The EMBO Journal Vol.18 No.21 pp.6155–6168, 1999

Identification and specificities of N-terminal acetyltransferases from *Saccharomyces cerevisiae*

Bogdan Polevoda, Joakim Norbeck[1], Hikaru Takakura[2], Anders Blomberg[1] and Fred Sherman[3]

Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA, [1]Department of Cell and Molecular Biology and Microbiology, Lundberg Laboratory, Göteborg University, Medicinaregatan 9C, S-413 90 Göteborg, Sweden and [2]Takara Shuzo Co. Ltd, Biotechnology Research Laboratory, Seta 3-4-1, Otsu, Shiga, 520-21, Japan

[3]Corresponding author
e-mail: Fred_Sherman@urmc.rochester.edu

N-terminal acetylation can occur cotranslationally on the initiator methionine residue or on the penultimate residue if the methionine is cleaved. We investigated the three N-terminal acetyltransferases (NATs), Ard1p/Nat1p, Nat3p and Mak3p. Ard1p and Mak3p are significantly related to each other by amino acid sequence, as is Nat3p, which was uncovered in this study using programming alignment procedures. Mutants deleted in any one of these NAT genes were viable, but some exhibited diminished mating efficiency and reduced growth at 37°C, and on glycerol and NaCl-containing media. The three NATs had the following substrate specificities as determined *in vivo* by examining acetylation of 14 altered forms of iso-1-cytochrome *c* and 55 abundant normal proteins in each of the deleted strains: Ard1p/Nat1p, subclasses with Ser-, Ala-, Gly- and Thr-termini; Nat3p, Met-Glu- and Met-Asp- and a subclass of Met-Asn-termini; and Mak3p subclasses with Met-Ile- and Met-Leu-termini. In addition, a special subclass of substrates with Ser-Glu-Phe-, Ala-Glu-Phe- and Gly-Glu-Phe-termini required all three NATs for acetylation.

Keywords: acetylation/N-terminal acetyltransferase/protein modification/*Saccharomyces cerevisiae*

Introduction

The two cotranslational processes, cleavage of N-terminal methionine residues and N-terminal acetylation, are by far the most common modifications, occurring on the vast majority of proteins. Proteins from prokaryotes, mitochondria and chloroplasts initiate with formyl-methionine, whereas proteins from the cytosol of eukaryotes initiate with methionine. The formyl group is removed from prokaryotic proteins by a deformylase, resulting in methionine at the N-termini. The methionine at the N-termini is cleaved from nascent chains of most prokaryotic and eukaryotic proteins. N-terminal acetylation occurs subsequently on certain of the proteins, either containing or lacking the methionine residue. This N-terminal acetylation occurs on more than one-half of eukaryotic proteins, but seldom on prokaryotic proteins (Driessen *et al*., 1985; Kendall *et al*., 1990).

Because the N-terminal region of yeast iso-1-cytochrome *c* (iso-1) is dispensable for biosynthesis, function and mitochondrial import (Baim *et al*., 1985; Wang *et al*., 1996), and because N-terminal processing is completed before mitochondrial import of apo-cytochrome *c*, N-terminal processing can be investigated freely with essentially any alteration. In fact, altered forms of iso-1 proved to be ideally suited to investigating the specificity of N-terminal methionine cleavage and N-terminal acetylation. Results on altered iso-1 (Tsunasawa *et al*., 1985) were the basis for the hypothesis that methionine is cleaved from penultimate residues having radii of gyration of 1.29 Å or less (glycine, alanine, serine, cysteine, threonine, proline and valine; Sherman *et al*., 1985), a hypothesis that was confirmed from the results of a complete set of altered iso-1 having all possible amino acids at the penultimate position (Moerschell *et al*., 1990).

Methionine excision occurs before completion of the nascent chain and before other N-terminal processing events, such as N-terminal acetylation (Driessen *et al*., 1985; Kendal *et al*., 1990). Eukaryotes, including *Saccharomyces cerevisiae*, contain two types of methionine aminopeptidases, Map1p and Map2p (Arfin *et al*., 1995; Li and Chang, 1995). Yeast mutants containing either *map1* or *map2* null mutations are viable, although *map2* mutants grow more slowly, and the *map1map2* double mutants are nonviable (Li and Chang, 1995). Thus, removal of N-terminal methionine is an essential function in yeast, as in prokaryotes, but the process can be carried out by either of the enzymes.

Although all of the eukaryotic methionine aminopeptidases have presumably been identified and their specificities defined, our understanding of N-terminal acetyltransferases (NATs) is far more limited. N-terminal acetylation of proteins is catalyzed by NATs that transfer acetyl groups from acetyl-CoA to the termini of α-amino groups. Similar to N-terminal methionine cleavage, N-terminal acetylation is one of the most common protein modifications in eukaryotes, occurring on ~85% of the different varieties of proteins. *In vitro* studies indicate that N-terminal acetylation occurs cotranslationally when there are 20–50 residues extruding from the ribosome (Driessen *et al*., 1985; Kendall *et al*., 1990). Acetylation occurs most often on penultimate residues after methionine cleavage, but it also occurs on initiator methionine residues. N-terminal acetylation also can occur at internal sites after specific proteolytic processing of the completely translated protein, as in the case of peptide hormones (Dores *et al*., 1990). Post-translational acetylation of each of the different proteins occurs with different NATs having different specificities, and these differ from the set of NATs carrying out cotranslational acetylation.

Eukaryotic proteins susceptible to N-terminal acetylation have a variety of different N-terminal sequences, with no simple consensus motifs and no dependence on a single type of residue. Eukaryotic proteins with serine and alanine termini are the most frequently acetylated and these residues, along with methionine, glycine and threonine account for >95% of the N-terminal acetylated residues (Driessen *et al*., 1985; Persson *et al*., 1985; Kendall *et al*., 1990; Bradshaw *et al*., 1998). However, not all proteins with any of these N-terminal residues are acetylated. In this study, we demonstrate that the complexity of the termini that are acetylated is due to the existence of multiple NATs, each acting on different groups of amino acid sequences whose specificity is determined by two or more residues at the terminal positions.

Mullen *et al*. (1989) established that the *NAT1* and *ARD1* genes encode a NAT in the yeast *S.cerevisiae*. A *nat1*$^-$ mutant was originally uncovered by screening a collection of heavily mutagenized strains for protein acetyltransferase activity *in vitro*. The previously identified *ard1*$^-$ mutant was first suspected to be related to *nat1*$^-$ because of certain similar phenotypes. In addition to lacking NAT activity, both *nat1*$^-$ and *ard1*$^-$ mutants exhibited slower growth, derepression of the silent mating type gene *HML*α and failure to enter G_0. Also, a 20-fold increase in NAT activity occurred when both *NAT1* and *ARD1* were concomitantly overexpressed, but not when either gene was overexpressed individually. Furthermore, Park and Szostak (1992) used epitope-tagged derivatives of Nat1p and Ard1p to provide evidence that both proteins interact and form a complex. *NAT1* and *ARD1* were also investigated by Lee *et al*. (1989a,b).

Most importantly, as confirmed and extended in this study, *nat1*$^-$ and *ard1*$^-$ mutants were not able to carry out N-terminal acetylation *in vivo* of the same subset of normally acetylated proteins, including those with Ser-termini (Mullen *et al*., 1989; Sherman *et al*., 1993). However, altered iso-1 with Met-Glu-, Met-Asp- and Met-Asn-Asn-termini were still acetylated in *nat1*- mutants, indicating that other NATs were responsible for acetylating proteins with N-terminal methionine residues.

In this regard, Tercero and Wickner (1992) and Tercero *et al*. (1992, 1993) described the *MAK3* gene that encodes a NAT which is required for the N-terminal acetylation of the viral major coat protein, *gag*, with an Ac-Met-Leu-Arg-Phe-terminus. *Mak3*$^-$ deficient mutants do not assemble or maintain the viral particle, and have reduced growth on media containing nonfermentable carbon as the sole energy source. Tercero *et al*. (1993) suggested that the diminished growth of *mak3*-Δ strains on glycerol medium was due to the lack of N-terminal acetylation of the mitochondrial proteins α-ketoglutarate dehydrogenase (Kdg1p), fumarate dehydratase (Fum1p) and a mitochondrial ribosomal protein (Mrp1p), which all contained Met-Leu-Arg-Phe-termini, similar to the L-A *gag* protein. They further suggested that acetylation may be part of the import signal sequence, which may be cleaved after entry into the mitochondrion. Such processed proteins derived from *mak3*-Δ strains would not have altered mobilities on two-dimensional gels. Tercero *et al*. (1993) investigated the specificity of Map3p acetylation by testing variants having various single amino acid replacements of the Met-Leu-Arg-Phe-Val-terminus. In this study, we confirmed the properties of *mak3*-Δ mutants and investigated the requirement of Mak3p for acetylating the other N-terminal sequences. Our results demonstrate that acetylation of the more frequently observed methionine termini, such as Met-Glu- or Met-Asn-Asn-, does not depend on *MAK3*.

In addition, by systematically screening a series of heavily mutagenized strains for peptide acetyltransferase activity *in vitro*, Kulkarni and Sherman (1994) uncovered *nat2-1*, a conditional mutant whose extracts were unable to acetylate peptides with methionine termini. However, in this study we were unable to demonstrate that the *nat2-1* temperature-sensitive mutation prevents acetylation *in vivo* using the iso-1 system.

Lee *et al*. (1997) reported the isolation and purification of a protein having NAT activity *in vitro* with certain proteins and peptides containing N-terminal methionine residues, including those having Met-Glu-termini. Analysis of HPLC-purified peptides obtained from trypsin digestion of the putative NAT resulted in the identification of three peptide sequences that corresponded to a protein, YLR410W, which is now denoted Vip1p and is similar to the *Schizosaccharomyces pombe* protein Asp1p (Feoktistova *et al*., 1999). In this study, we demonstrated that the *vip1*-Δ deletion does not prevent acetylation *in vivo* of proteins with N-terminal methionine residues, including those proteins having the proper Met-Glu- N-terminal sequences.

Most importantly, we used standard dynamic programming alignment methods to identify the NAT required for acetylating the more frequently occurring proteins with N-terminal methionine residues. The amino acid sequence of this NAT, denoted Nat3p, was highly similar to that of Ard1p and Mak3p. Furthermore, *nat3*-Δ mutants were not able to acetylate proteins with Met-Glu- and related termini *in vivo*.

Thus, in this study we established that *S.cerevisiae* contains three NATs, Ard1p/Nat1p, Nat3p and Mak3p, with each required for acetylating different groups of proteins. We also suggest that these three NATs account for all cotranslational N-terminal acetylation.

