Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 308 (2003) 1–11
BBRC
www.elsevier.com/locate/ybbrc

Breakthroughs and Views

Composition and function of the eukaryotic N-terminal acetyltransferase subunits☆

Bogdan Polevoda* and Fred Sherman

Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA

Received 23 June 2003

Abstract

Saccharomyces cerevisiae contains three N-terminal acetyltransferases (NATs), NatA, NatB, and NatC, composed of the following catalytic and auxiliary subunits: Ard1p and Nat1p (NatA); Nat3p and Mdm20p (NatB); and Mak3p, Mak10, and Mak31p (NatC). The overall patterns of N-terminally acetylated proteins and NAT orthologous genes suggest that yeast and higher eukaryotes have similar systems for N-terminal acetylation. The differential expression of certain NAT subunits during development or in carcinomas of higher eukaryotes suggests that the NATs are more highly expressed in cells undergoing rapid protein synthesis. Although Mak3p is functionally the same in yeast and plants, findings with TE2 (a human Ard1p ortholog) and Tbdn100 (a mouse Nat1p ortholog) suggest that certain of the NAT subunits may have functions other than their role in NATs or that these orthologs are not functionally equivalent. Thus, the vertebrate NATs remain to be definitively identified, and, furthermore, it remains to be seen if any of the yeast NATs contribute to other functions.

Keywords: Ard1p; Nat1p; N-terminal acetylation; N-terminal acetyltransferase; NatA; NatB; NatC; *Saccharomyces cerevisiae*

N-terminal acetylation

N-terminal acetylation is an enzyme-catalyzed reaction in which the protein N^{α}-terminal residues, such as α-Ser, α-Ala, α-Met, etc., accept the acetyl group from acetyl-CoA. This modification neutralizes positive charges that may influence the protein function, stability, interaction with other molecules, or other subsequent modifications, such as phosphorylation. The reaction is catalyzed by a number of acetyltransferases (NATs) that have been found in all kingdoms, prokaryotes, archaea, and eukaryotes [1,2]. N-terminal acetylation is one of the most common protein modifications in eukaryotes, occurring on approximately 80–90% of the different varieties of cytosolic mammalian proteins, on about 50% of yeast proteins, but rarely on prokaryotic or archaeal proteins [1,3–5]. It is believed that N-terminal acetylation is cotranslational only in eukaryotes but not in prokaryotes, where it occurs posttranslationally. In vitro studies indicated that N-terminal acetylation modification occurs when there are between 25 and 50 residues extruding from the ribosome [3].

Studies with yeast *Saccharomyces cerevisiae* revealed three major N-terminal acetyltransferases, NatA, NatB, and NatC (Table 1), that act on groups of substrates, each containing degenerate motifs [1,2]. As previously summarized, subclasses of proteins with Ser–, Ala–, Gly–, or Thr– termini are acetylated by NatA; proteins with Met–Glu– or Met–Asp– termini and subclasses of proteins with Met–Asn– and Met–Met– termini are acetylated by NatB; and subclasses of proteins with Met–Ile–, Met–Leu–, Met–Trp–, or Met–Phe– termini are acetylated by NatC [2,6]. In addition, a special subclass of NatA substrates with Ser–Glu–, Ser–Asp–, Ala–Glu–, or Gly–Glu– termini, designated NatA′, is

☆ *Abbreviations:* Ac-CoA, acetyl coenzyme A; α-Ser, α-Ala, etc., N^{α}-terminal residues of serine, alanine, etc.; 2D-gel, two dimensional gel electrophoresis; iso-1, iso-1-cytochrome *c*; MS, mass spectrometry; N-terminal, NH_2-terminal; NAT(s), N-terminal acetyltransferase(s); ts, temperature sensitive.

* Corresponding author. Fax: 1-585-275-2683.
E-mail address: Bogdan_Polevoda@urmc.rochester.edu (B. Polevoda).

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0006-291X(03)01316-0

Table 1
The three types of yeast N-terminal acetyltransferases

Type	NatA	NatB	NatC
Catalytic subunit	Ard1p (27 kDa)	Nat3p (23 kDa)	Mak3p (20 kDa)
Auxiliary subunit	Nat1p (99 kDa) Others?	Mdm20p (92 kDa)	Mak10p (84 kDa) Mak31p (10 kDa)
Substrates[a]	Ser– Ala– Gly– Thr– Cys–? Val–?	Met–Glu– Met–Asp– Met–Asn– Met–Met–	Met–Ile– Met–Leu– Met–Trp– Met–Phe–

[a] Acetylation occurs only on subclasses of proteins containing the indicated termini, except for Met–Glu– and Met–Asp– termini, which are apparently always acetylated. (Adapted from [6,8,9,28,31–35].)

only partially acetylated in *nat3*-Δ and *mak3*-Δ mutants [7]. Thus, almost all acetylated proteins can be assigned to one of the NatA, NatB or NatC substrates. Although we do not know of any acetylated proteins in yeast that could not reasonably be a NatA, NatB, or NatC substrate, it remains to be seen if there are other NATs, acting on rarer substrates, or acting on the similar substrates posttranslationally.

Proteins susceptible to N-terminal acetylation have a variety of different N-terminal sequences, with no simple consensus motifs and with no dependence on a single type of residue. Proteins with serine and alanine termini are the most frequently acetylated, and these residues, along with methionine, glycine, and threonine, account for over 95% of the N-terminal acetylated residues [1–4,6]. However, only subsets of proteins with any of these N-terminal residues are acetylated and none of these N-terminal residues guarantee acetylation, indicating that the enzymes recognize some structural characteristics of the N-terminal portion in addition to a particular amino acid at the N terminus [2,7]. Generally, acetylation cannot be definitively predicted from the primary amino-acid sequence. We previously suggested that NATs act on substrates with specific but degenerate sequences and that the activities are diminished by suboptimal residues [2]. We further suggested that the degree of acetylation is the net effect of positive optimal or suboptimal residues and negative inhibitory residues. Because the required and inhibitory residues may affect acetylation to various degrees and because inhibitory residues may occupy various sites in the nascent chain, predicting acetylated and non-acetylated sequences is still not absolutely reliable. Nevertheless, acetylation of many proteins can now be predicted with a high degree of accuracy and the NatB substrates, Met–Glu– and Met–Asp–, are apparently always acetylated [1].

Overall patterns of N-terminally acetylated proteins and orthologous genes possibly encoding NATs suggest that yeast and higher eukaryotes have the same or very similar system for N-terminal acetylation [1]. However, NATs from higher eukaryotes have not been functionally identified and their substrate specificities in vivo have not been characterized. Only limited information is available on the orthologous subunits of NatA in mammals and of NatC in *Arabidopsis thaliana*.

