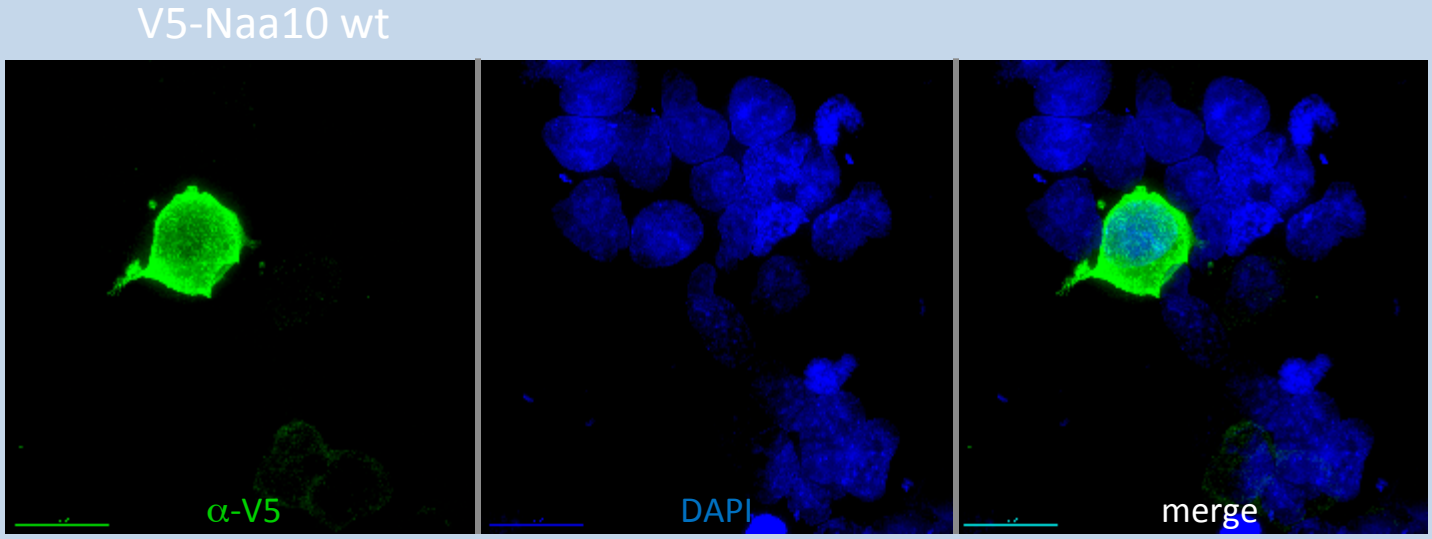
nuclear
transcription factors

- DNMT1 (E-Cad transformation)
- RelA/p65 (MCL-1 survival)
- Rip1/NEMO (DNA damage)
- β-Catenin/Lef-1/TEF (wnt-pathway)
- AR (proliferation)

localization of Naa10

Immunofluorescence

- HEK293 cells +/- V5-hNaa10 wt or mutants
- IF with anti-V5

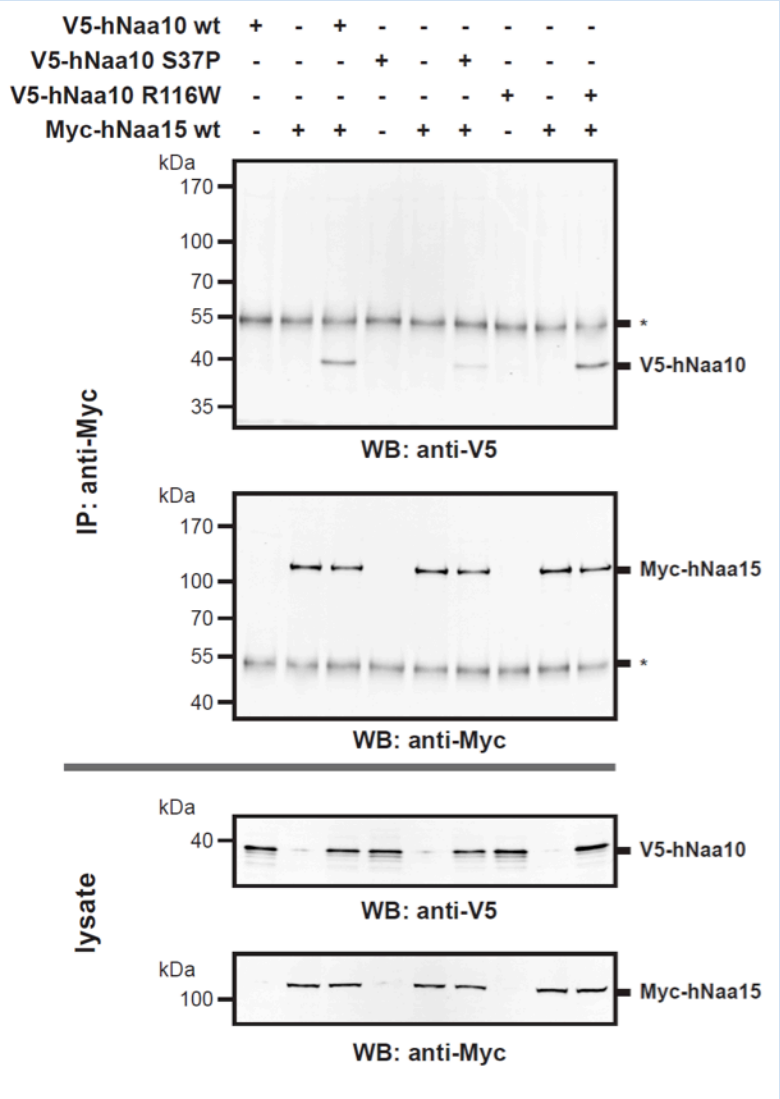


interaction of Naa10 and Naa15

Co-IP of Naa10 and Naa15 in HEK293 cells

- precipitating antibody: α-Myc

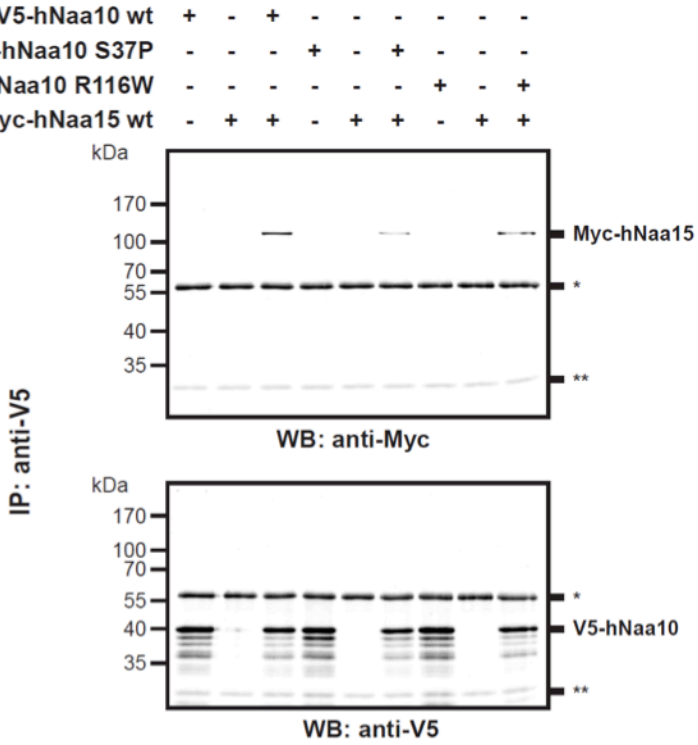
→ reduced interaction of Naa15 and Naa10 S37P?



interaction of Naa10 and Naa15

Co-IP of Naa10 and Naa15 in HEK293 cells

- precipitating antibody: α-V5



→ reduced interaction of Naa15 and Naa10 S37P?