Results

Identification of Nat3p by sequence similarity

Neuwald and Landsman (1997) previously used amino acid alignment procedures to assign Ard1p and Mak3p to the GNAT (GCN5-related *N*-acetyltransferase) family of *N*-acetyltransferases, which includes over 50 members, some of which are N^{α}- or N^{ε}-acetyltransferases. The GNAT superfamily of *N*-acetyltransferases is characterized by four conserved regions, motifs A–D, which span >100 residues. Neuwald and Landsman (1997) suggested that these conserved motifs would be involved in binding CoA because the ability to bind and transfer CoA is the only property that these diverse *N*-acetyltransferases have in common. Most importantly, the structures of the histone *N*-acetyltransferase, Hat1p (Dutnall *et al*., 1998), aminoglycosidase 3-*N*-acetyltransferase (Wolf *et al*., 1998) and serotonin *N*-acetyltransferase (Hickman *et al*., 1999) confirmed that motif A is involved in CoA binding. Furthermore, motif A is the longest and most highly conserved, having a short amino acid sequence Q/RxxGxG/A that is

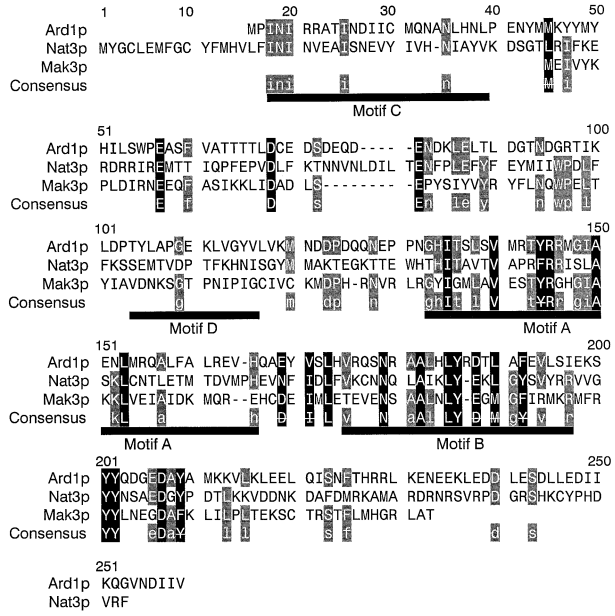

Fig. 1. Amino acid sequence alignment of Ard1p, Nat3p and Mak3p. Highly conserved residues are highlighted in black, where M = M or L, Y = Y or F, I = I or V and D = D or E. Moderately conserved residues are highlighted in gray. Motifs A–D, which correspond to the motifs described by Neuwald and Landsman (1997) and Dutnall *et al*. (1998), belong to the superfamily of various *N*-acetyltransferases. Motifs A and B are involved in the binding of acetyl CoA.

Table I. Genes investigated for possibly encoding *N*-terminal acetyltransferase activity *in vivo*

Gene	ORF	Reference
NAT1	YDL040C	Mullen *et al*. (1989)
ARD1	YHR013C	Mullen *et al*. (1989)
MAK3	YPR051W	Tercero *et al*. (1993)
NAT3	YPR131C	this study
NAT2	YGR147C	Kulkarni and Sherman (1994)
VIP1	YLR410W	Lee *et al*. (1997)

of particular interest. Point mutations in any of these residues lead to >90% reduced activity of Gcn5p, Hat1p, Mak3p, Esa1p, MOF and SSAT enzymes (Tercero and Wickner, 1992; Coleman *et al*., 1996; Hilfiker *et al*., 1997; Dutnall *et al*., 1998; Kuo *et al*., 1998; Smith *et al*., 1998).

We used the complete DNA sequence of *S.cerevisiae* to compare all deduced open reading frames (ORFs) with known or putative NATs (Table I). YPR131C, now denoted Nat3p, was found to be significantly similar to both Ard1p and Mak3p, using BLAST (Altschul *et al*., 1990), BESTFIT (Smith and Waterman, 1981) in different variations, and finally MULTIALIGN (Corpet, 1988) computer programs (Figure 1). Although the overall similarity between Nat3p, Ard1p and Mak3p is not extensive, the regions spanning motifs A and B have remarkably high degrees of similarity. Furthermore, the Nat3p sequence, RRISLA (amino acid residues 144–149), closely resembles the highly conserved sequence Q/RxxGxG/A. The corresponding regions of Ard1p and Mak3p, RRMGIA and RGHGIA, respectively, perfectly satisfy conserved sequence requirements. The presence of several hydrophobic amino acid residues at positions 134, 137, 139 and 143 also contributes to the formation of an active acetyl CoA pocket and is important

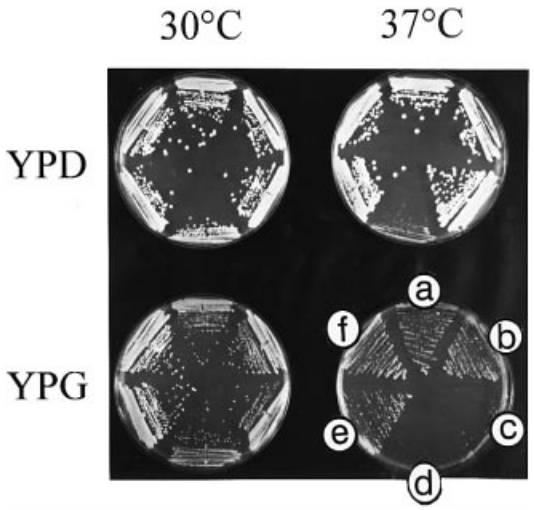

Fig. 2. The growth of the following isogenic series of strains at 30°C and 37°C on YPD and YPG media for 3 and 5 days, respectively: a, normal (B-11679); b, *vip1*-Δ (B-11707); c, *mak3*-Δ (B-11705); d, *nat3*-Δ (B-11852); e, *ard1*-Δ (B-11706); and f, *nat1*-Δ (B-11815).

for function, as demonstrated with Mak3p (Tercero *et al*., 1992), Hat1p and Gcn5p (Dutnall *et al*., 1998; Lin *et al*., 1999).

The degree of similarity between Ard1p, Mak3p and Nat3p in motif B is also high (Figure 1), including the conserved Asn179, Tyr186, Glu188 and Tyr191, and several surrounding hydrophobic amino acids, which help to bind acetyl CoA, as shown in the structures of Hat1p and Gcn5p. The similarity in two other motifs, C and D, is not as readily observed, except for two asparagine and three isoleucine residues in motif C, which are absent in Mak3p. It should be noted that Ard1p, Mak3p and Nat3p form a subfamily in the GNAT protein family having greater similarity among themselves than to the other members of GNAT, including histone acetyltransferases, the proteins involved in silencing or transcription co-activation and prokaryotic antibiotic acetyltransferases.

The other proteins considered in this investigation, Nat1p, Nat2p and Vip1p, do not share significant similarity to any GNAT family member, except Nat2p, which possibly has an acetyl CoA-binding site and a sequence resembling motif A, but the overall similarity to other motifs as well as to the Ard1p/Mak3p/Nat3p subfamily is low.

Phenotypes of known and putative NAT deletion mutants

Mullen *et al*. (1989) reported that *nat1*-Δ and *ard1*-Δ deletion mutants are viable but exhibit reduced growth and diminished mating efficiency compared with the *MAT***a** cells. Also Tercero and Wickner (1992) reported that *mak3*-Δ mutants have diminished growth on YPG medium. These phenotypes were confirmed and others were investigated in this study, along with the phenotypes of the *nat3*-Δ and *vip1*-Δ mutants, using the following isogenic series of strains that were also prepared to examine total proteins on two-dimensional gels (Table II): normal (B-11679); *nat1*-Δ (B-11815); *ard1*-Δ (B-11706); *mak3*-Δ (B-11705); *nat3*-Δ (B-11852); and *vip1*-Δ (B-11707). As shown in Figure 2, and summarized in Table III, *nat1*-Δ and *ard1*-Δ mutants exhibited diminished growth

Table II. Yeast strains

Strain no.	Genotype	
B-7528		*cyc1-31* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8236		*CYC1-1070* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-7471		*CYC1-795* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8262		*CYC1-1093* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8261		*CYC1-1092* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9423		*CYC1-1371* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9563	*nat1*-Δ::*URA3*	*CYC1-1371* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9564	*mak3*-Δ::*URA3*	*CYC1-1371* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8237		*CYC1-1071* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10645		*CYC1-987* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-7658	*nat1*-Δ::*URA3*	*CYC1-987* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10672	*mak3*-Δ::*URA3*	*CYC1-987* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11973	*nat3*-Δ::*kanMX2*	*CYC1-987* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10551	*vip1*-Δ::*URA3*	*CYC1-987* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8032		*CYC1-963* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8360	*nat1*-Δ::*URA3*	*CYC1-963* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9074	*mak3*-Δ::*URA3*	*CYC1-963* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11974	*nat3*-Δ::*kanMX2*	*CYC1-963* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10563	*vip1*-Δ::*URA3*	*CYC1-963* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8031		*CYC1-962* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8361	*nat1*-Δ::*URA3*	*CYC1-962* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9073	*mak3*-Δ::*URA3*	*CYC1-962* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11975	*nat3*-Δ::*kanMX2*	*CYC1-962* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10548	*vip1*-Δ::*URA3*	*CYC1-962* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-7687		*CYC1-853* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-7727	*nat1*-Δ::*URA3*	*CYC1-853* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9072	*mak3*-Δ::*URA3*	*CYC1-853* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11863	*nat3*-Δ::*kanMX2*	*CYC1-853* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10580	*vip1*-Δ::*URA3*	*CYC1-853* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-7723		*CYC1-872* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-7846	*nat1*-Δ::*URA3*	*CYC1-872* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9071	*mak3*-Δ::*URA3*	*CYC1-872* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11864	*nat3*-Δ::*kanMX2*	*CYC1-872* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10589	*vip1*-Δ::*URA3*	*CYC1-872* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8462		*CYC1-1162* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9024	*nat1*-Δ::*URA3*	*CYC1-1162* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9022	*mak3*-Δ::*URA3*	*CYC1-1162* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11865	*nat3*-Δ::*kanMX2*	*CYC1-1162* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10564	*vip1*-Δ::*URA3*	*CYC1-1162* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8566		*CYC1-1201* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-8565	*nat1*-Δ::*URA3*	*CYC1-1201* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-9076	*mak3*-Δ::*URA3*	*CYC1-1201* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-12040	*nat3*-Δ::*kanMX2*	*CYC1-1201* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10646		*CYC1-1286* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10673	*nat1*-Δ::*URA3*	*CYC1-1286* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-10688	*mak3*-Δ::*URA3*	*CYC1-1286* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-12041	*nat3*-Δ::*kanMX2*	*CYC1-1286* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11404	*vip1*-Δ::*URA3*	*CYC1-1286* *MAT***a** *cyc7-67 ura3-52 lys5-10*
B-11679		*MAT*α *ura3-52* (YFJN001-1A)
B-11706	*ard1*-Δ::*kanMX2*	*MAT*α *ura3-52*
B-11815	*nat1*-Δ::*kanMX2*	*MAT*α *ura3-52*
B-11852	*nat3*-Δ::*kanMX2*	*MAT*α *ura3-52*
B-11705	*mak3*-Δ::*kanMX2*	*MAT*α *ura3-52*
B-11707	*vip1*-Δ::*kanMX2*	*MAT*α *ura3-52*
B-10172	*nat2-1 CYC1-872*	*MAT*α *cyc7-67 trp2*
B-10173	*nat2-1 CYC1-1093*	*MAT*α *cyc7-67 trp2*
B-6924	*MAT***a** *cyc1-115 cyc7-67 trp2 can1 lys2*	
B-6925	*MAT*α *cyc1-115 cyc7-67 trp2 can1 lys2*	