NatA acetyltransferase

NatA is a major NAT in yeast cell with multiple substrates in vivo, resulting in potentially over 2500 protein targets. Among these, approximately 140 substrates were identified by 2D-gel electrophoresis by comparing the corresponding protein spot migration in normal and *nat1* or *ard1* strains followed by sequencing, or by mass spectrometry (MS) analysis, and by comparing the molecular masses of the proteins prepared from normal and mutant strains [1]. This protein list includes some abundant proteins, ribosomal proteins, 26S proteosomal subunits, and mutationally altered iso-1-cytochromes *c*. The NatA substrates appear to be the most degenerate, encompassing a wide range of sequences, especially those with N-terminal residues of serine or alanine. In fact, approximately 90% and 50% of the Ser– and Ala– proteins, respectively, are acetylated [1,3]. NatA activity requires two subunits, Ard1p itself and Nat1p [8]. The previously identified *ard1* mutant was suspected to be related to *nat1* because of certain similar phenotypes. The *nat1* and *ard1* mutants were unable to N-terminally acetylate in vivo the same subset of 24 normally acetylated proteins, including those with Ala– and Ser– termini [6]. In addition to lacking NAT activity, both *nat1*-Δ and *ard1*-Δ mutants exhibited slower growth, temperature sensitivity (ts), slower growth on non-fermentable carbon sources, derepression of the silent mating type gene *HMLα*, failure to enter Go [8,9], defects in sporulation, salt sensitivity, caffeine sensitivity, hydroxyurea sensitivity, and detergent SDS (sodium dodecyl sulfate) sensitivity. Overexpression of both Ard1p and Nat1p subunits is required for increased NAT activity in vivo [8] and both interact with each other to form an active complex in vitro [9].

Also by using biochemical purification procedures another protein, Nat5p (YOR253w), was found to be associated with NatA, but surprisingly the corresponding deletion mutant did not share the phenotypes of *nat1*-Δ or *ard1*-Δ strains [10]. *NAT5* was not required for acetylation of 24 abundant proteins [6], previously identified as NatA substrates by 2D-gel [1], and *nat5*-Δ strains has no obvious phenotype. A deletion of *NAT5*, also designated *ROG2*, was found to suppress the temperature sensitivity of the *mck1 mds1* and the *bul1 bul2* double mutants [11]. *MCK1* and *MDS1* are the two of a total of four genes in yeast that encode glycogen synthase kinase 3 (GSK-3), which plays an important role in both mitosis and meiosis, although the molecular mechanism of its action is not known in detail. Interestingly, *MCK1* acts on transcription of *IME2* at the beginning of meiosis and also on chromosomal segregation processes at mitosis (see also below). Bul1p and Bul2p have been reported to bind and to facilitate the action of Rsp5p, a E3-type family ubiquitin ligase that downregulates a number of proteins in vivo. While its physiological partners and genetic interactions have been explored, the function of Nat5p remains to be determined. Although Nat5p is not required for NatA activity, it is possible that Nat5p acetylates a limited number of proteins other then NatA substrates and that Nat1p serves as a shared subunit in triple heteromeric Nat1p–Ard1p–Nat5p complex (see also below). A direct biochemical purification of the yeast NatA complex revealed that Nat1p, Ard1p, and Nat5p proteins are present in a 1:1:1 stoichiometric ratio [10]. Furthermore, analysis with the tandem affinity purification (TAP) protocol revealed that several proteins, Asc1p, Eno1p, Mis1p, Myo1p, and YGR090w, were associated with the Ard1p; Asc1p also was associated with Nat1p; and Nat1p was found in Spo7p-tagged TAP-purified samples (http://yeast.cellzone.com). However their requirement for NatA activity has not been tested.

In the elegant crosslinking experiments, it was recently found that yeast Nat1p interacts with nascent polypeptide chain and also with the translating ribosome, indicating that Nat1p may serve as an anchor of NatA complex to ribosome [10]. Interestingly, Nat1p was located in close proximity to the nascent polypeptides and interacted with a variety of ribosome-bound nascent polypeptides, independently of whether they were substrates for NatA acetylation or not. Another subunit of NatA, Ard1p, did not interact directly with their substrates nor did it bind to ribosomes and it failed to form a crosslink product with the nascent polypeptides. Similarly, Nat5p did not directly interact with ribosomes or nascent polypeptides. Considering the trimeric NatA structure, one can suggest that during translation on ribosomes, Nat1p recognizes whether the N-terminal sequence of nascent polypeptide is a substrate for Ard1p or Nat5p and directs appropriate substrates to the corresponding catalytic subunit for acetylation. However, the unknown nature of the Nat5p substrates complicates the interpretation. It should be noted that if Nat5p modifies an internal amino-acid residue, for example N^{ε}-lysine, then unacetylated proteins from *nat5*-Δ mutants may not be detected by differences in migration on 2D-gels.

Both *NAT1* and *ARD1* influence the transcriptional silencing at telomeres in yeast [12,13]. NatA complex may function in the same pathway with the well-characterized Sir-proteins, *SIR2*, *SIR3*, and *SIR4*, that were shown to be involved in silencing together with *HHF2* (histone H4). Interestingly, the human *ARD1* ortholog gene, designated TE2, is located on X-chromosome that is involved in sex determination [14]. As mentioned above, the mutations in *NAT1* and *ARD1* genes caused the defects in mating due to derepression of the silent loci and prevented yeast cells from displaying G1-specific growth arrest in response to nitrogen deprivation [15]. The overexpression of the *SIR1* gene suppresses such defects of *nat1* and *ard1* mutants [13]; however it is not clear whether any of the Sir-proteins is a substrate for NatA or its acetylation affects Sir-function. *NAT1* also was identified in a yeast genome-wide search for genes that affect chromosome stability when overexpressed [16]. It was suggested that *NAT1*overexpression might act by dominant negative mechanisms preventing the formation of active Nat1p–Ard1p heterodimer. Similarly, the deletion of *NAT1* caused the chromosome loss and temperature and hydroxyurea sensitivity. Furthermore, it was reported that *NAT1* and *ARD1* in yeast are involved in DNA repair and recombination via DNA end joining.

Collectively, these NatA study results suggest that NatA acetylation is important for function of an unknown set of proteins involved in general growth control; cell growth on non-fermentable carbon sources; transition to Go phase growth; sporulation; cell growth on salt, caffeine, hydroxyurea, and SDS containing media; and derepression of silent loci. The latter phenomenon may occur via acetylation of the certain Sir-protein(s) and other chromatin-associated proteins or less likely the transcription factors and histones to regulate the DNA-remodeling complexes. Interestingly, in contrast to yeast, the *ARD1* ortholog is an essential gene in *Trypanosoma brucei*, necessary for viability in mammalian and insect-stage cells, as it was found in an attempt to study a possible involvement of NatA in telomeric silencing [17].

The role of NatA in cell proliferation, development, and cancer

Based on the overall patterns of N-terminally acetylated proteins and the NAT orthologous genes, we

suggested previously that yeast and higher eukaryotes have the same systems for N-terminal acetylation [1]. Indeed, the Nat1p and Ard1p homologs could be found in all genomes of the model organisms. Moreover, as in yeast, in higher eukaryotes Nat1p orthologous proteins have been implicated in several essential biological phenomenons, like tissue development, cell proliferation, and cancer.

Xenopus laevis Nat1p, designated Xat-1, was isolated in a screen to identify the mRNAs with stage-specific gene expression during early embryogenesis by using differential display [18]. Xat-1 transcripts were expressed at relatively constant levels throughout early embryonic stages, but also exhibited a dynamic pattern in brain, somites, branchial arches, pronephros, and otic vesicles. Xat-1 expression was highly concentrated in the neuroepithelial cells lining the wall of brain cavity, in contrast to axon-enriched marginal zone in the brain.

Similarly, a mouse mNAT1 (*NARG1*) was significantly expressed in the ventricular zone and intermediate zone in the developing brain and shown to be regulated by physiological levels of *N*-methyl-D-aspartate (NMDA) receptor function in developing neurons in vivo [19]. NMDA receptors are the class of glutamate receptors that transmit signals between neurons. *NARG1* was highly expressed in the neonatal brain in the regions of neuronal proliferation and migration, but is dramatically downregulated during early postnatal development.