Molecular basis for N-terminal acetylation by the heterodimeric NatA complex

Glen Liszczak[1,2], Jacob M Goldberg[2], Håvard Foyn[3], E James Petersson[2], Thomas Arnesen[3,4] & Ronen Marmorstein[1,2]

VOLUME 20 NUMBER 9 SEPTEMBER 2013 NATURE STRUCTURAL & MOLECULAR BIOLOGY

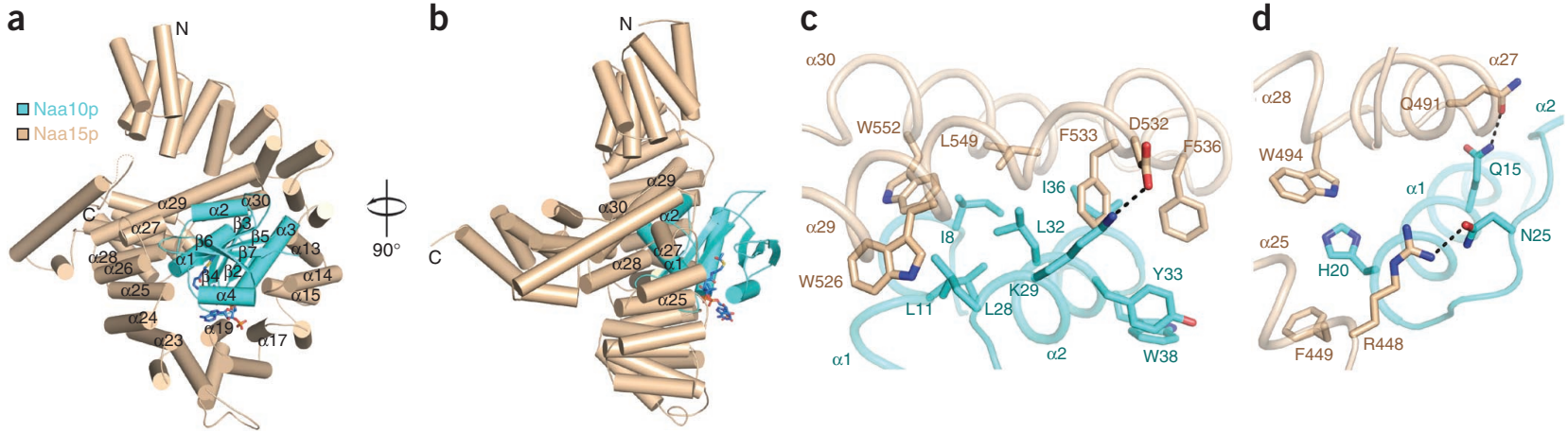


Figure 1 Overall structure of the NatA complex bound to acetyl CoA. (**a**) Naa10p (teal) and Naa15p (brown) subunits, shown in cartoon, bound to acetyl CoA (CPK coloring and stick format). Only Naa15p helices that contact Naa10p are labeled. The dashed brown line represents a disordered loop region in Naa15p. The dimensions of the complex are 107 Å × 85 Å × 70 Å. (**b**) A 90° rotation of the view in **a**. Helices that are depicted in **c** and **d** are labeled. (**c**) Zoom view highlighting key residues that compose the predominantly hydrophobic interface between Naa10p α1–α2 and Naa15p α29–α30. (**d**) Zoom view of the intersubunit interface at the C-terminal region of Naa10p α1 and the Naa15p α25–α27–α28 helices.

Crystal Structure of NAA50

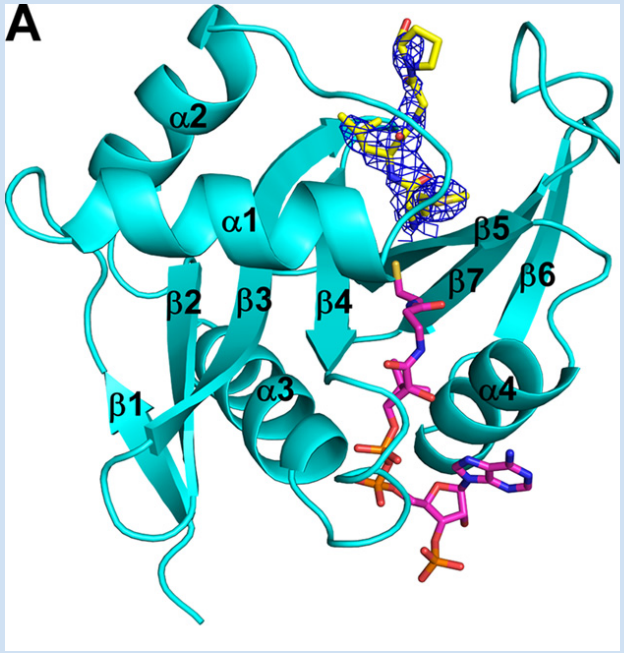


FIGURE 1. **Overall structure of the ternary Naa50p·CoA·peptide complex.** *A*, structure of the ternary complex showing Naa50p in *teal*; CoA as *magenta sticks* with carbon, nitrogen, and sulfur atoms in Corey-Pauling-Koltun coloring; and the substrate peptide as *yellow sticks* with carbon, nitrogen, and sulfur atoms in Corey-Pauling-Koltun coloring. Substrate peptide electron density obtained from a composite omit map (*blue*) is shown contoured to 1.5σ. *B*, superposition of the ternary Naa50p complex with the ternary Gcn5

Glen Liszczak[‡§], Thomas Arnesen[¶‖], and Ronen Marmorstein[‡§1]

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 42, pp. 37002–37010, October 21, 2011

LETTER

doi:10.1038/nature12141

De novo mutations in histone-modifying genes in congenital heart disease

Samir Zaidi[1,2]*, Murim Choi[1,2]*, Hiroko Wakimoto[3], Lijiang Ma[4], Jianming Jiang[3,5], John D. Overton[1,6,7], Angela Romano-Adesman[8], Robert D. Bjornson[7,9], Roger E. Breitbart[10], Kerry K. Brown[3], Nicholas J. Carriero[7,9], Yee Him Cheung[11], John Deanfield[12], Steve DePalma[3], Khalid A. Fakhro[1,2], Joseph Glessner[13], Hakon Hakonarson[13,14], Michael J. Italia[15], Jonathan R. Kaltman[16], Juan Kaski[12], Richard Kim[17], Jennie K. Kline[18], Teresa Lee[4], Jeremy Leipzig[15], Alexander Lopez[1,6,7], Shrikant M. Mane[1,6,7], Laura E. Mitchell[19], Jane W. Newburger[10], Michael Parfenov[3], Itsik Pe'er[20], George Porter[21], Amy E. Roberts[10], Ravi Sachidanandam[22], Stephan J. Sanders[1,23], Howard S. Seiden[24], Mathew W. State[1,23], Sailakshmi Subramanian[22], Irina R. Tikhonova[1,6,7], Wei Wang[15,25], Dorothy Warburton[4,26], Peter S. White[14,15], Ismee A. Williams[4], Hongyu Zhao[1,27], Jonathan G. Seidman[3], Martina Brueckner[1,28], Wendy K. Chung[4,29], Bruce D. Gelb[22,24,30], Elizabeth Goldmuntz[14,31], Christine E. Seidman[3,5,32] & Richard P. Lifton[1,2,6,7,33]