on YPG medium at 37°C, although growth on YPD medium at 30°C was nearly normal. The mating frequency of *MAT***a** *nat1*-Δ and *MAT***a** *ard1*-Δ cells was reduced by nearly two orders of magnitude (Table III). The *mak3*-Δ mutants exhibited an almost complete lack of growth on YPG medium at 37°C (Figure 2; Table III), more or less confirming the results of Tercero and Wickner (1992). Disruption experiments revealed that *NAT3* was not an essential gene, although the growth of *nat3*-Δ strains was severally retarded on both YPD and YPG media at both temperatures, but particularly on YPG medium at 37°C (Figure 2; Table III). Furthermore, the mating efficiency of *MAT*α *nat3*-Δ cells, but not *MAT***a** *nat3*-Δ cells, was reduced by almost three orders of magnitude (Table III). Also, examination of the growth of the strains on YPD medium containing 1 M NaCl revealed a marked diminution of growth of the *nat3*-Δ strain (Figure 3; Table III). In contrast, the *vip1*-Δ mutant did not exhibit any growth or mating defect (Figures 2 and 3; Table III).

Choice of mutationally altered iso-1

We attempted to create a series of altered iso-1 representing the most common types of N-terminal acetylated proteins having Met-, Ser-, Ala-, Gly- or Thr-terminal residues. In normal yeast strains, the N-terminal methionine of iso-1 is cleaved and the newly exposed threonine residue is not acetylated. However, during the course of several studies, many mutant forms of iso-1 were found to be acetylated (Tsunasawa *et al.*, 1985). Because of the dispensability of the N-terminal region and the ease of generating altered sequences by transformation with synthetic oligonucleo-

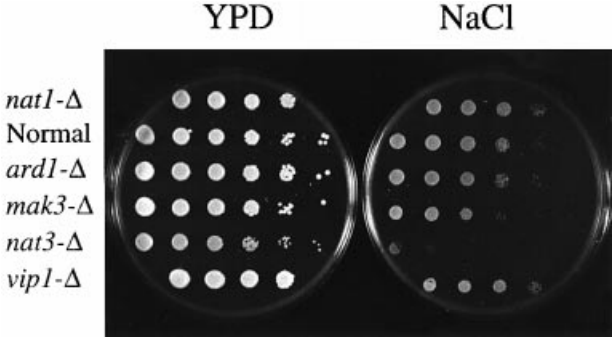

Fig. 3. The growth of serial one-tenth dilutions of the following isogenic series of strains at 30°C on YPD medium and NaCl medium (YPD + 1 M NaCl) for 3 days: *nat1*-Δ (B-11815); normal (B-11679); *ard1*-Δ (B-11706); *mak3*-Δ (B-11705); *nat3*-Δ (B-11852); and *vip1*-Δ (B-11707).

Table III. Mating efficiency and growth of the normal B-11679 strain and the isogenic mutants

	Approximate mating efficiency		Growth				
			YPD		YPG		NaCl
	× *MAT***a**	× *MAT*α	30°C	37°C	30°C	37°C	30°C
Normal	1	1	+++	+++	+++	++	+++
nat1-Δ	0.04	1	+++	+++	+++	++	++
ard1-Δ	0.04	1	+++	+++	+++	++	++
nat3-Δ	0.4	2×10^{-3}	++	+	++	0	±
mak3-Δ	1	1	+++	+++	+++	0	++
vip1-Δ	1	1	+++	+++	+++	+++	+++

NaCl denotes YPD containing 1.0 M NaCl.

tides, the iso-1 system has been used to systematically investigate N-terminal processing (Moerschell *et al*., 1990; Sherman *et al*., 1993). In the study of Moerschell *et al*. (1990), which included all 20 replacements in the Met-*Any*-Phe-Leu-Ala Gly- series, only Met-Glu- and Met-Asp- were completely acetylated, whereas Met-Asn- was partially acetylated. The types of acetylated iso-1 were augmented in this study by generating another series, Met-*Any*-Glu-Phe-Leu-Ala Gly-, whose acetylated derivatives are presented in Table IV. The difference between the two series further emphasizes the complex sequence requirement for acetylation.

In this study, we also included altered iso-1 having sequences corresponding to the N-terminal sequences of proteins presumably acetylated in yeast. Cook *et al*. (1991) reported that the N-terminal methionine was cleaved and the penultimate cysteine residue partially acetylated in the mutant form of actin having a Met-Cys-Asp-Ser-Glu-Val-terminus. However, the *CYC1-1071* iso-1, having the same five amino acid N-terminus, was not acetylated (Table IV). Also, the *CYC1-1371* iso-1 was not acetylated (Table IV), even though it had the same five N-terminal amino acids, Thr-Glu-Phe-Tyr-Ser-, as Erg7p which presumably is acetylated (Corey *et al*., 1994; E.J.Corey, unpublished results). In contrast, the two iso-1, *CYC1-1201 (*Ac-Met-Leu-Arg-Phe-Arg-) and *CYC1-1286 (*Ac-Met-Leu-Ala-Phe-Ala-), having similar N-terminal sequences to the L-A *gag* protein (Ac-Met-Leu-Arg-Phe-Val-) (Tercero *et al*., 1993) were both acetylated (Table IV).

The 14 altered iso-1 listed in Table IV were all functional, although the amount varied between 25 and 100% of the normal level. The variation in amount probably reflects amphipathic structures in the N-terminal region which influences the levels of iso-1 (R.P.Moerschell and F.Sherman, unpublished results).

Lack of acetylation of altered iso-1 in deletion mutants

The eight acetylated iso-1, listed at the bottom of Table IV, were used to investigate the requirement for acetylation by the NATs or presumptive NATs. The *NAT1*, *MAK3* and *NAT3* genes were disrupted in each of the eight normal strains and the percent acetylation of all of the iso-1 was determined from the HPLC analysis of peptides generated

Table IV. Percentage N-terminal acetylation of altered iso-1 from normal, *nat1*-Δ, *mak3*-Δ and *nat3*-Δ strains

		% Iso-1	% acetylation				
			Normal	*nat1*-Δ	*mak3*-Δ	*nat3*-Δ	
CYC1-1070	(Met)Cys - Glu - Phe - Leu - Ala - ATA ATG TGT GAA TTC TTG GCC	50	0 B-8236	–	–	–	
CYC1-795	(Met)Thr - Glu - Phe - Leu - Ala - ATA ATG ACT GAA TTC TTG GCC	100	0 B-7471	–	–	–	
CYC1-1093	(Met)Pro - Glu - Phe - Leu - Ala - ATA ATG CCT GAA TTC TTG GCC	80	0 B-8262	–	–	–	
CYC1-1092	(Met)Val - Glu - Phe - Leu - Ala - ATA ATG GTT GAA TTC TTG GCC	80	0 B-8261	–	–	–	
CYC1-1371	(Met)Thr - Glu - Phe - Tyr - Ser - ATA ATG ACT GAA TTC TAC TCG	90	0 B-9423	0 B-9563	0 B-9564	–	a
CYC1-1071	(Met)Cys - Asp - Ser - Glu - Val - Ala - ATA ATG TGT GAC TCC GAA GTT GCC	50	0 B-8237	–	–	–	b
CYC1-987	Ac (Met)Ser - Glu - Phe - Leu - Ala - ATA ATG TCT GAA TTC TTG GCC	25	100 B-10645	0 B-7658	0 B-10672	(0) B-11973	
CYC1-963	Ac (Met)Gly - Glu - Phe - Leu - Ala - ATA ATG GGT GAA TTC TTG GCC	100	100 B-8032	0 B-8360	0 B-9074	0 B-11974	
CYC1-962	Ac (Met)Ala - Glu - Phe - Leu - Ala - ATA ATG GCT GAA TTC TTG GCC	60	100 B-8031	0 B-8361	0 B-9073	0 B-11975	
CYC1-853	Ac - Met - Glu - Phe - Leu - Ala - ATA ATA ATG GAA TTC TTG GCC	80	100 B-7687	100 B-7727	100 B-9072	0 B-11863	
CYC1-872	Ac - Met - Asn - Asn - Leu - Ala - ATA ATA ATG AAC AAC TTG GCC	70	100 B-7723	~95 B-7846	100 B-9071	0 B-11864	
CYC1-1162	Ac - Met - Ile - Arg - Leu - Lys - Ala - AAT ATG ATT AGA TTG AAG GCC	30	94 B-8462	100 B-9024	0 B-9022	100 B-11865	
CYC1-1201	Ac - Met - Leu - Arg - Phe - Arg - Ala - AATAAT ATG TTG AGA TTC AGA GCC	30	96 B-8566	100 B-8565	0 B-9076	100 B-12040	c
CYC1-1286	Ac - Met - Leu - Ala - Phe - Ala - Val - Thr - Ala - TAAA ATG TTG GCT TTC GCC GTT ACT GCC	40	100 B-10646	100 B-10673	0 B-10688	100 B-12041	

[a]N-terminal sequence corresponds to the N-terminal sequence of Erg7p, Ac-Thr-Glu-Phe-Tyr-Ser- which is acetylated in *ARD1*$^+$ but not *ard1*$^-$ strains (Corey *et al*., 1994; E.J.Corey, unpublished).
[b]N-terminal sequence corresponds to a mutant form of actin, Ac-Met-Cys-Asp-Ser-Glu-Val-, partially acetylated in both *ARD1*$^+$ and *ard1*$^-$ strains (Cook *et al*., 1991).
[c]N-terminal sequence corresponds to N-terminal sequence of the major coat protein, Ac-Met-Leu-Arg-Phe-Val-, acetylated in *MAK3*$^+$ but not *mak3*$^-$ strains (Tercero and Wickner, 1992).
The (0%) value of acetylation of the *CYC1-987 nat3*-Δ iso-1 was estimated from the results of microsequencing.