A mouse *NAT1* homolog, designated in an other study *tubedown-1* (*tbdn-1*), was isolated as a gene highly expressed in unstimulated IEM cells during blood vessel development [20]. Its expression was high in developing embryos and placenta. Generally, in adult *tbdn-1* expression was low or undetected in most organs examined with exception of atrial endocardium, endothelium of retinal vessels, bone marrow tissue, and ovary [20]. The protein was also downregulated in myeloid leukemia cells after differentiation in response to leukemia inhibitory factor in vitro. Thus, in mammals NatA may play a role in neuronal maturation as NMDA receptor-regulated gene [19], and in vascular and hematopoietic development [20].

Human homolog of *NAT1*, designated NATH, was found to be highly expressed in a study to identify the genes differentially expressed in papillary thyroid carcinomas (PTC) [21]. It was also overexpressed in other tumors, especially clinically aggressive and with histological evidence of poorly differentiated or undifferentiated areas. Normally, NATH mRNA is expressed at low level in most adult tissues, including the normal thyroid, liver, pancreas, mammary and salivary glands, lung, ovary, urogenital system, and upper gastrointestinal tract. However, a somewhat higher level of its expression was seen in parts of human brain, the heart, bone marrow, and in several leukemia and carcinoma cell lines. The highest level was found in adult testis and Burkitt's lymphoma cell line. As it was noted previously, the NATH expression data with elevated mRNA levels in aggressive carcinomas primarily suggest that NATH may be involved in cellular proliferation [21]. The data on mouse *tbdn-1* and *NARG-1* clearly support such notion [19,20] (see the paragraph above). However, we believe that this could be explained by the higher translation rates in actively growing cells. Because NatA acetylation is cotranslational and *NARG-1* is expected to be cooperatively expressed with other genes involved in protein translation, it is not surprising to observe a higher level of *NAT1* orthologs in proliferating tissues. Therefore, the elevated levels of NATH in cancer cells probably are not a primary, but a secondary effect, and are due to a NAT role in the modification of the newly synthesized proteins. For this reason *NAT1* expression is low in most differentiated tissues (see above). Similarly, in yeast *NAT1* expression is significantly downregulated in stationary phase or after diauxic shift when the translation levels are low (*NAT1* basic information in: The Saccharomyces Genome Database (SGD); http://www.yeastgenome.org/).

Another important observation made previously is that although NATH is highly expressed in cancer cell lines, the transfection of embryonal kidney and papillary thyroid carcinoma cells by NATH did not alter significantly the cellular proliferation rate [21]. Although it is not completely relevant here, this indicates that the overexpression of NATH alone may not be sufficient and requires another subunit for the specific activity. This is consistent with experiments in yeast where the concomitant expression of Nat1p and Ard1p was required for a higher acetyltransferase activity [9]. At present, there is no information on NatA subunit composition in higher eukaryotes, although Ard1p and Nat1p orthologs could be easily deciphered in the genomes of sequenced eukaryotic organisms.

Additional Nat1p and Ard1p protein characteristics

Several studies predicted that NATH protein and the other Nat1p orthologs contain at least four tetratricopeptide repeats (TPRs) in two tandems located between amino-acid residues 46–113 and 374–441 (NATH coordinates), respectively [18,20,21]. TPRs are degenerate 34 amino-acid repeats often associated with protein–protein interactions, present in diverse proteins widespread among organisms from all kingdoms, and are important for cell cycle regulation, transcription, and chaperone function [22]. A single TPR contains two antiparallel α-helices that pack into an open structure. It has been proposed that TPR proteins preferentially interact with WD-40 repeat proteins, but in many instances several TPR-proteins seem to form multi-

protein complexes. There are 24 predicted TPR-containing proteins and 93 WD-40 containing proteins coded by the yeast genome [23]. Prominent examples of TPR-proteins include yeast Cdc16p, Cdc23p, and Cdc27p, the components of cyclosome; Pex5p-receptor for peroxisomal targeting signals; Tom70p co-receptor for mitochondrial targeting signals; and Ser/Thr phosphatase 5C.

Protein sequence analysis of yeast Nat1p showed the presence of up to 7 TPR motifs, 4 or 5 of those with a strong match to the conserved amino-acid residues (Fig. 1, see also the last section). Similar to the structure of TPR-containing Pex5p protein, Nat1p interaction with nascent polypeptide might be mediated by its TPR motifs. It may also serve as the site(s) for interaction with the other NatA subunits, Ard1p and Nat5p.