Table 2 | Genes of interest with *de novo* mutations in probands

ID	Gene	Mutation	Dx	Other structural/neuro/ht-wt
1-00596	*MLL2*†	p.Ser1722Arg fs*9	LVO	Y/Y/N
1-00853	*WDR5*†	p.Lys7Gln	CTD	N/Y/N
1-00534	*CHD7*†	p.Gln1599*	CTD	Y/Y/Y
1-00230	*KDM5A*†	p.Arg1508Trp	LVO	N/N/Y
1-01965	*KDM5B*†	p.IVS12 + 1 G>A	LVO	N/N/Y
1-01907	*UBE2B*†	p.Arg8Thr	CTD	N/N/N
1-00075	*RNF20*†	p.Gln83*	HTX	Y/Y/Y
1-01260	*USP44*†	p.Glu71Asp	LVO	N/N/N
1-02020	*SMAD2*††	p.IVS6 +1 G>A	HTX	Y/N/N
1-02621	*SMAD2*††	p.Trp244Cys	HTX	Y/NA/N
1-01451	*MED20*	p.IVS2 + 2 T>C	HTX	N/Y/Y
1-01151	*SUV420H1*	p.Arg143Cys	CTD	N/Y/N
1-00750	*HUWE1*	p.Arg3219Cys	LVO	N/Y/N
1-00577	*CUL3*	p.Iso145Phe fs*23	LVO	Y/Y/N
1-00116	*NUB1*	p.Asp310His	CTD	Y/Y/Y
1-01828	*DAPK3*	p.Pro193Leu	CTD	N/N/NA
1-03151	*SUPT5H*	p.Glu451Asp	LVO	N/NA/N
1-00455	*NAA15*	p.Lys336Lys fs*6	HTX	Y/Y/N
1-00141	*NAA15*	p.Ser761*	CTD	N/NA/Y
1-01138	*USP34*	p.Leu432Pro	LVO	N/NA/N
1-00448	*NF1*	p.IVS6 +4 del A	CTD	N/NA/N
1-00802	*PTCH1*	p.Arg831Gln	LVO	N/NA/N
1-02458	*SOS1*	p.Thr266Lys	Other	Y/Y/Y
1-02952	*PITX2*	p.Ala47Val	LVO	N/NA/N
1-01913	*RAB10*	p.Asn112Ser	Other	N/NA/N
1-00638	*FBN2*	p.Asp2191Asn	CTD	N/NA/N
1-00197	*BCL9*	p.Met1395Lys	LVO	N/NA/N
1-02598	*LRP2*	p.Glu4372Lys	HTX	N/NA/N

Gene symbols are as in NCBI RefSeq database. Other structural/neuro/ht-wt denotes presence (Y) or absence (N) of other structural abnormalities, impaired cognitive speech or motor development, and height (ht) and/or weight (wt) less than 5th percentile for age, respectively. Further clinical details in Supplementary Tables 10 and 11. Associated syndromes: *MLL2*, Kabuki syndrome; *CHD7*, CHARGE syndrome; *CUL3*, pseudohypoaldosteronism, type 2E.

* Premature termination mutation.

† Gene involved in production, removal or reading of H3K4 methylation mark.

†† Gene involved in removal of H3K27 methylation mark.

Del, deletion; Dx, diagnosis; fs, frameshift mutation; fs*n*, frameshift mutation followed by premature termination *n* codons later; NA, data not available.

Table S10. Chromatin modifying and other genes of interest with *de novo* mutations in CHD probands

ID	Gene	Heart Exp†	Mutation	Primary Classification: Specific Cardiovascular Diagnoses§	Extracardiac Structural Anomalies	Neuro-Develop-mental	Somatic Growth	
							Ht(%)	Wt(%)
1-00455	*NAA15*	214	p.Lys335Lys fs*6	HTX: Dextrocardia, TAPVR, LSVC, hypoplastic TV, DORV, hypoplastic RV, D-TGA, PS	Hydronephrosis, asplenia, malrotation	normal	50	50
1-00141	*NAA15*	214	p.Ser761*	CTD: TOF, single LCA	No	n/a	<5	20

†Heart expression refers to # reads per million at murine e14.5. Mutation denotes the impact on encoded protein in three letter code; * denotes termination mutation. *Frameshift* mutation in *MLL2*, *CUL3 and NAA15*. 'IVS' stands for intervening sequence. 'fs' stands for frameshift. *Splice site* mutation in *KDM5B, MED20*, and *SMAD2* occur at 1st base of canonical splice donor of intron 12, at 2nd base of canonical splice donor of intron 2 and 1st base of canonical splice donor of intron 6 respectively.

§HLHS-hypoplastic left heart syndrome; Dbl AA-double aortic arch; TOF-tetralogy of Fallot; PAPVR-partial anomalous pulmonary venous return; LSVC-left superior vena cava; LA-left atrium; CAVC-complete atrioventricular canal defect; TAPVR- total anomalous pulmonary venous return; MV-mitral valve; BAV-bicuspid aortic valve; ASD-atrial septal defect; VSD- ventricular septal defect; PA-pulmonary atresia; RAI- right atrial isomerization.

Table S4. All *de novo* mutations with Bayesian QS≥50

												Proband			Father		Mother		
ID	Primary Cardiac Classification	Gene	Mutation	Amino Acid Change	dbSNP	RefSeq NM Accesion IDs	RefSeq NP Accesion IDs	Chr	Position (hg19)	Base change	Mean Heart Exp.	Variant Quality Score	Ref Cov	Nonref Cov	Ref Cov	Nonref Cov	Ref Cov	Nonref Cov	Bayesian Quality Score
1-00141	CTD	NAA15	Nonsense	S761X	Novel	NM_057175	NP_476516	4	140306112	C>A	213.74	196	16	12	30	0	41	0	92
1-00455	HTX	NAA15	Frameshift	D335	Novel	NM_057175	NP_476516	4	140272757	-AAAG	213.74	*Indel-Pass*	68	17	57	0	72	0	100
1-01119	CTD	NAA16	Missense	R70C	Novel	NM_024561	NP_078837	13	41893010	C>T	12.47	228	142	121	234	0	219	0	100

“Model Organisms” to study this in?

- ***S. cerevisiae (yeast)***
- ***S. pombe (yeast)***
- ***C. elegans (worms)***
- ***D. melanogaster (flies)***
- ***M. musculus (mouse)***
- ***Microcebus (mouse lemur- primate)***
- ***C. Jacchus (marmoset- primate)***
- ***Cell lines from the Ogden Syndrome boys***
- ***Induced Pluripotent Stem (iPS) cells - human***