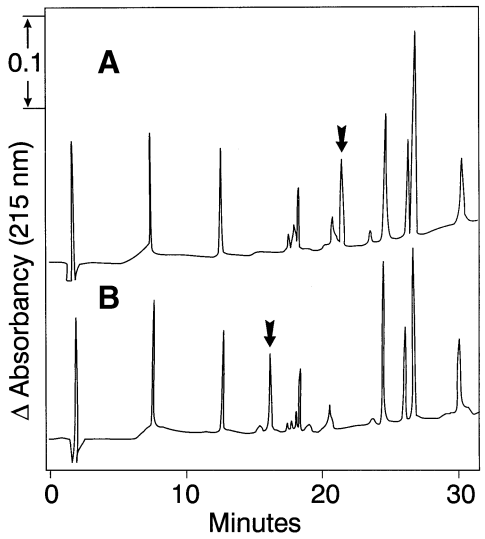

Fig. 4. Examples of HPLC profiles of peptides obtained by lysylendopeptidase digestion of mutant forms of iso-1-cytochrome *c*. (**A**) *CYC1-987 NAT1*$^{+}$ (B-10645). (**B**) *CYC1-987 nat1*-Δ (B-7658). Arrows in (A) and (B) indicate, respectively, Ac-Ser-Glu-Phe-Leu-Ala-Gly-Ser-Ala-Lys and Ser-Glu-Phe-Leu-Ala-Gly-Ser-Ala-Lys. A total of 2.5 nmol of each digests was applied to a μBondaspere column and eluted with a linear gradient as described in the Materials and methods. The peptides, indicated by arrows, derived from N-termini were collected and subjected to amino acid analysis.

by digestion with lysylendopeptidase (Table IV). As illustrated in Figure 4, this procedure is quantitative and allows the detection of partially acetylated forms with a sensitivity greater than 5%. By and large, the samples were either completely or almost completely acetylated, or were completely unacetylated. As discussed below, certain altered forms uncovered in previous studies were only partially acetylated.

The lack of acetylation of an iso-1 in one or another of the *nat1*-Δ, *mak3*-Δ or *nat3*-Δ mutants clearly revealed the dependence on one or more of the NATs as presented in Table IV. The Ac-Met-Glu- (*CYC1-853*) and Ac-Met-Asn-Asn- (*CYC1-872*) iso-1 are not acetylated only in the *nat3*-Δ mutant, and the Ac-Met-Ile-Arg- (*CYC1-1162*), Ac-Met-Leu-Arg- (*CYC1-1201*) and Ac-Met-Leu-Ala- (*CYC1-1286*) iso-1 are not acetylated only in the *mak3*-Δ mutant. However, the lack of acetylation of the Ac-Ser-Glu-Phe- (*CYC1-987*), Ac-Gly-Glu-Phe- (*CYC1-963*) and Ac-Ala-Glu-Phe- (*CYC1-962*) iso-1 in the *nat1*-Δ, *nat3*-Δ and *mak3*-Δ mutants, and acetylation in the normal strain indicate a role in for all three NATs. Most importantly, the results with two-dimensional gel electrophoresis of total yeast proteins, described below, revealed that the acetylation of most Ser- and Ala-termini was not dependent on Mak3p or Nat3p.

Acetylation of abundant proteins

Two-dimensional gel electrophoresis can be used to investigate N-terminal acetylation by comparing the protein patterns of mutant strains with the protein patterns of control strains. The lack of acetylated proteins in the mutant strains is characterized by a shift in their isoelectric points towards a more basic pH. This method has been used previously in the analysis of identified yeast proteins from total extracts of *nat1* and *ard1* mutants (Boucherie *et al*., 1996; Garrels *et al*., 1997). However, this procedure would not reveal acetylated proteins that underwent further N-terminal processing, such as post-translational cleavage of N-terminal regions. Some such proteins with cleaved N-terminal regions are presented in Table V.

In this study, we used the two-dimensional gel method to investigate N-terminal acetylation by comparing the protein patterns of the normal strain (B-11679) with the protein patterns of the *nat1*-Δ (B-11815), *ard1*-Δ (B-11706), *mak3*-Δ (B-11705), *nat3*-Δ (B-11852) and *vip1*-Δ (B-11707) mutants (Table I). A total of 55 distinct spots, denoted 1–55 (Figure 5), from the protein sample extracted from the normal strain B-11679 were identified on the two-dimensional gel. These 55 proteins, previously identified by microsequencing of tryptic peptides (Norbeck and Blomberg, 1997) or by mass spectrometry (MS-MALDI) (J.Norbeck, T.Larsson, K.-A.Karlsson and A.Blomberg, unpublished data), are listed in Table V, along with their N-terminal sequences. When necessary, the identities of ambiguous proteins were confirmed by MS-MALDI analysis, and ambiguous shifts in position were confirmed by the two-dimensional gel analysis of mixtures of proteins from normal and mutant strains.

The largest number of protein shifts were observed with the *nat1*-Δ and *ard1*-Δ mutants, whose protein patterns were identical to each other, but differed from the normal in 24 of the 55 identified proteins (Tables V and VI). Portions of the normal, *nat1*-Δ and *ard1*-Δ two-dimensional gels are shown in Figure 6, illustrating the alkaline shifts with the Gln1p (No. 1), Tal1p (No. 8) and Adh1p (No. 11) proteins. Most significantly, all of the proteins affected by *nat1*-Δ and *ard1*-Δ (Table V, Nos 1–24) had penultimate residues of alanine or serine, consistent with the results of altered iso-1 (Table IV), and indicating that the N-terminal methionine residue was cleaved and the penultimate residues were acetylated in the normal strain, but not in the *nat1*-Δ or *ard1*-Δ mutants. The identity of the *nat1*-Δ and *ard1*-Δ protein patterns is consistent with the view that Nat1p and Ard1p are components of the same NAT (Mullen *et al*., 1989; Park and Szostak, 1992).

Only two of the 55 proteins identified from the *nat3*-Δ mutant were shifted, as shown in a portion of the two-dimensional gel presented in Figure 7. The two affected proteins, Act1p (No. 25) and Rnr4p (No. 26), have Met-Glu- and Met-Asp- residues (Tables V and VI), consistent with results of the altered iso-1, where Met-Glu- (*CYC1-853*) and Met-Asn- (*CYC1-873*) N-termini were not acetylated in *nat3*-Δ strains (Table IV).

Surprisingly, the protein pattern of the *mak3*-Δ mutant was normal, indicating that none of the 55 proteins listed in Table V were substrates for the Mak3p NAT. In contrast, *mak3*-Δ prevented the acetylation of not only the Ac-Met-Ile-Arg- (*CYC1-1162*), Ac-Met-Leu-Arg- (*CYC1-1201*) and Ac-Met-Leu-Ala- (*CYC1-1286*) iso-1, but also the Ac-Ser-Glu- (*CYC1-987*), Ac-Gly-Glu- (*CYC1-963*) and Ac-Ala-Glu- (*CYC1-962*) iso-1 that were also affected by *nat1*-Δ. Also, *mak3*-Δ is not only required for the N-terminal acetylation of *gag*, the major coat protein of the L-A double-stranded RNA virus, but also for the artificial β-galactosidase constructs having *gag*-like termini (Tercero *et al*., 1993). We suggest that the lack of effect of *mak3*-Δ on the two-dimensional gel protein pattern is simply due to the rarity of Map3p substrates.

Table V. Proteins whose positions on two-dimensional gels are affected or not affected by *nat1*-Δ, *ard1*-Δ, *nat3*-Δ or *mak3*-Δ

Affected by *nat1*-Δ and *ard1*-Δ		
(1)	Gln1p	(M) AEASIE···
(2)	Ssb1p	(M)AEGVFQ···
(3)	Ssb2p	(M)AEGVFQ···
(4)	Vma1p	(M)AGAIEN···
(5)	YMR116c	(M)ASNEVL···
(6)	YLR109w	(M)SDLVNK···
(7)	Pdc1p	(M)SEITLG···
(8)	Tal1p	(M)SEPAQK···
(9)	Gdh1p	(M)SEPEFQ···
(10)	YIL041w	(M)SFNAFA···
(11)	Adh1p	(M)SIPETQ···
(12)	Ssa1p	(M)SKAVGI···
(13)	Ssa2p	(M)SKAVGI···
(14)	Arg1p	(M)SKGKVC···
(15)	Sam2p	(M)SKSKTF···
(16)	Yst1p	(M)SLPATF···
(17)	Yst2p	(M)SLPATF···
(18)	Pgk1p	(M)SLSSKL···
(19)	Sti1p	(M)SLTADE··[a]
(20)	Pgi1p	(M)SNNSFT···
(21)	Adk1p	(M)SSSESI···
(22)	Sse1p	(M)STPFGL···
(23)	Bmh1p	(M)STSRED···
(24)	Bmh2p	(M)SQTRED···
Affected by *nat3*-Δ		
(25)	Act1p	MDSEVAA···
(26)	Rnr4p	MEAHNQF···
Affected by *mak3*-Δ		
None		
Not affected by *nat-1*Δ, *ard1*-Δ, *nat3*-Δ or *mak3*-Δ		
(27)	Hsc82p	(M)AGETFE···
(28)	Sam1p	(M)AGTFLF···
(29)	Tpi1p	(M)ARTFFV···
(30)	Eno1p	(M)AVSKVY···
(31)	Eno2p	(M)AVSKVY···
(32)	Pdb1p	MFSRLPT··[b]
(33)	RplA0p	(M)GGIREK···
(34)	Lys9p	(M)GKNVLL···
(35)	Fba1p	(M)GVEQIL···
(36)	YKL056c	MIIYKDI···
(37)	Psa1p	MKGLILV···
(38)	Gpp1p	MKRFNVL···
(39)	Hsp60p	MLRSSVV··[c]
(40)	Ilv5p	MLRTQAA··[d]
(41)	Gpm1p	(M)PKLVLV···
(42)	Thr4p	(M)PNASQV···
(43)	Met17p	(M)PSHFDT···
(44)	Shm2p	(M)PYTLSD···
(45)	Tuf1p	(M)SALLPR··[e]
(46)	Pyk1p	(M)SRLERL···
(47)	YJR105w	(M)TAPLVV···
(48)	Ald6p	(M)TKLHFD···
(49)	Ipp1p	(M)TYTTRQ···
(50)	Hxk2p	(M)VHLGPK···
(51)	Atp2p	(M)VLPRLY··[f]
(52)	Met6p	(M)VQSAVL···
(53)	Sod1p	(M)VQAVAV···
(54)	Tdh2p	(M)VRVAIN···
(55)	Tdh3p	(M)VRVAIN···

Cleavage of *N*-terminal methionine residues, shown in parentheses, was deduced from the radius of gyration of the penultimate residue (Moerschell *et al*., 1990).
[a]Partially acetylated.
[b]The first 33 amino acids are removed, and the N-terminus of the resolved form is SSTKTM.
[c]The first 25 amino acids are removed, and the N-terminus of the resolved form is ECKFGV.
[d]The first 47 amino acids are removed, and the N-terminus of the resolved form is LKQINF.
[e]An unknown number of amino acids are removed at the N-terminal region.
[f]The first 33 amino acids are removed, and the N-terminus of the resolved form is ASAAQS.