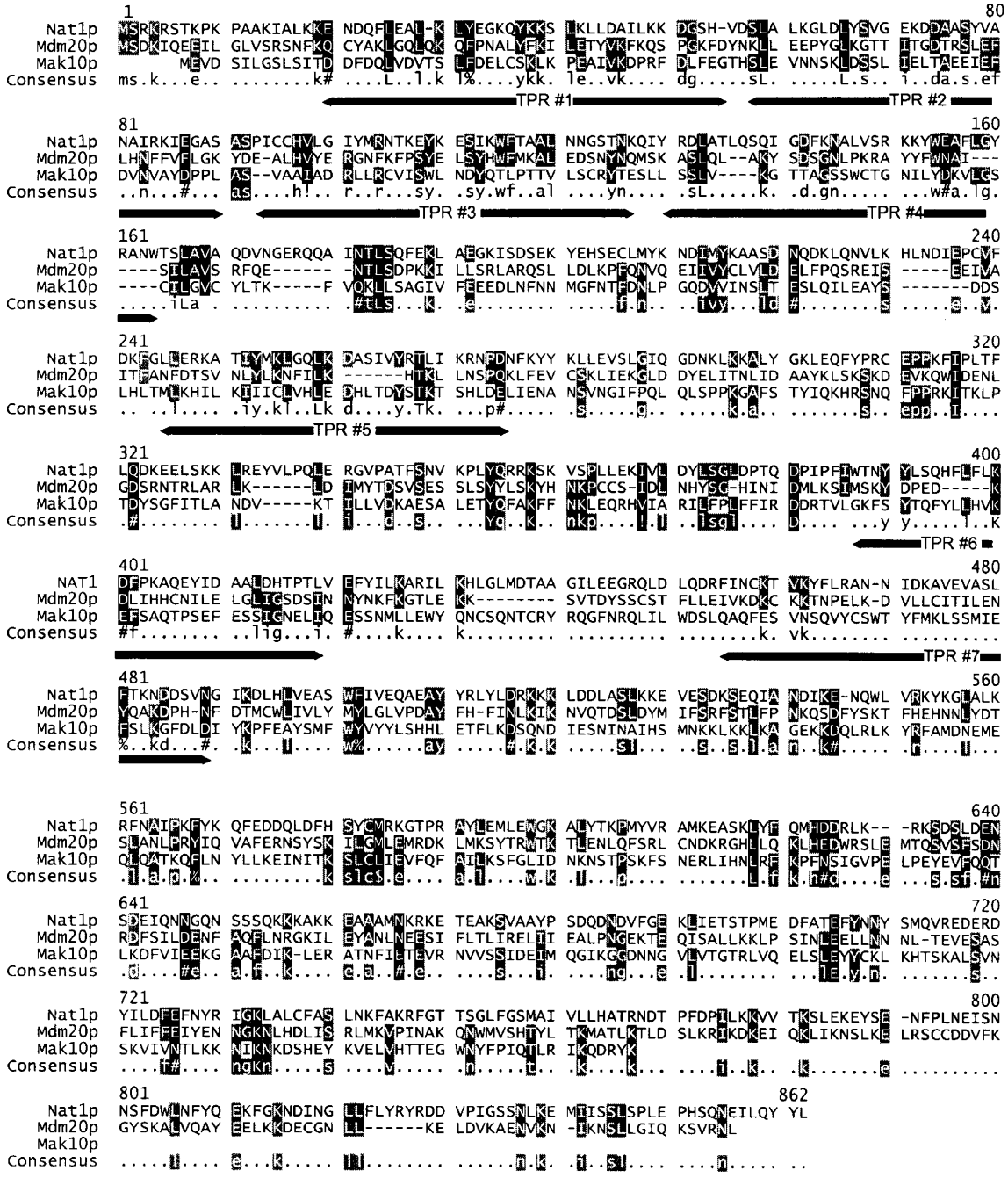

Fig. 1. Protein sequence alignment of the yeast auxiliary subunits Nat1p, Mdm20p, and Mak10p, from NatA, NatB, and NatC, respectively. Protein sequences were aligned by Multalin version 5.4.1. Highly conserved residues are highlighted in black, whereas moderately conserved residues are highlighted in gray. Consensus symbols are: ! is anyone of IV; $ is anyone of LM; % is anyone of FY; and # is anyone of NDQEBZ. The potential TPR motifs in Nat1p, TPR #1 to TPR #7, are indicated by black bars at the bottom of the sequence.

However, the functional significance of the TPR motifs in Nat1p remains to be investigated. In addition, the highly charged Nat1p regions between 550–590 and 620–670 with predicted coiled coil structures could be important for protein–polypeptide interaction or interaction with other subunits. On the other hand, C-terminal region of Ard1p containing a potential coiled coil domain may be involved in the formation of Ard1p–Ard1p homodimer and Ard1–Nat1p complexes, as C-terminally truncated Ard1p did not form a stable complex with Nat1p [9]. Park and Szostak [9] also pointed out the weak sequence similarity of the Ard1p protein to the homeodomain, which, however, may not be functionally significant.

A short protein region in *tbdn-1*, the mouse Nat1p ortholog, was previously noted to be similar to the polycomb-like PHD finger B DNA-binding domain [20]. In addition, *tbdn-1* was shown to undergo autoacetylation, as revealed in immunoprecipitation experiment [18]. However, the similarity of the *tbdn-1* segment to the PHD domain is weak and it is unknown if it has any functional significance. Gendron et al. [20] suggested that *tbdn-1* autoacetylation may facilitate DNA binding, a process that often occurs with transcription factors possessing acetyltransferase activity [24]. The analysis of *tbdn-1* autoacetylation site suggested that it probably occurs on the internal lysine residue(s) located at amino-acid residues 377–398 and suggested that *tbdn-1* may be a N^{ε}-acetyltransferase [20]. On the other hand, the autoacetylation of *tbdn-1* in the immunoprecipitated sample could be explained by the presence of the catalytic subunit, Ard1p or Nat5p orthologs, or other proteins present in the sample, considering that the mouse NatA complex is similar in composition to the yeast NatA. Since Nat1p does not contain a region with significant similarity to an acetyl coenzyme A (Ac-CoA) binding site, then the only reasonable explanation for the observed *tbdn-1* autoacetylation is that the mouse Ard1p or Nat5p can acetylate at the internal ε-lysine residue (see also below). It would also be of interest to determine whether Nat1p contains any DNA-binding motifs, such as a bromodomain.

Another structural feature of Nat1p orthologs is the presence of bipartite nuclear localization signal (NLS), as, for example, at amino-acid residues 594–611 of Xat-1 [18]. A putative bipartite NLS signal was found also in human NATH [20] between amino-acid residues 612–628, KKNAEKEKQQRNQKKKK. However, NATH-myc fusion protein was predominantly localized in cytoplasm and only rarely in the nucleus of the transfected cells [21]. Similarly, a predicted NSL site could be predicted in the yeast Nat1p, KKKAKKEAAAMN KRK, at positions 647–662. Furthermore, potential NLS signal, KKVLKLEELQISNFTHRRLR, could be detected in Ard1p between residues 190 and 210. However, a systematic investigation of the yeast protein localization revealed that both Nat1p and Ard1p proteins localize primarily to cytoplasm, although some faint staining was observed in the nucleus, especially for Nat1p (TRIPLES: a Database of Gene Function in *S. cerevisiae*, http://ygac.med.yale.edu/triples/triples.htm). The cellular localization of TE2, a human ortholog of Ard1p [14], is presently unknown. Obviously, additional studies are needed to clarify this issue, including the use of biochemical fractionation techniques for NatA subunit localization.

Do NatA subunits of higher eukaryotes act in N^{ε}-acetylation?

Most interestingly, the human homolog of Ard1p, TE2, was reported to acetylate an internal lysine of hypoxia-inducible factor, HIF-1α, resulting in the regulation of its stability (HIF-1α plays a central role in cellular adaptation to changes in oxygen availability) [25]. ARD1-mediated acetylation enhances interaction of HIF-1α with pVHL ubiquitination complex and HIF-1α ubiquitination, suggesting that the acetylation of HIF-1α by ARD1 is critical for its proteasomal degradation [25]. Biochemical experiments demonstrated that ARD1 acetylated Lys532 of HIF-1α protein and that the acetylated site could be detected by Ac-Lys monoclonal antibody, although no evidence from in vivo studies was presented. It is surprising that one protein (57% sequence similarity between yeast and human Ard1p) can act in both N^{α}-acetylation and N^{ε}-acetylation. Hence, based on this report and the data on *tbdn-1* autoacetylation (see above), one can conclude that mammalian ARD1 can function as N^{ε}-acetyltransferase. It would be of interest to determine if the acetylation of HIF-1α mediated by mammalian ARD1 is reversible, which is a typical feature of N^{ε}-acetylation, but not N^{α}-terminal acetylation. Second, it is important to determine whether ARD1 can act on both N^{ε}-type and N^{α}-type substrates in mammals or only N^{ε}-substrates, or whether it acts on different substrates when recruited by different acetyltransferase complexes to modify different targets. In the latter case, the set of interacting proteins should be different. If ARD1 acts only on N^{ε}-substrates, then its role has diverged from the one in yeast and the new catalytic subunit of NatA N^{α}-terminal acetyltransferase remains unidentified in mammals.

Moreover, recently a 100-kDa variant of Tubedown-1, named Tbdn100, was identified as a constituent of the purified homeoprotein Msx2-regulated DNA-binding OCFRE (osteocalcin FGF response element—the promoter region of the osteocalcin gene) complex from osteoblastic cell line, which also contained Ku70 and Ku80, the multipurpose nucleic acid binding proteins [26]. Based on the fact that fibroblast growth factor treatment regulates Ku, but not Tbdn100 accumulation,

it was proposed that Tbdn100 might mediate the stable assembly of the complex at OCFRE site. The protein is expressed at detectable levels in heart, skeletal muscle, and testis, with much lower levels in other tissues that are similar to those of the other mammalian Nat1p orthologs (see the previous section). Importantly, Western blot analysis confirmed the accumulation of the Tbdn100 in nuclear fractions prepared from osteoblasts. This is essentially a first report on a possible function of Tbdn100 as a transcriptional co-regulator, cooperating with the Ku antigen and Runx2 to up-regulate transcription from the OC promoter. Considering the *tbdn-1* autoacetylation described above [20], the protein may function as a co-adapter that mediates transactivation in part via regulation of chromatin or associated protein acetylation, as described for other transcription factors [24]. In this respect and taking into account the results on human ARD1 presented above, it is reasonable to reconsider the role of NatA complex in derepression of silent loci in yeast. If the yeast Nat1p indeed functions similar to the mammalian ortholog, Tbdn100, then NatA would be expected to acetylate proteins not only at N^{α}-termini, but also at certain internal lysine residues. In this regard, it is tempting to speculate that NatA acetylates certain internal lysine residues of certain Sir-proteins or other proteins involved in gene silencing.