S. cerevesiae (yeast)

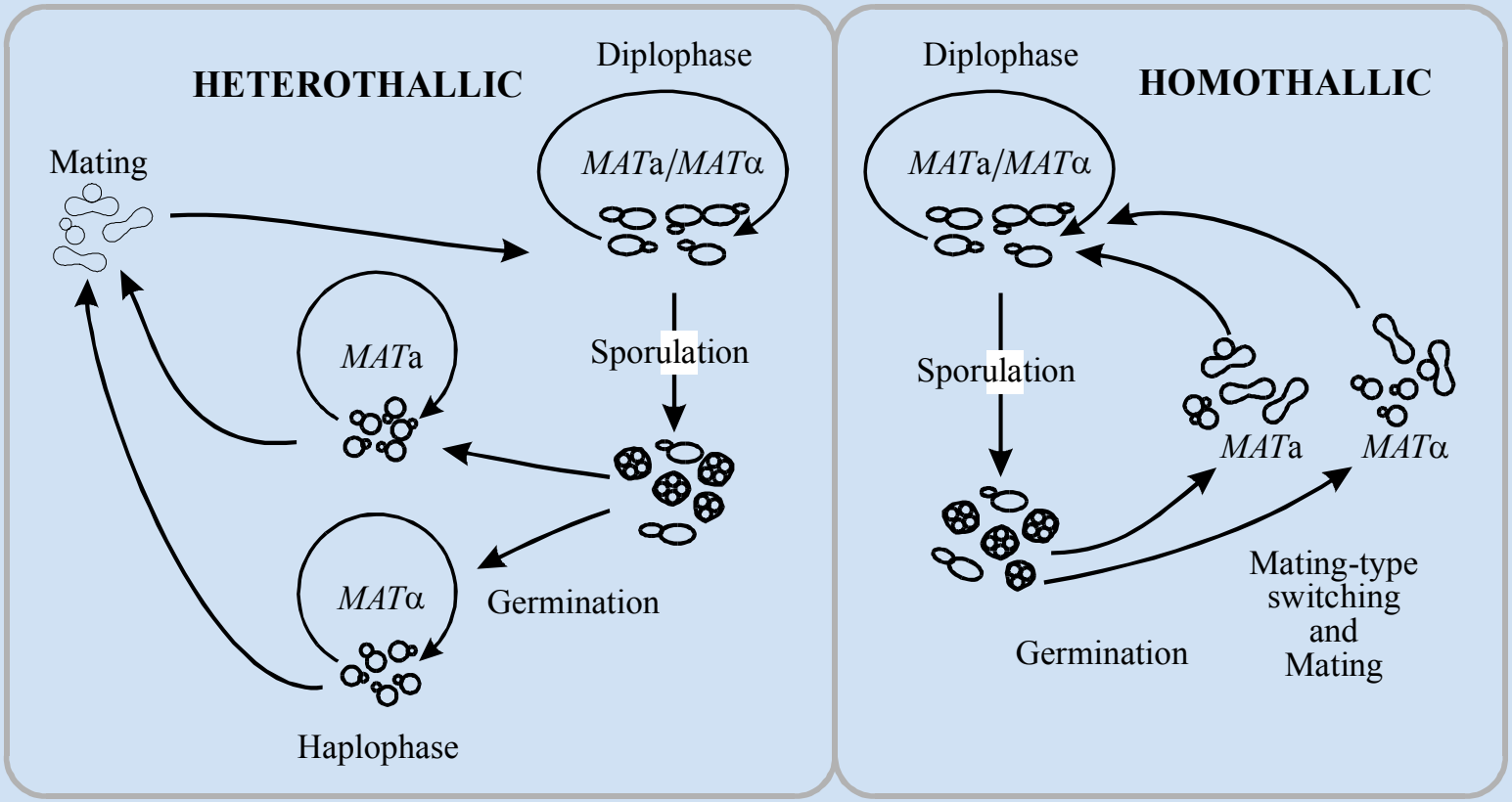
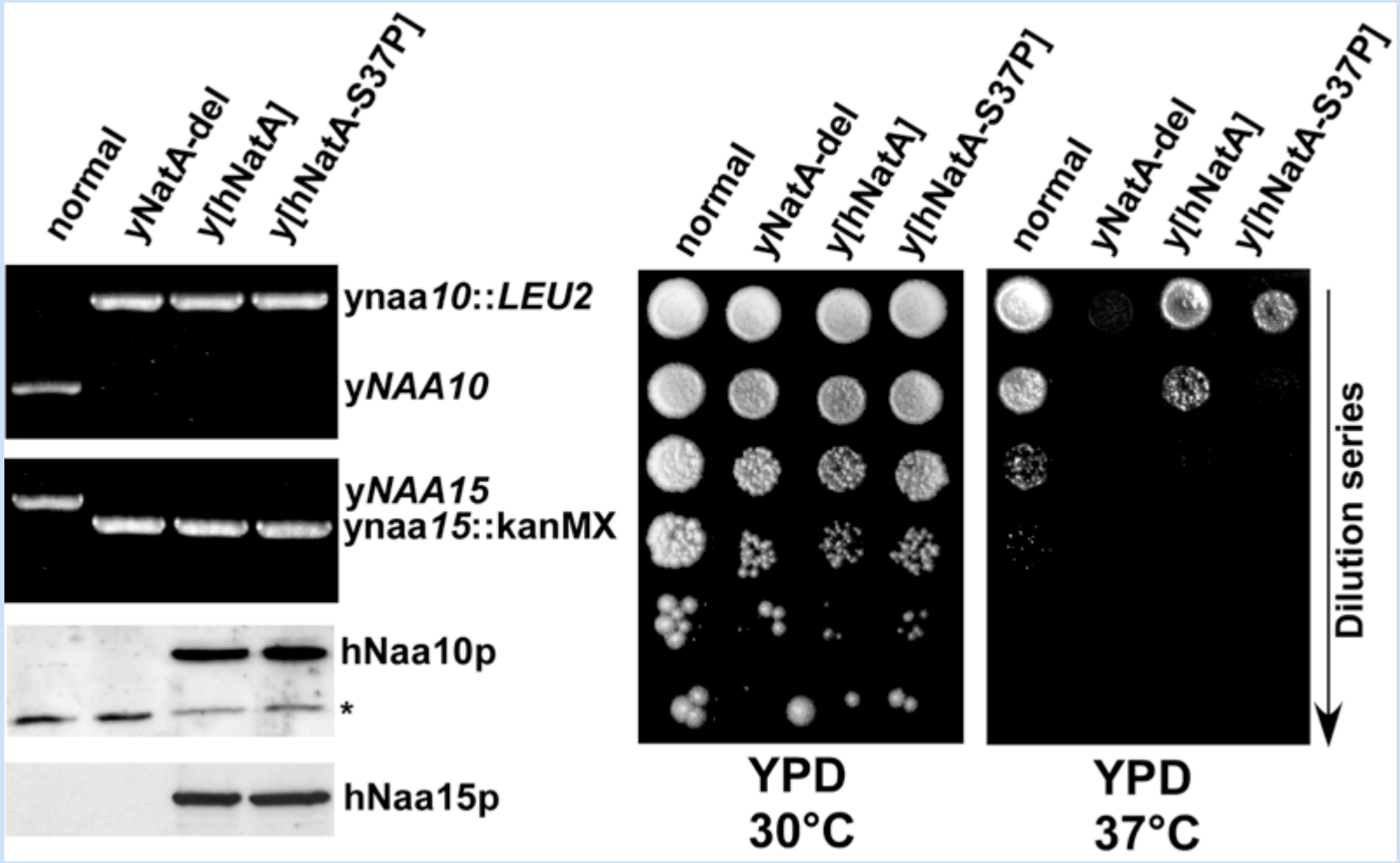


Figure 1. Life cycles of heterothallic and homothallic strains of *S. cerevisiae*. Heterothallic strains can be stably maintained as diploids and haploids, whereas homothallic strains are stable only as diploids, because the transient haploid cells switch their mating type, and mate.

hNaa10p-S37P is functionally impaired *in vivo* using a yeast model.



Unpublished data, do not further distribute.

Open question:
Function of N-terminal acetylation?

Protein stability? Protein secretion?