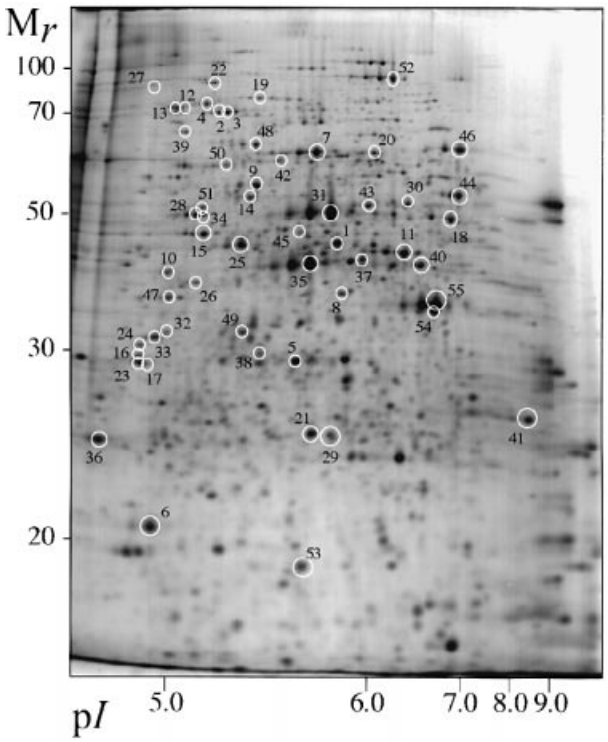

Fig. 5. A phosphoimager produced image of a two-dimensional polyacrylamide gel of protein extract prepared from the normal strain B-11679 labeled with [^{35}S]methionine during exponential growth in SD medium (minimal medium with glucose as the carbon source). The numbered positions in the figure correspond to the 55 proteins numbered in Table V.

Table VI. Summary of the 55 proteins examined by two-dimensional gels (summarized from Table V)

Acetylated			Not acetylated	
Terminus	No.	Acetylase	Terminus	No.
Ac-Ser-	19	Ard1p/Nat1p	Ser-	1
Ac-Ala-	5	Ard1p/Nat1p	Ala-	5
–			Gly-	3
–			Thr-	3
–			Val-	6
–			Pro-	4
Ac-Met-Glu-	1	Nat3p	Met-Glu-	0
Ac-Met-Asp-	1	Nat3p	Met-Asp-	0
–			Met-Ile-	1
–			Met-Leu-	2
–			Met-Lys-	2

Cleavage of *N*-terminal Met- was deduced from the radius of gyration of the penultimate residue (Moerschell *et al*., 1990).

However, the overlap of the iso-1 affected by both *nat1*-Δ and *mak3*-Δ mutations (Table IV), and the lack of effect of the *mak3*-Δ mutation on proteins 1–24 (Table V) is indeed puzzling.

The lack of acetylation of the Ac-Ser-, Ac-Ala- and Ac-Gly- iso-1 in the *mak3*-Δ and *nat3*-Δ mutants (Table IV), and the acetylation of the Ac-Ser- and Ac-Ala- number 1–24 proteins in the *mak3*-Δ and *nat3*-Δ mutants (Table V) cannot be explained by a difference in the strains. Identical results of two-dimensional gel analysis were obtained from the protein samples prepared from the following set of isogenic strains used in the

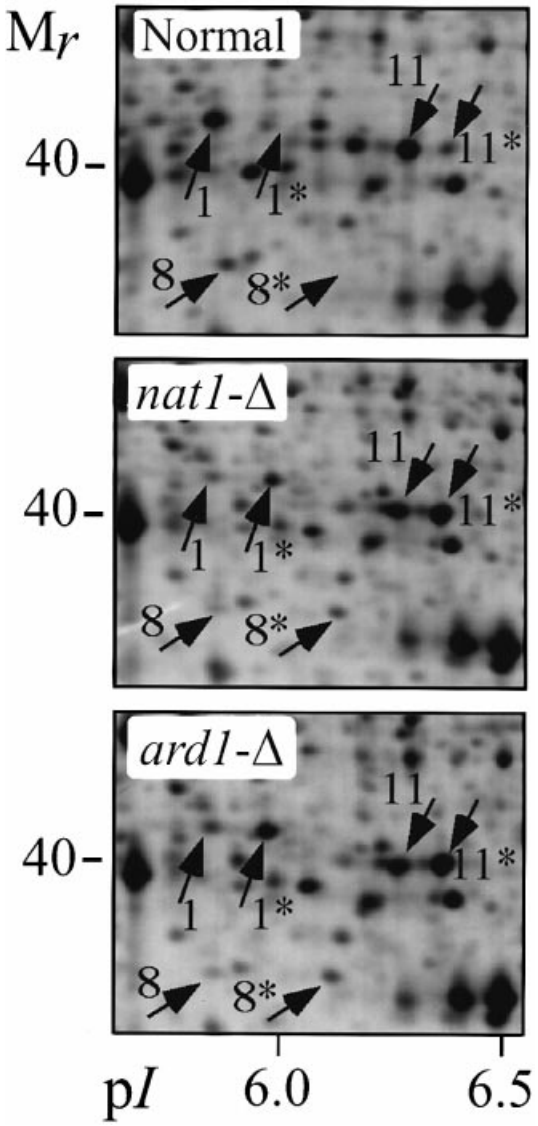

Fig. 6. Portions from gels of protein extracts prepared from exponentially growing cells of the normal strain B-11679 and the *nat1*-Δ and *ard1*-Δ deletion strains. The positions of three identified proteins, Gln1p, Tal1p and Adh1p, which show an alkaline shift in the *nat1*-Δ and *ard1*-Δ deletion mutants, are indicated, with the numbers corresponding to the proteins listed in Table V, and with an asterisk indicating the position of the non-acetylated form.

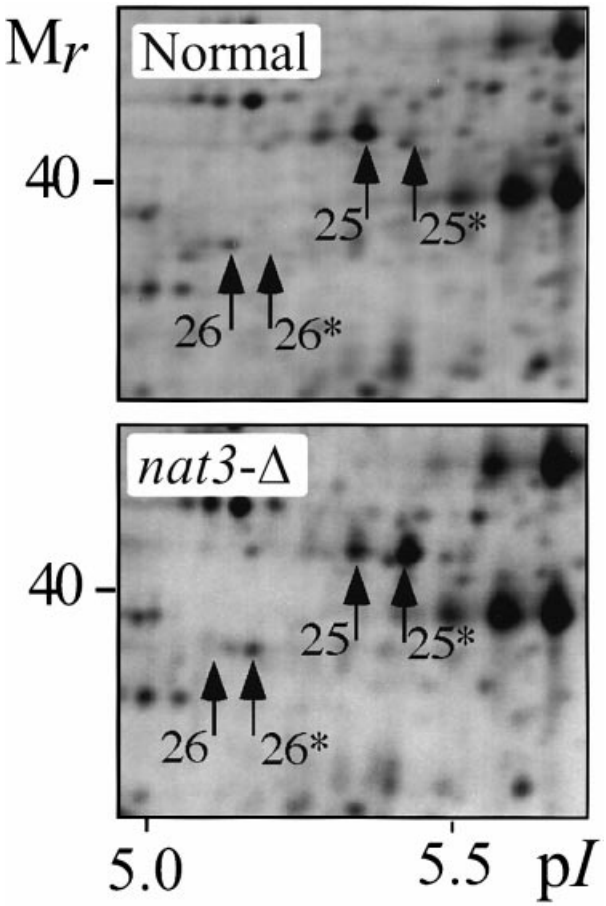

Fig. 7. Portions from gels of protein extracts prepared from exponentially growing cells of the normal strain B-11679 and the *nat3*-Δ deletion strain. The positions of Act1p (25) and Rnr4p (26), which both show an alkaline shift in the *nat3*-Δ deletion mutant, are indicated by numbers corresponding to the proteins listed in Table V, and with an asterisk indicating the position of the non-acetylated form.

analysis of altered iso-1: normal (B-10645); *nat1*-Δ (B-7658); *mak3*-Δ (B-10672); *nat3*-Δ (B-11973) and *vip1*-Δ (B-10551) (data not shown).

vip1-Δ or nat2-1 does not prevent acetylation in vivo

Lee *et al*. (1997) reported that purified Vip1p had NAT activity *in vitro* with a variety of peptides having methionine residues at the N-terminus, including sequences that, from our results, appear to be Nat3p substrates. However, in our study, the *vip1*-Δ deletion did not prevent N-terminal acetylation *in vivo* of the following iso-1 having these as well as other termini: Ac-Ser-Glu- (*CYC1-987*, B-10551); Ac-Met-Glu- (*CYC1-853*, B-10580); Ac-Met-Asn- (*CYC1-872*, B-10589); Ac-Met-Ile- (*CYC1-1162*, B-10564); and Ac-Met-Leu- (*CYC1-1286*, B-11404). Furthermore, the protein patterns on two-dimensional gels were normal in extracts prepared from the *vip1*-Δ strains B-11707 and B-10551. There is also a discrepancy in the molecular mass reported for Vip1p by Lee *et al*. (1997), 70 000 ± 5000 Da, and the molecular mass deduced from the sequence, 129 766 Da. Furthermore, Vip1p is orthologous to the *Schizosaccharomyces pombe* protein Asp1p, which is important for the function of the cortical actin cytoskeleton (Feoktistova *et al*., 1999). We suggest that Vip1p is not a NAT.