The *Drosophila melanogaster* protein, SAN (Accession GI: 6980078), a yeast Nat5p ortholog, was recently found to have a function in sister chromatid cohesion (SCC) (B. Williams, personal communication). A *san* mutation resulted in a SCC disruption in mitosis, as well as in metaphase arrest. The substrates of this putative acetyltransferase have not been determined, although SAN overexpressed in bacteria or baculovirus had no acetyltransferase activity on native histones, suggesting that proteins involved in SCC may be the targets. SAN was found to be a part of a 150–200 kDa protein complex containing an associated 25 kDa subunit GI: 24662081, an Ard1p homolog, as well as a 100 kDa subunit GI:24643243, a Nat1p homolog. Thus, similar to yeast, the corresponding *Drosophila* orthologs, Ard1p, Nat1p, and Nat5p, form a complex.

Importantly, two new proteins critical for SCC in *Drosophila* were uncovered in the initial screen: first, the already mentioned SAN; and second, a protein designated "Deco", which is the *Drosophila* ortholog of yeast Eco1p that was previously shown to acetylate internal lysine residues of Scc1p, Scc3p, and Pds5p, proteins involved in SCC of yeast [27]. The phenotypes of both *san* and *deco* mutations were nearly identical, in which there were early sister chromatid separation and activation of the spindle checkpoint. Obviously, it would be important to determine if the yeast Nat5p and the corresponding mammalian ortholog are involved in the acetylation of internal lysine residues, and if deficiencies of these proteins affect SCC.

NatB acetyltransferase

NatB acetyltransferase contains the catalytic subunit Nat3p, [6] and the auxiliary subunit Mdm20p that is also required for NatB activity [28]. *MDM20* was identified as a gene that is necessary for mitochondrial inheritance and organization of the actin cytoskeleton [29]. Some of the NatB substrates in yeast were identified by different techniques and include the following: two essential proteins actin [6] and tropomyosin [30]; small subunit of ribonucleotide reductase Rnr4p [6]; ribosomal proteins S21 and S28 [7]; 20 S proteosomal subunit Pre1p; and 19 S proteosomal subunits Rpt3p and Rpn11p [31]. The NatB substrates have common sequences with Met–Asp– or Met–Glu– N-termini that can be easily deciphered and normally are modified. Indeed, all Met–Asp– or Met–Glu– eukaryotic proteins that have undergone N-terminal analysis were demonstrated to be N-terminally acetylated [1].

The *nat3* and *mdm20* deletion mutants showed similar if not identical multiple phenotypes, including the following: slow growth; temperature and osmotic sensitivity; deficiency in utilization of non-fermentable carbon sources; reduced mating efficiency; inability to form functional actin cables; defects in mitochondrial and vacuolar inheritance; random polarity in budding; sensitivity to the antimitotic drugs; and susceptibility to the number of DNA damaging agents [28]. In fact, among the three NATs, mutants of the NatB subunits, the *nat3* and *mdm20* mutants, have the most severe phenotypes. Furthermore, the comparative phenotype analysis of the NatB mutants with *act1* and *tpm1* mutants altered in the N-terminal protein regions indicate that the *nat3*-Δ and *mdm20*-Δ phenotypes are due primarily, if not entirely, to the lack of actin and tropomyosin acetylation [28]. We have also demonstrated that unacetylated actin and actins altered in the N-terminal region have similar defective properties in vitro.

As mentioned above, NatB acetylates two essential proteins, actin and tropomyosin, in yeast. Acetylation of these proteins is important for tropomyosin–F-actin interactions, which are required to stabilize actin filaments that are bundled together to form actin cables [29]. Tropomyosin must be acetylated before it can be associated with actin filaments in vivo. The lack of actin acetylation itself leads to significantly decreased actomyosin interactions and sliding velocity, disrupting the normal cytoskeleton function [25]. The lack of both actin and tropomyosin acetylation causes defects in the transport from the mother cell to the bud of a variety of membrane-bound organelles, including mitochondria, vacuoles, peroxisomes, Golgi elements, and secretory vesicles. In addition, it probably affects nuclear division and the functioning of the DNA-chromatin-containing complexes, as revealed by the sensitivity of *nat3*-Δ and

mdm20-Δ mutants to the antimitotic drugs and the DNA damaging agents [28].

The Nat3p and Mdm20p are conserved proteins and corresponding orthologous proteins are readily identified in the genomes of all model eukaryotic organisms, including *A. thaliana*, *Caenorhabditis elegans*, *D. melanogaster*, *Homo sapiens*, *Mus musculus*, *Schizosaccharomyces pombe*, and *X. laevis* [28]. However, at present none of the orthologous proteins were investigated for NatB acetyltransferase activity or function in vivo.

NatC acetyltransferase

MAK3 gene that encodes a catalytic subunit of NatC was first described as required for the N-terminal acetylation of the viral major coat protein, *gag*, with an Ac–Met–Leu–Arg–Phe– terminus [32,33]. The co-purification of Mak3p, Mak10p, and Mak31p suggested that these three subunits form a complex. Moreover, protein–protein interactions between Mak3p and Mak10p, as well as between Mak31p and Mak10 were detected in two-hybrid screen. Furthermore, we have shown that all three subunits are required for NatC activity but not for acetylation of NatA or NatB substrate types [34]. In addition, all the three deletion strains showed similar phenotypes, including slower growth on non-fermentable carbon sources at elevated temperature. This suggests that some mitochondrial proteins might be the substrates for NatC in addition to two subunits of 26S proteasome, Pup2p and Pre5p, and L-A virus *gag* protein [31].

MAK10 gene was identified as a gene necessary for the propagation of the L-A dsRNA virus in yeast [35]. Also it has been shown that it is essential for optimal growth on non-fermentable carbon sources, like glycerol or ethanol, and such effect is independent of its effect on L-A virus. In addition, *MAK10* is expressed at a low level and is glucose repressible. It was noted that Mak10p protein sequence contains several regions of similarity to T-cell receptor α-subunit V (variable) regions [35], but it remains to be seen if this similarity has any functional significance. The mammalian Mak10p ortholog, designated T4a, was identified in a study directed to uncover genes upregulated in the healing corneal epithelium [36]. The high expression of this gene was also found in cultured, autonomously replicating embryonic and neointimal smooth muscle cells, but not in normal adult cells, where the expression remained low [37]. As in case of mammalian orthologs of Nat1p, it is likely that the higher expression of T4a in actively proliferating cells is related to its cotranslational mode of action.

Although biological function of *MAK31* gene is not known, the protein sequence indicates that it could be characterized as a Sm-like protein [38]. The Sm-like proteins are small nuclear ribonucleoproteins (snRNPs) found to be associated with U1, U2, and U5 snRNAs as well as with U4/U6 double snRNP and the U4/U6–U5 triple snRNP. snRNPs are involved in various functions including pre-mRNA splicing, histone mRNA 3′ formation, tRNA processing, rRNA maturation, and telomeric DNA synthesis. Mak31p is more divergent than the other yeast Sm-like proteins because it is the only one that does not contain a glycine or cysteine residue at position 107 and did not precipitate with any of the tested RNAs [39]. It is puzzling that Mak31p, a required subunit of a NAT, is highly similar to specialized protein family having different functions and that analogous subunits were not observed in the NatA and NatB complexes. Mak31p may function in determining the specificity of protein substrates or of attachment sites to the ribosome.