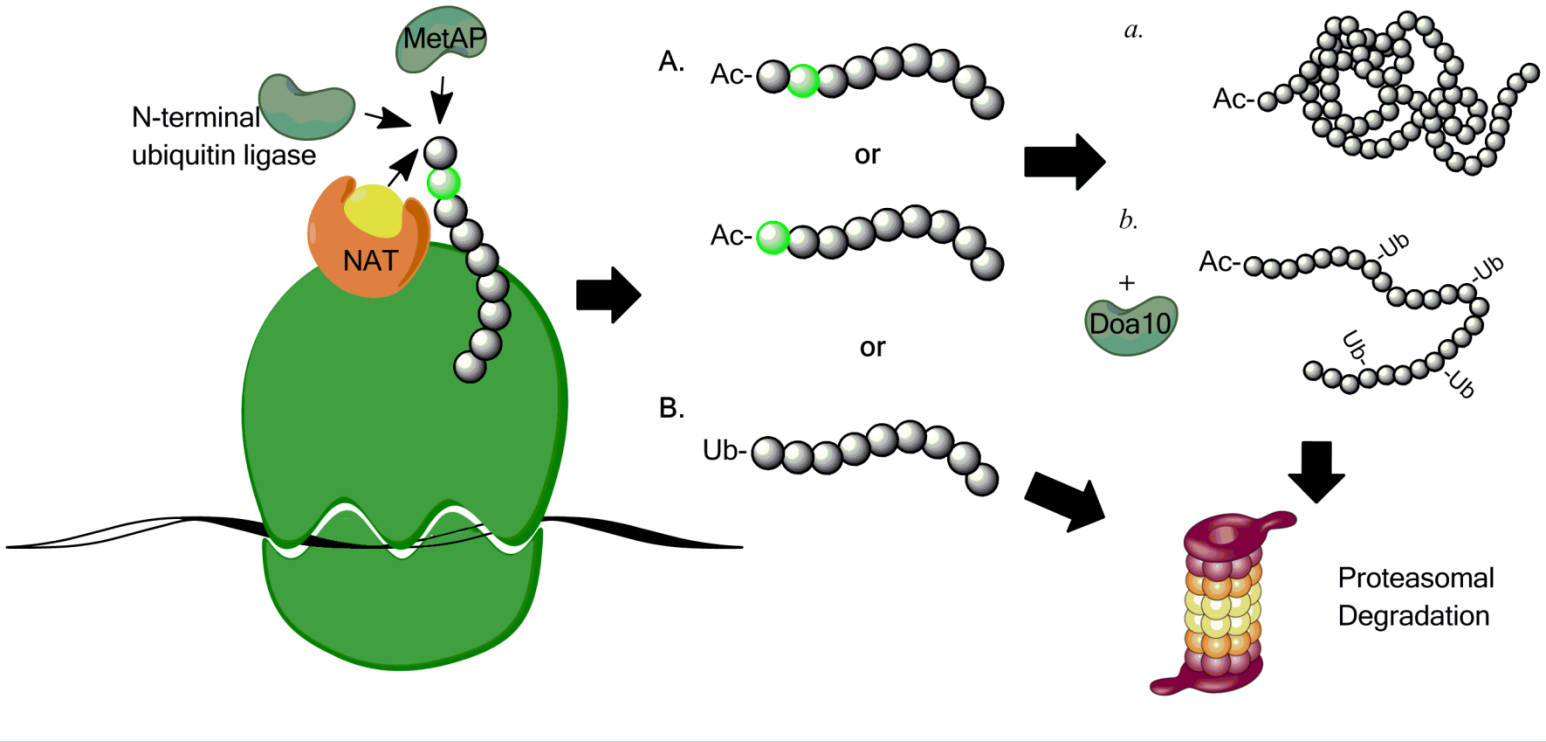
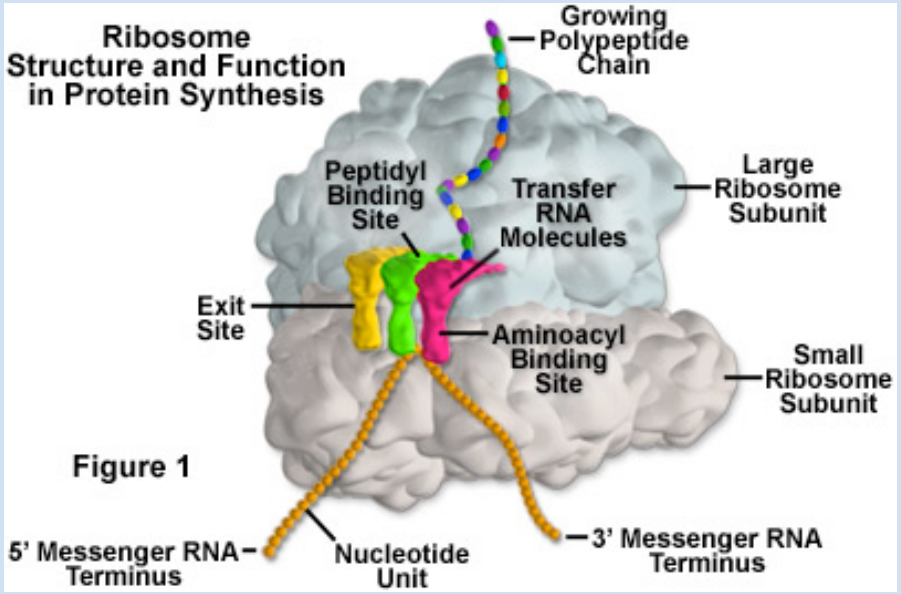
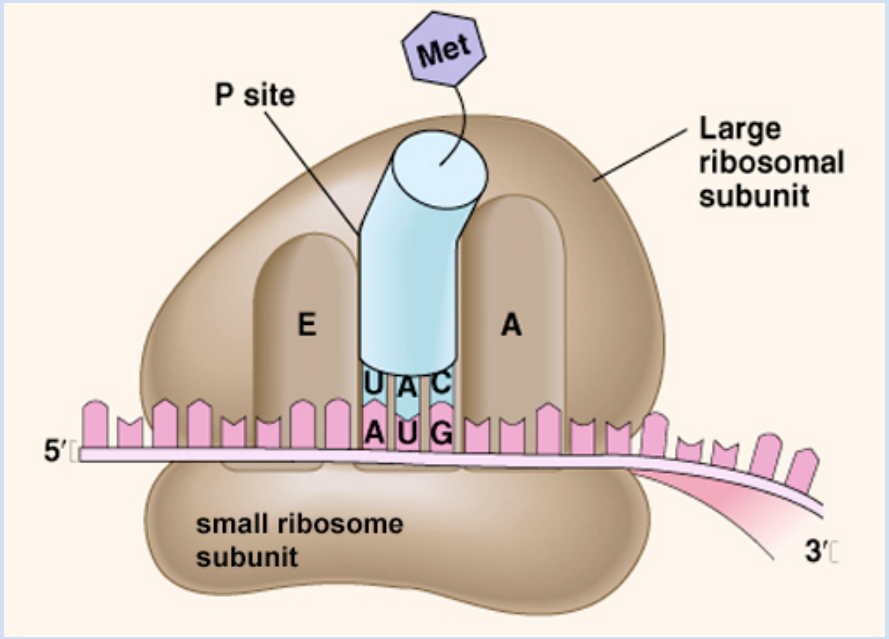
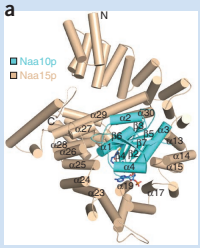