We have also investigated the strains B-10172 (*CYC1-872*) and B-10173 (*CYC1-1093*) containing the conditional mutation *nat2-1* that was reported to diminish N-terminal acetylation *in vitro* (Kulkarni and Sherman, 1994). Because of the temperature sensitivity, the *nat2-1* strains were first grown at 23°C and the culture was subsequently shifted to 37°C for 14 h. The iso-1 isolated from yeast grown using this procedure were fully acetylated, and thus we were unable to demonstrate that the *nat2-1* temperature-sensitive mutation prevents acetylation *in vivo*. It is possible that the growth of cells may be more sensitive to the deficiency of Nat2p activity than the acetylation of iso-1, preventing the recovery of unacetylated forms *in vivo*.

Discussion

The three NATs and their substrates

In this study we further investigated the yeast NATs, Ard1p/Nat1p and Mak3p, and identified and characterized the novel NAT, Nat3p. Because the combination of Ard1p and Nat1p subunits produced NAT activity *in vivo* and *in vitro* (Mullen *et al*., 1989; Park and Szostak, 1992), and because of the amino acid sequence similarity of Ard1p, Mak3p and Nat3p, and the presence of presumptive CoA-binding sites (Figure 1), we suggest that these three proteins are catalytic proteins of NATs. While Nat1p enhanced the activity of the presumably catalytic Ard1p subunit, it is not known whether Mak3p or Nat3p are similarly associated with subunits that stimulate activity. Either of the *ard1*-Δ or *nat1*-Δ mutations completely prevent N-terminal acetylation of the same set of proteins. The suggestion that Ard1p and not Nat1p represents the catalytic subunit is at variance with the conclusion of Lee *et al*. (1988) who reported that Nat1p, acting as a dimeric protein, was sufficient to produce NAT activity.

Because of the lack of methionine cleavage from penultimate residues that are large, the N-terminal residues of proteins consist of Gly, Ala, Ser, Cys, Thr, Pro and Val after cleavage and, of course, Met without cleavage, unless additional processing occurs. Previous work, and the work presented here with the iso-1 system, as well as the work of Huang *et al*. (1987) on the thaumatin plant protein expressed in yeast suggest that subsets of proteins with only the N-termini residues Ser, Ala, Gly, Thr and Met are cotranslationally acetylated. Although rare examples of N-terminal acetylation of proteins with Cys-, Pro- and Val-termini have been reported (Driessen *et al*., 1985; Flinta *et al*., 1986; Cook *et al*., 1991), it is not clear whether they involved cotranslational processing. Also, we have not considered the results obtained *in vitro* with the rat NAT, which acetylates short synthetic peptides with methionine, serine and other terminal residues (Yamada and Bradshaw, 1991a,b).

The three NATs, Ard1p/Nat1p, Nat3p and Mak3p, can be considered to act on different groups of substrates, NatA, NatB, NatC and NatD, as summarized in Table VII. The major NAT, Ard1p/Nat1p, acetylates subsets of proteins with termini of Ser, Ala, Gly and Thr, but not Met. The lack of normally acetylated Ac-Ser- and Ac-Ala- proteins in *ard1*-Δ or *nat1*-Δ is clearly exemplified by the results of the two-dimensional gels presented in Tables V and VI, showing that 18 of 21 Ser proteins are normally acetylated, and five of 10 Ala proteins are normally acetylated. Similar results were reported by Garrels *et al*. (1997), who reported that 48 of 51 Ser proteins are normally acetylated and seven of 15 Ala proteins are normally acetylated, and that acetylation of all of these proteins did not occur in the *nat1*-Δ mutant. Furthermore, Takakura *et al*. (1992) reported that all eight ribosomal proteins initiating with serine were acetylated in normal but not *nat1*-Δ strains. There are less examples of NatA substrates having Ac-Gly- and Ac-Thr-termini, and these are presented in Table VII.

The NatB substrates acetylated solely by Nat3p are also presented in Table VII. In addition, numerous Ac-Met-Met-Asn- iso-1 are acetylated, but it has not been tested whether the acetylation is lacking in *nat3*-Δ strains.

Table VII. The substrate classes

N-termini	Examples
NatA substrates, acetylation requiring Nat1p or Ard1p	
Ac-Ser-	*CYC1-987*; etc.
Ac-Ala-	*CYC1-962*; etc.
Ac-Gly-	*CYC1-963*
Ac-Thr-	*pup1ΔLS*; *pre3ΔLS*[a]
NatB substrates, acetylation requiring Nat3p	
Ac-Met-Glu-	Rnr4p; *CYC1-853*
Ac-Met-Asp-	Act1p; *CYC1-848*
Ac-Met-Asn-Asn-	*CYC1-872*; *CYC1-345-H*
(Ac-Met-Met-Asn-)	*CYC1-183-T, CYC1-242-N*; etc.
NatC substrates, Met- proteins acetylation requiring Mak3p	
Ac-Met-Ile-Arg-Leu-Lys-	*CYC1-1162*; *CYC1-31-Y*
Ac-Met-Leu-Arg-Phe-Arg-	*CYC1-1201*
Ac-Met-Leu-Ala-Phe-Ala-	*CYC1-1286*
(Ac-Met-Ile-Lys-Phe-Lys-)	*CYC1-9-BU*
NatD substrates, a subset of NatA substrates requiring all three NATs for acetylation	
Ac-Ser-Glu-Phe-	*CYC1-987*
Ac-Ala-Glu-Phe-	*CYC1-962*
Ac-Gly-Glu-Phe-	*CYC1-963*

The acetylated sequences in parentheses are similar to the sequences in the same group, but have not been demonstrated to require the NAT of the group. [a]*pup1ΔLS*; *pre3ΔLS* are from Arendt and Hochstrasser (1999).

Similarly, the NatC substrates, Ac-Met-Ile- and AcMet-Leu- iso-1, acetylated solely by Mak3p, are presented in Table VII. Also, *gag* and *gag–lacZ* fusion protein derivatives having Ac-Met-Leu and Ac-Met-Trp sequences are considered to be NatC substrates (Tercero *et al*., 1993).

The NatD substrates are a subset of NatA substrates that denote the proteins that are acetylated in normal, but not in *nat1*-Δ (or presumably *ard1*-Δ), *nat3*-Δ or *mak3*-Δ strains (Tables IV and VII), and these include the Ac-Ser-Glu-Phe- (*CYC1-987*), Ac-Ala-Glu-Phe- (*CYC1-962*) and Ac-Gly-Glu-Phe- (*CYC1-963*) iso-1. Clearly, these NatD substrates are expected primarily to require the Ard1p/Nat1p NAT for acetylation. Why the NatD substrates also require Mak3p and Nat3p is unknown. One cannot argue that either Ard1p/Nat1p, Nat3p or Mak3p can acetylate NatD substrates because an acetylation deficiency would require three mutations, such as in a triple *ard1*-Δ *mak3*-Δ *nat3*-Δ strain. Thus, it appears that acetylation of NatD substrates has an additional requirement beyond that needed for acetylation of the bulk of NatA substrates, and that Mak3p and Nat3p are essential for this additional requirement. It is possible that acetylation of NatD substrates requires Ard1p/Nat1p, as well as auxiliary factors whose activities are dependent on acetylation by Nat3p and Mak3p. We are currently determining whether any normal protein is a NatC substrate, and if any altered iso-1 are NatA but not NatD substrates.

Sequences required for N-terminal acetylation

The sequences of the substrates required for N-terminal acetylation can now be better evaluated by separately considering the NatA, NatB and NatC substrates. Only the NatB substrates have common sequences that can be easily deciphered, composed of Ac-Met-Glu-, Ac-Met-Asp-, Ac-Met-Asn-Asn- and probably Ac-Met-Met-Asn- sequences (Table VII). As emphasized by Moerschell *et al*.

Table VIII. Examples of similar sequences that are completely, partially or not acetylated

Ac-Ser-Leu-Pro-Ala-	(100%)	Yst1p
Ser-Leu-Thr-Ala-	(50%)	Sti1p
Ac-Ala-Gly-Ala-Ile-	(100%)	Vma1p
Ala-Gly-Thr-Phe-	(0%)	Sam1p
Ac-Gly-Glu-Phe-Leu-	100%	*CYC1-963*
Ac-Gly-Asp-Val-Glu-	70%	*CYC1-RAT1*
Ac-Thr-Glu-Phe-Tyr-	(100%)	Erg7p
Thr-Glu-Phe-Tyr-	0%	*CYC1-1371*
Thr-Glu-Phe-Leu-	0%	*CYC1-795*
Ac-Met-Glu-Phe-Leu-	100%	*CYC1-853*
Ac-Met-Asp-Phe-Leu-	100%	*CYC1-848*
Ac-Met-Glu-Phe-Lys-	55%	*CYC1-838*
Ac-Met-Asp-Pro-Leu-	67%	*CYC1-878*
Ac-Met-Asn-Asn-Ser-	100%	*CYC1-872*
Ac-Met-Asn-Phe-Leu-	79%	*CYC1-849*
Met-Asn-Gln-Phe-	(0%)	*CYC1-9-CB*
Ac-Met-Met-Asn-Ser-	100%	*CYC1-183-T*
Met-Met-Ile-Met-	0%	*CYC1-242-O*
Ac-Met-Ile-Arg-Ile	100%	*CYC1-31-Y*
Ac-Met-Ile-Lys-Phe	100%	*CYC1-9-BV*
Met-Ile-Thr-Gln-	0%	*CYC1-345-K*
Ac-Met-Leu-Arg-Phe-	100%	*CYC1-1201*
Ac-Met-Leu-Arg-Ala-	(100%)	JC33A
Met-Leu-Arg-Glu-	(0%)	JC33B
Met-Leu-Arg-Thr-	0%	Ilv5p
Ac-Met-Leu-Ala-Phe-	100%	*CYC1-1286*
Met-Leu-Phe-Leu-	0%	*CYC1-850*

The percentage acetylation, shown in parentheses, are estimations. Erg7p is from Corey *et al*. (1994) and unpublished results; and JC33A and JC33B from Tercero *et al*. (1993).

(1990), all seven eukaryotic Met-Glu- and Met-Asp- proteins uncovered in literature and database searches were N-terminally acetylated, but none of the six prokaryotic proteins with the same terminal residues. However, there are Met-Glu- and Met-Asp- iso-1 with reduced efficiency of acetylation, including *CYC1-838* (55%) and *CYC1-878* (67%) (Table VIII). We suggest that all the NatB substrates contain any one of these required sequences, but acetylation is diminished by inhibitory residues. For example, from the result with the Ac-Met-Asp-Pro- iso-1 (*CYC1-878*) having only 67% acetylation, one can suggest that adjacent proline residues diminish the action of Nat3p. Similarly, Moerschell *et al*. (1990) demonstrated that antepenultimate (the third residue) proline residues can inhibit methionine cleavage from certain residues with intermediate sizes of side chain. Also, methionine cleavage was completely inhibited from the Met-Val-Pro- sequence of a mutant human hemoglobin (Barwick *et al*., 1985; Prchal *et al*., 1986).