Protein sequence analysis revealed that species containing orthologs of the yeast Mak3p include *A. thaliana*, *C. elegans*, *D. melanogaster*, *H. sapiens*, *M. musculus*, and *S. pombe* [1], although the proteins forming this subfamily are more diverged compared to Ard1p and Nat3p subfamilies. Also the proteins with high similarity to Mak3p are present in some bacteria and archaea, but it is doubtful that they represent true N-terminal acetyltransferases that act co-translationally on a wide range of proteins because prokaryotic proteins are rarely acetylated. The Mak10p and Mak31p orthologs were also detected in all model higher eukaryotes as well, but not in prokaryotes, with exception of Sm-like proteins with unknown function in archaea.

Recently At*MAK3*, an *A. thaliana* gene orthologous to *MAK3* of *S. cerevisiae*, was identified and characterized [40]. The At*mak3-1* mutation affected the growth of the plant and the synthesis and accumulation of plastome-encoded photosynthetic gene products, particularly the photosystem II core protein D1. This defect was associated with a diminished level of most of the thylakoid multiprotein complexes and, consequently, with an abnormal photosynthetic electron flow. The reduction in both leaf pigmentation and growth rate can be attributed to the photosynthetic lesion. This phenotype predominated in young At*mak3-1* leaves, indicating that At*MAK3* function is limited in fast-growing tissues.

Most importantly, At*MAK3* can functionally replace the catalytic Mak3p subunit of the NatC complex in yeast and rescue the two known defects associated with the yeast *mak3* mutation, propagation of killer virus particles and growth on non-fermentable carbon sources [40]. Moreover, At*MAK3* could complement not only *mak3*, but also the *mak10* and *mak31* mutations in yeast. Surprisingly, knock-out of the putative plant At*MAK10* gene (the ortholog of the yeast *MAK10*) did not produce a phenotype. However, AtMAK3 and AtMAK10 interact in vivo, as determined

by two-hybrid analysis, although this interaction is not required for N-terminal acetylation of photosynthetic components. Therefore, the formation of a NatC-type complex is not required for AtMAK3 function in *A. thaliana* [39]. Whether or not AtMAK3 acts alone to *N*-acetylate proteins or as part of a different multi-subunit complex remains to be determined. Direct purification of AtMAK3 and associated proteins using biochemical procedures should reveal the nature of NatC-type complex in plants.

The viability of both the yeast *mak3* and plant At*mak3* mutants indicates that NatC substrates are not essential for growth or that the lack of acetylation of the essential proteins is not critical. The fact that a higher plant protein can substitute for the yeast Mak3p indicates the conserved function of *N*-acetylation during evolution, in spite of the fact that AtMAK3 is acting on a set of chloroplast proteins that does not exist in yeast. On the other hand, mitochondrial respiration was not affected in At*mak3* mutants [40]. Thus, NAT substrates diverge during evolution, while the NAT enzymes apparently maintain the same specificities.

Table 2
Amino-acid sequence similarities (%) of the three major yeast *Saccharomyces cerevisiae* NAT subunits: (left) similarities among the catalytic NAT subunits, Ard1p, Mak3p, and Nat3p; (right) similarities among the auxiliary large subunits of the yeast NATs

Catalytic subunits			Auxiliary large subunits		
	Nat3p	Mak3p		Mdm20p	Mak10p
Ard1p	15.4	18.8	Nat1p	11.3	11.1
Nat3p	—	13.1	Mdm20p	—	12.6

Sequence pair distances were made by using Clustal method with PAM250 residue weight table (MegAlign, DNAStar, Inc.).

Similarities of catalytic and auxiliary subunits of NatA, NatB, and NatC

Ard1p, Nat3p, and Mak3p, the catalytic subunits of NatA, NatB, and NatC, respectively, are related to each other by amino-acid sequence (Table 1). These subunits are the representatives of a large acetyltransferase family, the so-called GNAT [41]. In addition, as described above, all NATs have auxiliary subunits that interact with their respective catalytic subunits and that are required for N-terminal acetyltransferase activity. The function of the auxiliary subunits is unknown, but they may play a role in recognition of the proper NAT substrate and in ribosome binding. NatA, NatB, and NatC are required to acetylate different subclasses of proteins and mutants of each N-terminal acetyltransferase have different phenotypes. On the other hand, mutations of the catalytic subunit and the associated subunits for each of the NATs cause similar if not identical phenotypes. The sequences of the catalytic subunits Ard1p, Nat3p, and Mak3p are similar to each other, especially in the region encompassing Ac-CoA binding domain, but more diverged in other regions [6] (Table 2).

The Ard1p, Nat3p, and Mak3p orthologs could be grouped in the separate subfamilies, in which members of each subfamily were predicted to be derived from a common ancestor, as judged by phylogenetic analysis [1]. For example, the phylogenetic tree of the Ard1p orthologs, presented in Fig. 2A, reflects the global evolution of the organisms.

The evolution of the Nat1p, Mdm20p, and Mak10p orthologs from the model organisms also could be evaluated by protein sequence alignment. Phylogenetic analysis of the Nat1p orthologs, presented in Fig. 2B, showed that they evolved in the same direction and coordinately with Ard1p protein. A minor variation in

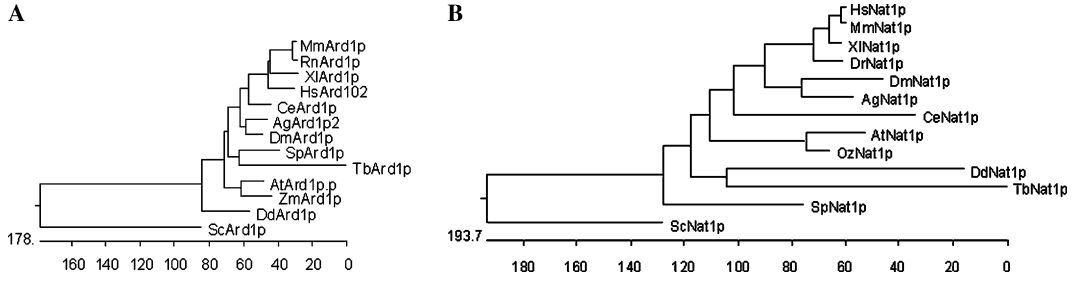

Fig. 2. Phylogenetic analysis of the Ard1p and Nat1p orthologs. Sequence pair distances were made by using Clustal method with PAM250 residue weight table (MegAlign, DNAStar, Inc.). (A) Phylogenetic tree of the Ard1p orthologs. Accession numbers for the Ard1p orthologs are as follows: AgArd1p, *Anopheles gambiae* str. PEST GI:21299226; AtArd1p, *Arabidopsis thaliana* gGI:15450947; CeArd1p, *Caenorhabditis elegans* GI:17541270; DdArd1p, *Dictyostelium discoideum* GI:28850429; DmArd1p, *Drosophila melanogaster* GI:24662081; HsArd1p, *Homo sapiens* GI:17149828; MmArd1p, *Mus musculus* GI:20071186; RtArd1p, *Rattus norvegicus* GI:27682107; ScArd1p, *Saccharomyces cerevisiae* Ard1p; SpArd1p, *Schizosaccharomyces pombe* GI:19115221; TbArd1p, *Trypanosoma brucei* GI:7649677; XlArd1p, *Xenopus laevis* GI: 284223549651963; and ZmArd1p, *Zea mays* GI:14550116. (B) Phylogenetic tree of the Nat1p orthologs. Accession numbers for the Nat1p orthologs are as follows: AgNat1p, *A. gambiae* str. PEST GI:21291634; AtNat1p, *A. thaliana* gGI:20259244; CeNat1p, *C. elegans* GI:25144955; DdNat1p, *D. discoideum* GI:28829826; DmNat1p, *D. melanogaster* GI:24643243; DrNat1p, *Danio rerio* GI:28279630; HsNat1p, *H. sapiens* GI:17149828; MmNat1p, *M. musculus* GI:25048321; OsNat1p, *Oryza sativa* (japonica cultivar-group) GI:13365578; ScNat1p, *S. cerevisiae* Nat1p; SpNat1p, *S. pombe* GI:19075660; TbNat1p, *T. brucei* GI:7649679; and XlNat1p, *X. laevis* GI:9651963.