Figure courtesy of Kris Gevaert

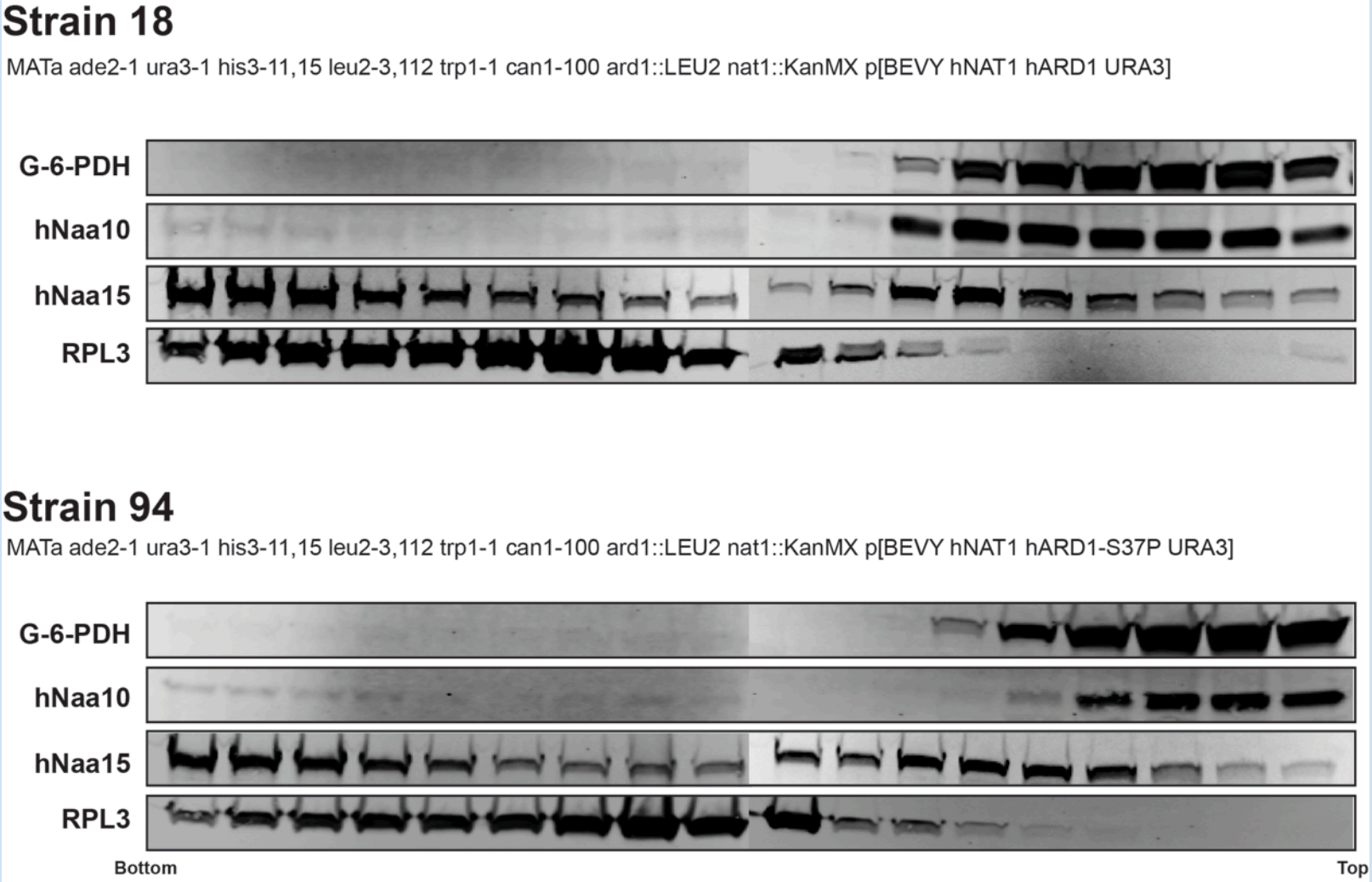

Ribosome
Structure and Function
in Protein Synthesis
Growing
Polypeptide
Chain
Peptidyl
Binding
Site
Transfer
RNA
Molecules
Large
Ribosome
Subunit
Exit
Site
Aminoacyl
Binding
Site
Small
Ribosome
Subunit
Figure 1
5' Messenger RNA
Terminus
Nucleotide
Unit
3' Messenger RNA
Terminus


Met
P site
Large
ribosomal
subunit
E
A
U A C
A U G
5'
small ribosome
subunit
3'

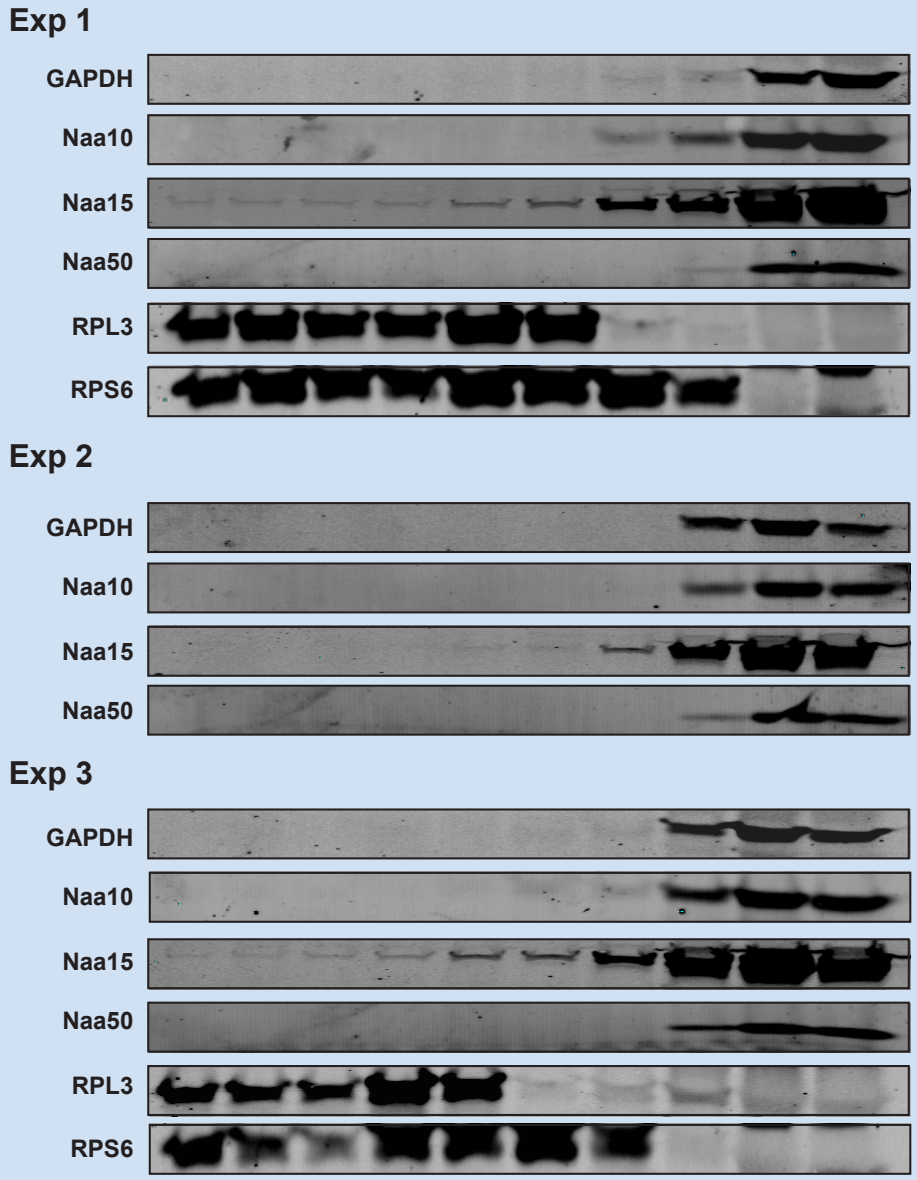

a
N
Naa10p
Naa15p
C

NAA10 + NAA15

Cell Fractionation

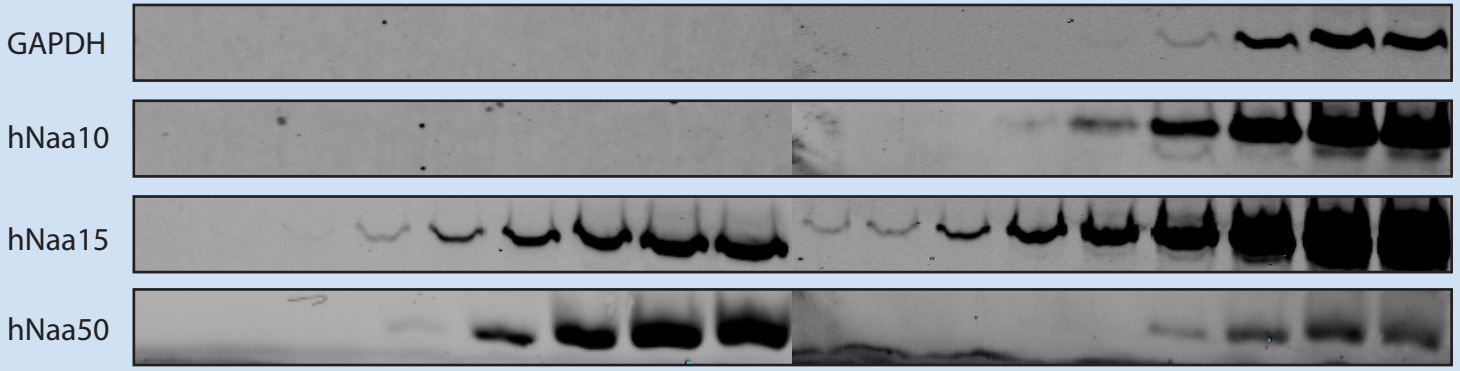


Cell Fractionation with HEK293 cells



Cell Fractionation with HEK293 cells

7-47% Linear Sucrose Gradient
HEK293 Cells
All antibodies against endogenous proteins
Experiment 1 of 3 - 10-12-13
Lysis conditions as per Arnesen 2009, BMC Biochem 10:15.