We suggest that the methionine aminopeptidases and NATs act on substrates with specific, but degenerate, sequences and that the activities of these cotranslational processing enzymes can be diminished by suboptimal residues. We further suggest that processing can be diminished by the inhibitory residues situated anywhere on the nascent chain at the time of acetylation. Thus, the degree of processing is the net effect of positive optimal or suboptimal residues and negative inhibitory residues. The lack of acetylation could be due to the absence of required residues or the presence of inhibitory residues. For example, the antepenultimate residues, Asn, in the *CYC1-872* sequence Ac-Met-Asn-Asn-, can be considered optimal, allowing complete acetylation, whereas the antepenultimate residues, Phe, in the *CYC1-849* sequence Ac-Met-Asn-Phe-, can be considered suboptimal, resulting in only 79% acetylation (Table VIII). However, the antepenultimate residue, Gln, in the *CYC1-9-CB* sequence Met-Asn-Gln-, prevents acetylation because it is not part of a required sequence. From the other point of view, the Lys residue in the *CYC1-838* sequence Ac-Met-Glu-Phe-Lys- and the Pro residue in the *CYC1-878* Ac-Met-Asp-Pro-Leu- can be considered inhibitory residues. Because the identities of required and inhibitory residues are unknown, the ability of a protein to be acetylated cannot be predicted definitively from the primary sequence. Also, it is unclear whether an amino acid position is occupied by required or inhibitory residues. In the NatC series Met-Leu-Arg-*Any*-, represented by the *CYC1-1201*, JC33A, JC33B and Ilv5p proteins (Table VIII), it is not known whether the '*Any*' residue is part of the required sequence, or if it depicts residues inhibiting the action of the Met-Leu-Arg- sequence.

Furthermore, because Erg7p, but not the *CYC1-1371* iso-1, is acetylated (Table VIII), it appears as though inhibitory residues could occupy sites further than five amino acid residues from the terminus.

In this regard, Huang *et al.*'s (1987) major conclusion on the generality of acetylation based on one or two residues has been misleading and has led to erroneous predictions that were incorporated in some databases.

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, ~90 and 50% of the Ser- and Ala- proteins, respectively, are acetylated. The diversity and high degeneracy prevent prediction of acetylated and nonacetylated sequences. Early attempts to predict N-terminal acetylation based on the properties of amino acid residues distributed along the N-terminal region were unsuccessful (Jörnvall, 1975; Driessen *et al*., 1985; Persson *et al*., 1985; Augen and Wold, 1986; Flinta *et al*., 1986).

Are there other NATs that act cotranslationally?

Evidence for the existence of NATs, other than the three described in this paper, would come from retention of acetylation in all of the *ard1*-Δ (or *nat1*-Δ), *mak3*-Δ and *nat3*-Δ mutants. Such findings would come from testing iso-1 mutants, as described here, or testing proteins known to be acetylated. As clearly revealed in Table IV, all acetylated iso-1 can be assigned to one of the NatA, NatB or NatC substrates, and these iso-1 represent diverse groups. Also, the acetylated iso-1 not tested with NAT-deficient mutants have N-terminal sequences that appear to be closely related to one or another of the NatA, NatB or NatC substrates (Table VIII). Furthermore, we do not know of any acetylated proteins in yeast that could not reasonably be a NatA, NatB or NatC substrate. Nevertheless, it remains to be seen whether there are other NATs acting on still rarer substrates not uncovered in this study.

In this regard, Tercero *et al*. (1993) reported that the blocked Ac-Met-Leu-Arg-Phe-Ala- variant apparently remained acetylated in the *mak3*-Δ mutants. It is not clear to us whether the sequence is responding to an as yet unidentified NAT or whether there was an error in the analysis.

We also wish to point out that human orthologues of the yeast Ard1p and Nat3p, but not of Mak3p or Nat1p, were uncovered in literature and database searches. A 100 amino acid region encompassing portions of motifs A, B and D of *h*Ard1p (Tribioli *et al*., 1995) and *h*Nat3p (R.Strausberg, NCI accession No. AA262571) have similarities of 74 and 58%, respectively, to the yeast counterparts.

Biological importance of N-terminal acetylation

The finding that N-terminal acetylation, occurring post-translationally, causes increased melanotropic effects of α-melanocyte-stimulating hormone, while reducing the analgesic action of β-endorphin, represents the clearest example of the biological importance of this modification (Smyth *et al*., 1979; Smyth and Zakarian, 1980; Glembotski, 1982a,b). However, there are surprisingly few examples demonstrating the biological importance of N-terminal acetylation occurring cotranslationally. Most of the examples cited are based on comparisons of acetylated and unacetylated forms that had additional differences in amino acid sequence (Sherman *et al*., 1993).

A significant means of assessing the general importance of N-terminal acetylation comes from the phenotypic defects in the *ard1*-Δ (or *nat1*-Δ), *mak3*-Δ and *nat3*-Δ mutants (Table III; Figures 2 and 3). The interpretation of the defects due to the lack of N-terminal acetylation in the *nat1*-Δ, *ard1*-Δ (Mullen *et al*., 1989) and *mak3*-Δ mutants (Tercero *et al*., 1993) has been discussed above, including the importance of N-terminal acetylation of the viral major coat protein, *gag*, for assembly or maintenance of the viral particle. Similarly to the other mutants, *nat3*-Δ mutants exhibit multiple defective phenotypes, including lack of growth on YPG medium at 37°C, reduced growth on medium containing NaCl and reduced mating of the *MAT*α cells (Figures 2 and 3; Table III). Such defects could arise from the lack of acetylation of any of a number of proteins essential for different processes. While the unacetylated proteins responsible for these defects are not easily identified, the temperature and NaCl sensitivity could be attributed to a lack of acetylation of Act1p, which contains a normal N-terminal sequence Ac-Met-Asp-Ser-Glu- (Table V). Many actin mutants are temperature and NaCl sensitive, including *act1-136*, which has Asp→Ala replacement at position 2 (Wertman *et al*., 1992).

The viability of *ard1*-Δ, *nat1*-Δ, *mak3*-Δ and *nat3*-Δ mutants lacking NATs suggests that the role of acetylation may be subtle and not absolute for most proteins. It is possible that only a subset of proteins actually requires this modification for activity or stability, whereas the remainder are acetylated only because their termini fortuitously correspond to consensus sequences.

Materials and methods

Yeast strains and media

Unless stated otherwise, yeast were grown at 30°C in YPD or YPG media or SD medium containing appropriate supplements (Sherman, 1991). NaCl medium refers to YPD medium with 1.0 M NaCl. The strains used in this study are listed in Table II. Analysis of N-terminal acetylation was carried out using two isogenic series (Table I), with each set originally derived from either B-7528 (*cyc1-31 MAT***a** *cyc7-67 ura3-52 lys5-10*) or B-11679 (*MAT*α *ura3-52*), also denoted YFJN001-1A. The series used for producing altered iso-1 (Table IV) were constructed by first using synthetic oligonucleotides to generate *CYC1* mutations encoding the desired iso-1 and then by separately disrupting each of the *NAT1*, *MAK3*, *NAT3* or *VIP1* genes (see below). Similarly, the series used for phenotypic studies and two-dimensional gel analysis were generating from B-11679 by separately disrupting each of the *ARD1*, *NAT1*, *MAK3*, *NAT3* or *VIP1* genes. In this paper, *nat1*-Δ, for example, denotes either *nat1*-Δ::*URA3* or *nat1*-Δ::*kanMX2* disruptions.

Mating efficiencies

Quantitative matings were determined by plating serial dilutions of logarithmically growing yeast cells onto SD plates containing a lawn of tester mating strain, B-6924 or B-6925, and determining the frequencies of prototrophic diploid colonies arising after incubation for 3 days. Dilutions of the haploid strains were also plated on YPD plates to determine the total number of cells. Mating efficiencies were expressed as the number of diploid colonies divided by the number of haploid cells plated on the lawn of the tester strain.

Construction of iso-1 mutants

The altered forms of iso-1 were conveniently produced by transforming yeast directly with synthetic oligonucleotides as described previously by Yamamoto *et al*. (1992), using the *cyc1-31* mutant B-7528 (Table II) and 50 μg of the oligonucleotides, listed in Table IX, that have minimal mismatches.

Iso-1 content

The amounts of iso-1 in the transformants were estimated by low temperature (–196°C) spectroscopic examinations of intact cells (Sherman and Slonimski, 1964) and comparing the intensities of the C_α band at 547 nm to the corresponding band of strains having known amounts of cytochrome *c*.

Construction of deletion mutants

Standard molecular biological procedures were performed as described previously (Sambrook *et al*., 1989). The *NAT3*, *NAT1*, *ARD1*, *MAK3* and *VIP1* genes were disrupted by replacing a portion of the genes with either the *kanMX2* or *URA3* genes and using the appropriate fragment for gene replacements.

The PCR-generated fragment required for producing the *nat3*-Δ::*kanMX2* disruption was prepared by the method of Baudin *et al*. (1993), using the pFA6-kanMX2 plasmid (Wach *et al*., 1994) as template and the primers oligomers 1 and 2 (Table X). The correct disruption was identified by PCR, using the set of primer oligomers 3 and 4. Similarly, the fragment required for producing the *nat1*-Δ::*kanMX2* disruption was prepared with oligomers 5 and 6, and the fragment required for producing the *nat1*-Δ::*URA3* disruption was prepared from the plasmid pAA1132 (also denoted JM111) (Mullen *et al*., 1989). The correct *nat1*-Δ disruption was identified with oligomers 7 and 8. The fragment required for producing the *ard1*-Δ::*kanMX2* disruption was prepared from plasmid pAB2341, which was constructed by inserting the entire *ARD1* ORF, from –1 to +973, in the pCR2.1 TOPO plasmid vector (Invitrogen, Carlsbad, CA) and subsequently replacing the *ARD1* coding region between the *Bgr*GI and *Ssp*I restriction sites with the *kanMX2* gene from pFA6-kanMX2 (Wach *et al*., 1994). The correct *ard1*-Δ disruption was identified with oligomers 9 and 10. Similarly, the fragment required for producing the *mak3*-Δ::*kanMX2* disruption was prepared from the plasmid pAB2340 that was constructed by inserting the entire *MAK3* ORF, from –133 to +668, in pCR2.1 TOPO plasmid vector and subsequently replacing the *MAK3* coding region between the *Acc*I and *Sna*BI restriction sites with the *kanMX2* gene from pFA6-kanMX2. The fragment required for producing the *mak3*-Δ::*URA3* disruption was prepared from the plasmid pJC11C (pAA1131) (Tercero *et al*., 1992). The correct *mak3*-Δ disruptions were identified with oligomers 11 and 12.