phylogenetic trees is probably due to inaccuracy of sequencing and ORF identification in different species. For example, we previously corrected the ATG start site of the *NAT3* ORF [28]. Overall, the NAT phylogenetic trees are consistent with the view that yeast and higher eukaryotes have the same N-terminal acetyltransferase systems, although the NATs may be involved in other functions (see above).

The NAT protein sequence analysis also revealed a significant sequence similarity among the auxiliary subunits, Nat1p, Mdm20p, and Mak10p (Table 2), but not Mak31p, which is obviously a member of the Sm-like protein family [39]. The protein similarities between pairwise combinations of Nat1p, Mdm20p, and Mak10p range from 11.1% to 12.6 %, which are comparable to the catalytic subunits Ard1p, Nat3p, and Mak3p similarities that range from 13% to 18.8%. In contrast, no significant similarity is detected between the catalytic subunits and the auxiliary subunits; for example, Ard1p and Nat1p show less than 2% similarity.

Furthermore, all auxiliary subunits, Nat1p, Mdm20p, and Mak10p, contain at least 6–7 TPR motifs non-randomly distributed throughout the protein molecules, as discussed above for Nat1p (Fig. 1). The functional significance of TPRs remains to be investigated, but most likely they are involved in nascent polypeptide binding. Furthermore, there is no detectable Ac-CoA binding site within protein sequences of the auxiliary subunits. The mammalian Nat1p orthologs also lack a significant similarity to an Ac-CoA binding site, making it unlikely that autoacetylation of *tbdn-1* occurs without another protein (see above) [20]. It is also unknown if there are still other auxiliary subunits required for function of NatA and NatB in vivo. In this regard, it would be of interest to determine whether Asc1p and other proteins associating with Ard1p and Nat1p (see above) are required for NatA function. The increased activity of NatA in strains overexpressing just *ARD1* and *NAT1* suggests that Asc1p is not required at least at a stoichiometric level for NatA function in vitro [9].

As noted above, yeast Nat1p interacts with nascent polypeptide chain and also with ribosome and may serve as an anchor of NatA complex to translating ribosomes [10]. We also found that Ard1p-myc fusion protein is ribosome associated as detected by polysome fractionations (B. Polevoda and F. Sherman, unpublished). It will be interesting to determine whether NatB and NatC are attached to yeast ribosomes and whether their attachment to ribosomes is mediated by auxiliary subunits, as in NatA. Based on our extensive studies of acetylation of iso-1-cytochrome *c* proteins with a large variety of altered N-termini, we can expect that no protein with potential acetylation site can escape co-translational N-terminal acetylation [2]. The detailed analysis of the acetylated proteins with similar N-terminal sequences but with different cellular compartment localization also supports such a suggestion [1]. However, proteins translated in mitochondria and chloroplasts may not be or may be only rarely N-terminally acetylated, similar to prokaryotes.

Although they have pronounced phenotypes, it is somewhat surprising that NatA, NatB, and NatC deletants are viable in yeast. Approximately one-half of yeast proteins are N-terminally acetylated, with most being NatA substrates [1,5]. We can consider that the lack of N-terminal acetylation diminishes activity or stability to various degrees, from no detectable level to the complete loss of function. Clearly, the complete loss of function of essential proteins does not occur by the lack of N-terminal acetylation in yeast. For example, the lack of N-terminal acetylation of the essential proteins, actin and tropomyosin, leads to only a partial loss of activity [28]. On the other hand, the lack of acetylation of the viral major coat protein, *gag*, leads to the loss of maintenance of the dsRNA virus, and this deficiency can be considered as complete [32]. In fact, the common use of N-terminal fusion of peptides and proteins for convenient identification and purification procedures indicates that the state of the N-terminus is by-and-large innocuous for most proteins. But it remains to be determined whether this is true for higher eukaryotes and whether the certain NATs have diverged from the ones in yeast by acting on new types of substrate or incorporating new subunit.

Acknowledgments

This work has been supported by the National Institute of Health Grant R01 GM12702. We are grateful to Dr. S. Rospert (Max-Plank Research Unit "Enzymology of Protein Folding"), Dr. B. Williams (Cornell University), Dr. R. Wickner (Laboratory of Biochemistry and Genetics, NIDDKD), and Dr. D. Leister (Max-Plank Institut für Züchtungsforschung) for communication of their results prior to publication.

References

[1] B. Polevoda, F. Sherman, N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins, J. Mol. Biol. 325 (2003) 595–622.
[2] B. Polevoda, F. Sherman, N^{α}-terminal acetylation of eukaryotic proteins, J. Biol. Chem. 275 (2000) 36479–39482.
[3] H.P.C. Driessen, W.W. de Jong, G.L. Tesser, H. Bloemendal, The mechanism of N-terminal acetylation of proteins, CRC Crit. Rev. Biochem. 18 (1985) 281–325.
[4] B. Persson, C. Flinta, G. von Heijne, H. Jörnvall, Structures of N-terminally acetylated protein, Eur. J. Biochem. 152 (1985) 523–527.
[5] F.J. Lee, L.W. Lin, J.A. Smith, N^{α}-acetyltransferase deficiency alters protein synthesis in *Saccharomyces cerevisiae*, FEBS Lett. 256 (1989) 139–142.
[6] B. Polevoda, J. Norbeck, H. Takakura, A. Blomberg, F. Sherman, Identification and specificities of N-terminal acetyltransferases from *Saccharomyces cerevisiae*, EMBO J. 18 (1999) 6155–6168.