N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals

Cheol-Sang Hwang, Anna Shemorry, Alexander Varshavsky*

N-Terminal Acetylation Inhibits Protein Targeting to the Endoplasmic Reticulum

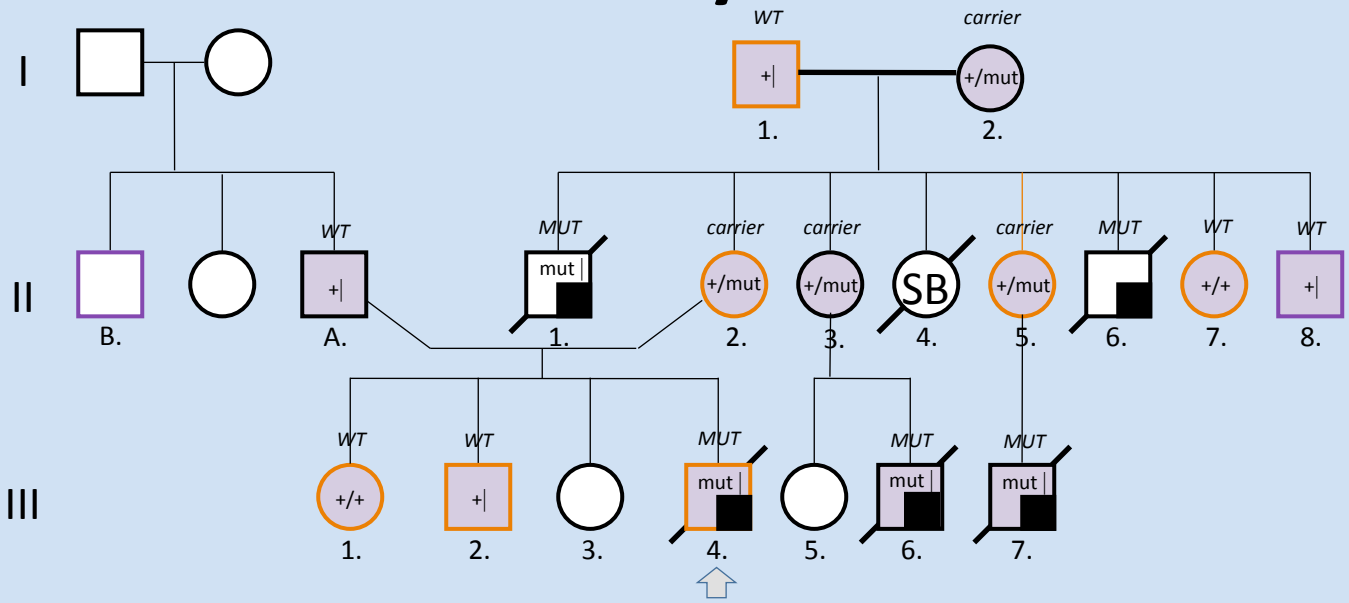
Gabriella M. A. Forte, Martin R. Pool*, Colin J. Stirling*

Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

Abstract

Amino-terminal acetylation is probably the most common protein modification in eukaryotes with as many as 50%–80% of proteins reportedly altered in this way. Here we report a systematic analysis of the predicted N-terminal processing of cytosolic proteins versus those destined to be sorted to the secretory pathway. While cytosolic proteins were profoundly biased in favour of processing, we found an equal and opposite bias against such modification for secretory proteins. Mutations in secretory signal sequences that led to their acetylation resulted in mis-sorting to the cytosol in a manner that was dependent upon the N-terminal processing machinery. Hence N-terminal acetylation represents an early determining step in the cellular sorting of nascent polypeptides that appears to be conserved across a wide range of species.

Proteomics Analysis of EBV-transformed cell lines from family members



= male ♂

= female ♀

= FFPE DNA (for patient III.7.) or DNA from blood available (and for some of them: EBV transformed cell lines available + skin fibroblast of patient III.6.)

SB = stillborn

= proband

= patient samples analyzed by N-terminal COFRADIC analyses **(#1 to #5)**

= patient samples prepared for N-terminal COFRADIC analyses (but still to be analyzed) **(#8 and #9)**

III.4. proband hemizygous, mutant (89323) **(#1a)** (#1b)
II.2. mother of proband, carrier (89324) **(#2)**
II.A. married-in father of proband, WT(89325)
III.2. brother of proband, WT(90526) **(#3)**
III.1. sister of proband, WT (90527) **(#4)**
I.2. grandmother of proband, carrier (90528)
I.1. married-in grandfather of proband, WT(90529) **(#5)**
II.7. aunt of proband, WT (90530) **(#6)**
II.3. aunt of proband, carrier (90531)
II.B. married-in uncle of proband, WT(90532) (#8)
II.8. uncle of proband, WT(90688) (#9)
II.5. aunt of proband, carrier with deceased boy (90797) **(#7)**

N-terminal COFRADIC

PROTOCOL

Selecting protein N-terminal peptides by combined fractional diagonal chromatography

An Staes[1,2], Francis Impens[1,2], Petra Van Damme[1,2], Bart Ruttens[1,2], Marc Goethals[1,2], Hans Demol[1,2], Evy Timmerman[1,2], Joël Vandekerckhove[1,2] & Kris Gevaert[1,2]

[1]Department of Medical Protein Research, Vlaams Instituut voor Biotechnologie (VIB), Ghent, Belgium. [2]Department of Biochemistry, Ghent University, Ghent, Belgium. Correspondence should be addressed to K.G. (kris.gevaert@vib-ugent.be).

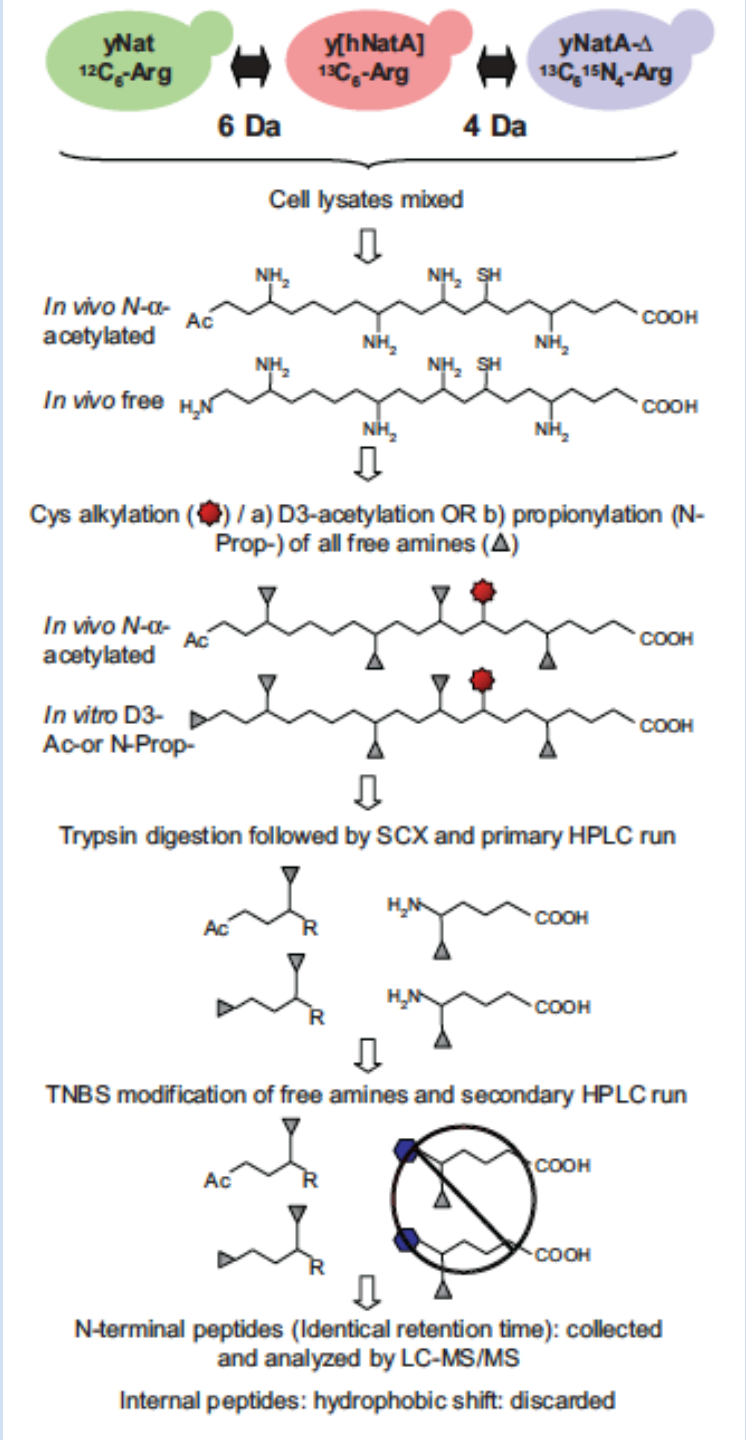
Published online 14 July 2011; doi:10.1038/nprot.2011.355