[7] R.J. Arnold, B. Polevoda, J.P. Reilly, F. Sherman, The action of N-terminal acetyltransferases on yeast ribosomal proteins, J. Biol. Chem. 274 (1999) 37035–37040.
[8] J.R. Mullen, P.S. Kayne, R.P. Moerschell, S. Tsunasawa, M. Gribskov, M. Colavito-Shepanski, M. Grunstein, F. Sherman, R. Sternglanz, Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast, EMBO J. 8 (1989) 2067–2075.
[9] E.-C. Park, J.W. Szostak, *ARD1* and *NAT1* proteins form a complex that has N-terminal acetyltransferase activity, EMBO J. 11 (1992) 2087–2093.
[10] M. Gautschi, S. Just, A. Mun, S. Ross, P. Rucknagel, Y. Dubaquie, A. Ehrenhofer-Murray, S. Rospert, The yeast N^{α}-acetyltransferase NatA is quantitatively anchored to the ribosome and interacts with nascent polypeptides, Mol. Cell. Biol. 23 (2003) (in press).
[11] T. Andoh, Y. Hirata, A. Kikuchi, Yeast glycogen synthase kinase 3 is involved in protein degradation in cooperation with Bul1, Bul2, and Rsp5, Mol. Cell. Biol. 20 (2000) 6712–6720.
[12] M.S. Longtine, S. Enomoto, S.L. Finstad, J. Berman, Telomere-mediated plasmid segregation in *Saccharomyces cerevisiae* involves gene products required for transcriptional repression at silencers and telomeres, Genetics 133 (1993) 171–182.
[13] E.M. Stone, M.J. Swanson, A.M. Romeo, J.B. Hicks, R. Sternglanz, The *SIR1* gene of *Saccharomyces cerevisiae* and its role as an extragenic suppressor of several mating-defective mutants, Mol. Cell. Biol. 11 (1991) 2253–2262.
[14] C. Tribioli, M. Mancini, E. Plassart, S. Bione, S. Rivella, C. Sala, G. Torri, D. Toniolo, Isolation of new genes in distal Xq28: transcriptional map and identification of a human homologue of the ARD1 *N*-acetyltransferase of *Saccharomyces cerevisiae*, Hum. Mol. Genet. 3 (1994) 1061–1067.
[15] M. Whiteway, R. Freedman, S. Van Arsdell, J.W. Szostak, J. Thorner, The yeast ARD1 gene product is required for repression of cryptic mating-type information at the HML locus, Mol. Cell. Biol. 7 (1987) 3713–3722.
[16] I.I. Ouspenski, S.J. Elledge, B.R. Brinkley, New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability, Nucleic Acids Res. 27 (1999) 3001–3008.
[17] A.K. Ingram, G.A. Cross, D. Horn, Genetic manipulation indicates that ARD1 is an essential N^{α}-acetyltransferase in *Trypanosoma brucei*, Mol. Biochem. Parasitol. 111 (2000) 309–317.
[18] S.C. Choi, J.Y. Chang, J.K. Han, A novel *Xenopus* acetyltransferase with a dynamic expression in early development, Biochem. Biophys. Res. Commun. 285 (2001) 1338–1343.
[19] N. Sugiura, R.G. Patel, R.A. Corriveau, *N*-methyl-D-aspartate receptors regulate a group of transiently expressed genes in the developing brain, J. Biol. Chem. 276 (2001) 14257–14263.
[20] R.L. Gendron, L.C. Adams, H. Paradis, Tubedown-1, a novel acetyltransferase associated with blood vessel development, Dev. Dyn. 218 (2000) 300–315.
[21] O. Fluge, O. Bruland, L.A. Akslen, J.E. Varhaug, J.R. Lillehaug, NATH, a novel gene overexpressed in papillary thyroid carcinomas, Oncogene 21 (2002) 5056–5068.
[22] G.L. Blatch, M. Lassle, The tetratricopeptide repeat: a structural motif mediating protein–protein interactions, Bioessays 21 (1999) 932–939.
[23] M.A. Andrade, C.P. Ponting, T.J. Gibson, P. Bork, Homology-based method for identification of protein repeats using statistical significance estimates, J. Mol. Biol. 298 (2000) 521–537.
[24] B. Polevoda, F. Sherman, The diversity of acetylated proteins, Genome Biol. 3 (2002) 0001–0006, reviews.
[25] J.W. Jeong, M.K. Bae, M.Y. Ahn, S.H. Kim, T.K. Sohn, M.H. Bae, M.A. Yoo, E.J. Song, K.J. Lee, K.W. Kim, Regulation and destabilization of HIF-1α by ARD1-mediated acetylation, Cell 111 (2002) 709–720.
[26] D.M. Willis, A.P. Loewy, N. Charlton-Kachigian, J.S. Shao, D.M. Ornitz, D.A. Towler, Regulation of osteocalcin gene expression by a novel Ku antigen transcription factor complex, J. Biol. Chem. 277 (2002) 37280–37291.
[27] D. Ivanov, A. Schleiffer, F. Eisenhaber, K. Mechtler, C.H. Haering, K. Nasmyth, Eco1 is a novel acetyltransferase that can acetylate proteins involved in cohesion, Curr. Biol. 19 (2002) 323–328.
[28] B. Polevoda, T.S. Cardillo, T.C. Doyle, G.S. Bedi, F. Sherman, Nat3p and Mdm20p are required for function of yeast NatB N^{α}-terminal acetyltransferase and of actin and tropomyosin, J. Biol. Chem. 278 (2003) (in press).
[29] G.J. Hermann, E.J. King, J.M. Shaw, The yeast gene, *MDM20*, is necessary for mitochondrial inheritance and organization of the actin cytoskeleton, J. Cell. Biol. 137 (1997) 141–153.
[30] J.M. Singer, J.M. Shaw, Mdm20 protein functions with Nat3 protein to acetylate Tpm1 protein and regulate tropomyosin–actin interactions in budding yeast, Proc. Natl. Acad. Sci. 100 (2003) 7644–7649.
[31] Y. Kimura, M. Takaoka, S. Tanaka, H. Sassa, K. Tanaka, B. Polevoda, F. Sherman, H. Hirano, N^{α}acetylation and proteolytic activity of the yeast 20 S proteasome, J. Biol. Chem. 275 (2000) 4635–4639.
[32] J.C. Tercero, R.B. Wickner, MAK3 encodes an *N*-acetyltransferase whose modification of the L-A gag NH_2 terminus is necessary for virus particle assembly, J. Biol. Chem. 267 (1992) 20277–20281.
[33] J.C. Tercero, J.D. Dinman, R.B. Wickner, Yeast MAK3 *N*-acetyltransferase recognizes the N-terminal four amino acids of the major coat protein (gag) of the L-A double-stranded RNA virus, J. Bacteriol. 175 (1993) 3192–3194.
[34] B. Polevoda, F. Sherman, NatC N^{α}-terminal acetyltransferase of yeast contains three subunits, Mak3p, Mak10p, and Mak31p, J. Biol. Chem. 276 (2001) 20154–20159.
[35] Y.J. Lee, R.B. Wickner, MAK10, a glucose-repressible gene necessary for replication of a dsRNA virus of *Saccharomyces cerevisiae*, has T cell receptor α-subunit motifs, Genetics 132 (1992) 87–96.
[36] X.-J. Yi, X. Li, F.-S. Yu, A novel epithelial wound-related gene is abundantly expressed in developing rat cornea and skin, Curr. Eye Res. 20 (2000) 430–440.
[37] M.C. Weiser-Evans, P.E. Schwartz, N.A. Grieshaber, B.E. Quinn, S.S. Grieshaber, J.K. Belknap, P.M. Mourani, R.A. Majack, K.R. Stenmark, Novel embryonic genes are preferentially expressed by autonomously replicating rat embryonic and neointimal smooth muscle cells, Circ. Res. 87 (2000) 608–615.
[38] T. Achsel, H. Brahms, B. Kastner, A. Bachi, M. Wilm, R. Luhrmann, A doughnut-shaped heteromer of human Sm-like proteins binds to the 3′-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro, EMBO J. 18 (1999) 5789–5802.
[39] B. Séraphin, Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs, EMBO J. 14 (1995) 2089–2098.
[40] P. Pesaresi, N. Haigh, S. Masiero, A. Dietzmann, L. Eichacker, R. Wickner, F. Salamini, D. Leister, Cytoplasmic N-terminal protein acetylation is required for efficient photosynthesis in Arabidopsis, Plant Cell 15 (2003) (in press).
[41] A.F. Neuwald, D. Landsman, GCN5-related histone *N*-acetyltransferase belongs to a diverse superfamily that includes the yeast SPT10, Trends Biochem. Sci. 22 (1997) 154–155.