In recent years, procedures for selecting the N-terminal peptides of proteins with analysis by mass spectrometry have been established to characterize protease-mediated cleavage and protein α-*N*-acetylation on a proteomic level. As a pioneering technology, N-terminal combined fractional diagonal chromatography (COFRADIC) has been used in numerous studies in which these protein modifications were investigated. Derivatization of primary amines—which can include stable isotope labeling—occurs before trypsin digestion so that cleavage occurs after arginine residues. Strong cation exchange (SCX) chromatography results in the removal of most of the internal peptides. Diagonal, reversed-phase peptide chromatography, in which the two runs are separated by reaction with 2,4,6-trinitrobenzenesulfonic acid, results in the removal of the C-terminal peptides and remaining internal peptides and the fractionation of the sample. We describe here the fully matured N-terminal COFRADIC protocol as it is currently routinely used, including the most substantial improvements (including treatment with glutamine cyclotransferase and pyroglutamyl aminopeptidase to remove pyroglutamate before SCX, and a sample pooling scheme to reduce the overall number of liquid chromatography—tandem mass spectrometry analyses) that were made since its original publication. Completion of the N-terminal COFRADIC procedure takes ~5 d.

Staes A *et al.* (**2011**) *Nat. Protoc.* **6**, 1130-1141

Gevaert K *et al.* (**2003**) *Nat. Biotechnol.* **21**, 566-569

Proteomics Strategy With Thomas Arnesen, Petra van Damme And Kris Gevaert



Some downstream substrates for NatA

AAAEEEDGGPEGPNR	66	86	92	91	87	84	91	Q99942	RNF5_HUMAN	E3 ubiquitin-protein ligase RNF5	Membrane; Multi-pass membrane protein. Mitochondrion membrane. Endoplasmic reticulum membrane.
AADTQVSETLKR	52	80	84	84	80	81	85	Q92616	GCN1L_HUMAN	Translational activator GCN1	
AAESALQVVEKLQAR	58	89	92	92	87	87	92	Q14241	ELOA1_HUMAN	Transcription elongation factor B polypeptide 3	Nucleus.
AVFADLDLR	66	95	96	96	96	95	100	P78346	RPP30_HUMAN	Ribonuclease P protein subunit p30	Nucleus;nucleolus.
MVEKEEAGGGISEEEAAQYDR	69	90	91	94	91	95	96	Q9UBE0	SAE1_HUMAN	SUMO-activating enzyme subunit 1	Nucleus.
MLGAPDESSVR	51	79	74	79	70	72	80	Q7Z4S6	KI21A_HUMAN	Kinesin-like protein KIF21A	Cytoplasm;cytoskeleton.
MLSPEAER	74	97	97	97	96	97	97	Q9NUG6	PDRG1_HUMAN	p53 and DNA damage-regulated protein 1	Cytoplasm.
AAGGGGGSSKASSSSASSAGALESSLDR	72	85	84	85	82	84	87	Q5VT52	RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2	
GEEANDDKKPTTKFELER	79	91	93	93	94	93	92	Q92989	CLP1_HUMAN	Polyribonucleotide 5'-hydroxyl-kinase Clp1	Nucleus.

ACS chemical biology

Letters

pubs.acs.org/acschemicalbiology

Design, Synthesis, and Kinetic Characterization of Protein N-Terminal Acetyltransferase Inhibitors

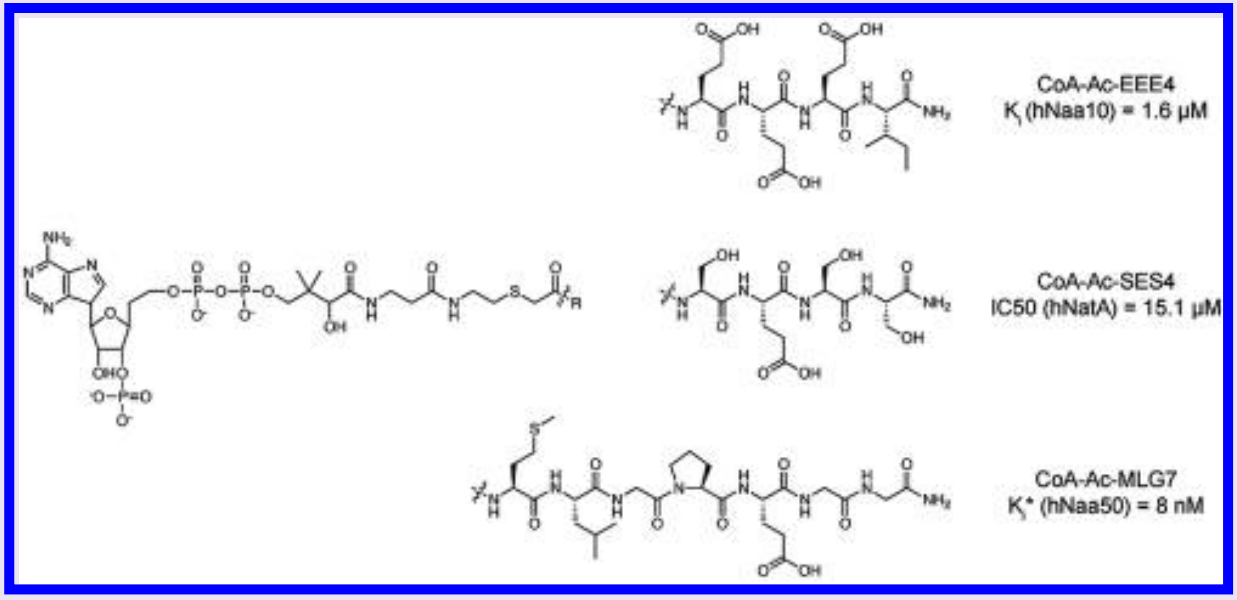
Håvard Foyn,[†,‡,§] Justin E. Jones,[§] Dan Lewallen,[§] Rashmi Narawane,[†] Jan Erik Varhaug,[‡,∥] Paul R. Thompson,[*,§] and Thomas Arnesen[*,†,‡]

[†]Department of Molecular Biology, University of Bergen, N-5020 Bergen, Norway

[‡]Department of Surgery, Haukeland University Hospital, N-5021 Bergen, Norway

[§]Department of Chemistry, The Scripps Research Institute, 120 Scripps Way, Jupiter, Florida 33458, United States

[∥]Department of Surgical Sciences, University of Bergen, N-5020 Bergen, Norway



Summary

- Found first human genetic disease involving Nt-acetylation of proteins
- Characterizing the Nt-acetylation pathway both *in vitro* and *in vivo*
- Working toward identifying interacting components and more downstream substrates of NatA complex.