Similarly, the fragment required to produce the *vip1*-Δ::*kanMX2* disruption was prepared with oligomers 13 and 14, the fragment required for producing the *vip1*-Δ::*URA3* disruption was prepared with oligomers 15 and 16, and the correct *vip1*-Δ disruption was identified with oligomers 17 and 18.

Quantitative analysis of acetylated iso-1

Altered iso-1 were isolated as described previously (Moerschell *et al*., 1990) and further purified by HPLC on a μBondasphere column (3.9 × 150 mm, C4, 300 Å, Waters Associates), using a linear gradient elution of 0–80% 2-propyl alcohol:acetonitrile (7:3, v/v) in 0.1% trifluoroacetic acid at a flow rate of 1 ml/min for 40 min. The degree of N-terminal acetylation of the altered iso-1 was estimated by direct

Table IX. Oligonucleotides used to produce *CYC1* mutants by transformation of the *cyc1-31* strain B-7528[a]

Allele	Strain no.	Oligomer no.	Oligonucleotide sequence
CYC1-795	B-7471	OL86.123	ATACACACACTAAATTAATAATGACTGAATTCTTGGCCGGTTCTGCTAAGAAAGGTGCTA
CYC1-853	B-7687	OL87.314	AAATTAATAATAATGGAATTCTTGGCCGGTTCTGCTAAGA
CYC1-872	B-7723	OL88.083	AAATTAATAATAATGAACAACTTGGCCGGTTCTGCTAAGA
CYC1-962	B-8031	OL89.161	AAATTAATAATGGCTGAATTCTTGGCCGGTTCTGCTAAGA
CYC1-963	B-8032	OL89.162	AAATTAATAATGGGTGAATTCTTGGCCGGTTCTGCTAAGA
CYC1-987	B-10645	OL89.200	TCTTAGCAGAACCGGCCAAGAATTCAGACATTATTAATTT
CYC1-1070	B-8236	OL91.133	AAATTAATAATGTGTGAATTCTTGGCCGGTTCTGCTAAGA
CYC1-1071	B-8237	OL91.134	CACTAAATTAATGTGTGACTCCGAAGTTGCCGGTTCTGCT
CYC1-1092	B-8261	OL91.245	AAATTAATAATGGTTGAATTCTTGGCCGGTTCTGCTAAGA
CYC1-1093	B-8262	OL91.246	AAATTAATAATGCCTGAATTCTTGGCCGGTTCTGCTAAGA
CYC1-1162	B-8462	OL92.103	AAATTAATAATGATTAGATTGAAGGCCGGTTCTGCTAAGA
CYC1-1201	B-8566	OL92.176	AAATTAATAATGTTGAGATTCAGAGCCGGTTCTGCTAAGA
CYC1-1286	B-10646	OL93.175	TACACACACTAAAATGTTGGCTTTCGCCGTTACTGCCGGTTCTGCT
CYC1-1371	B-9423	OL95.036	AAATTAATAATGACTGAATTCTATTCCGGTTCTGCTAAG

[a]B-7528 (*MAT***a** *cyc1-31 cyc7-67 ura3-52 lys5-10*). ATG and CAT initiator codons are underlined in, respectively, the sense and antisense oligonucleotides.

Table X. Oligonucleotides used in the construction and testing of disrupted genes

ORF	Oligomer no.	Sequence (5′→3′)
NAT3	oligo. 1	(–70) GCAAAAAACATAGCGGTGG GCGCACGTTGGTGGGCCTGGGAAGAACA GCTGAAGCTTCG
	oligo. 2	(+814) TGAATAGCACAGAGGTTCATTATTATGTTCTGAGTATGAGGACGAAGGCCACTAGTGGATCTG
	oligo. 3	(–142) TTCCAATGCGGTAGTGCTTAGC
	oligo. 4	(+906) TATATCATTCAGTATGTACATAC
NAT1	oligo. 5	(+9) GAAAAGAAGTACTAAGCCCAAGCCAGCAGCTAAAATAGCTTTGAACAGCTGAAGCTTCGT
	oligo. 6	(2670) TAACAATGTATCTTTTATGAGCATATGCGCATGCATATAAGCATAGGC CACTAGTGGATC
	oligo. 7	(–41) GACAAATACCATTGAGGAAGGC
	oligo. 8	(+2741) GAAAAGGTCCTGGCTGTCCTGG
ARD1	oligo. 9	(–41) TAAATACATACGATCAAGCTCC
	oligo. 10	(+777) CTAACAGAGCTTGTGAAGAAGC
MAK3	oligo. 11	(–143) CCGGGTAGCCAAAAGCGATAAA
	oligo. 12	(+679) ACTTGCTTTATTATCTCTCTCC
VIP1	oligo. 13	(–100) ATTAATCATTGAATCAATATGTTCACCACTTAAATTACCCACTGAAGCTTCGTACGCTGC
	oligo. 14	(+3607) CGATGAAA ATATCTTGATGTTTTGTAGAATGCTCTTATTCAATAGGCCACTAGTGGATCT
	oligo. 15	(–234) TGCTTAGGACTTATTAAATTTCATTAGCAATCTCATCGCGGAAATTTCTTTTTTTGATTTCGG
	oligo. 16	(+3695) TACTGATTTCATCGCTACCTAATTCATCGGCCCCAAAAAGAGCTCGTCATTATAAAAATCATT
	oligo. 17	(–301) CCTCTACTTCTTACGTCTCTCTCG
	oligo. 18	(+3731) TGGCGGCATTACTGTCATCTA

The position of the first nucleotide is presented, where A of the ATG initiation codon is assigned position 1. The underlined sequences correspond to segments in the plasmid pFA6-kanMX2.

sequence analysis of the intact proteins and recoveries of pertinent peptides derived from their N-termini by protease digestion. The sequence analysis was carried out with a Hewlett Packard model G1000A sequencer. A total of 5 nmol of each cytochrome *c* sample were digested with lysylendopeptidase (Wako Pure Chemicals) at an enzyme to substrate ratio of 1:400 (mol/mol) in 100 µl of 100 mM Tris–HCl buffer (pH 9.0) at 37°C for 6 h. The resulting peptides were separated by HPLC on a µBondasphere column (3.9 × 150 mm, C18, 100 Å, Waters Associates), using a linear gradient of 0–40% 2-propyl alcohol:acetonitrile (7:3, v/v) in 0.1% trifluoroacetic acid at flow rate of 0.8 ml/min for 40 min. The peptides derived from the N-terminal region were identified by comparison of their HPLC profiles with that from one of a series of altered iso-1 followed by analyses of their amino acid compositions and sequences. The amino acid analyses were carried out with a PICO-TAG system (Waters Associates) after hydrolysis at 110°C *in vacuo* for 24 h in 5.7 M HCl; the recovery of each N-terminal peptide was calculated based on the values obtained. This proportion of acetylated iso-1 molecules can be estimated at a precision of 5% with this method.

Qualitative determination of blocked iso-1

In some instances, N-terminal acetylation of the altered iso-1 was estimated qualitatively from the degree of blockage of the Edman degradation process. Iso-1 were partially purified using a weak cation-exchange BioRex70 column, 200–400 mesh (BioRad, Hercules, CA) in potassium phosphate buffer pH 7.0 with a 0–1.0 M potassium chloride linear gradient. After chromatography the protein samples were dialyzed against 0.1% ammonium carbonate and, if necessary, concentrated by Centricon-3 (Amicon-Millipore, Bedford, MA).

Protein samples were separated by SDS–PAGE using the Tricine-based separation system (Schagger and von Jagow, 1987) and proteins were transferred onto a PVDF-membrane, Immobilon P (Millipore) in 10 mM 3-[cyclohexylamino]-1-propane-sulfonic acid buffer (Sigma, St Louis, MO), 10% methanol, pH 11.0, for 1 h at 350 mA. Half the membrane was used for iso-1 protein detection by Western analysis, using an anti-cytochrome *c* antibody. The other half of the membrane was stained in Coomassie Brilliant Blue G-250 for 2 min, then destained three times in 20% methanol, 10% acetic acid and finally 3–5 times in distilled water. The protein band corresponding to cytochrome *c* was excised from the membrane and subjected to N-terminal protein sequencing with a microsequencer, Model 473A (Applied Biosystems, Inc.).

Growth conditions, [^{35}S]methionine labeling and protein extraction

Cells were grown in 100 ml flasks with 10 ml of SD medium supplemented with 20 µg/ml of uracil, as described previously (Blomberg *et al*., 1995). Cultures were inoculated from an overnight culture to a density of ~10^6 cells/ml and growth was monitored as optical density at 610 nm (OD_{610}). At a density of 5×10^6 cells/ml (an OD_{610} of ~0.35), 150 µCi of [^{35}S]methionine were added to the culture for 30 min, after which cells were harvested by centrifugation.

Protein extract was made from the cells and the d.p.m./µl were

determined as described previously (Blomberg *et al*., 1995). Protein extract equivalent to 2×10^6 d.p.m. was lyophilized and dissolved in 50 µl of the urea buffer used to re-swell the IPG strips.

Two-dimensional gel electrophoresis

Two-dimensional PAGE was performed essentially as described previously (Norbeck and Blomberg, 1997) using 18 cm IPG-strips pH 3–10/NL (Amersham-Pharmacia Biotech, Uppsala, Sweden). These strips were reswollen overnight and protein extract equivalent to 2×10^6 d.p.m. was applied by sample cups at the anodic end. Gels were then focused for ~3×10^4 Vh, using a linear voltage gradient as described previously (Norbeck and Blomberg, 1997), with the last phase shortened to ~2 h. The second dimension SDS–PAGE gels were 12.5% T, 2.1% C (Duracryl 0.65% bis, Genomic Solutions Inc., Ann Arbor, MI), and the electrophoresis was carried out for ~6 h at 1.6×10^4 mW/gel (with a maximum voltage of 500 V) and at 20°C.

Following completion of the second dimension run, the gels were dried on filter paper. The dry gels were exposed to image plates for ~72 h and subsequently scanned at a resolution of 176 µm in a phosphoimager (Molecular Dynamics). Image files were imported to the PDQuest software (PDI) for further analysis. Protein identifications were performed either by microsequencing of tryptic peptides (Norbeck and Blomberg, 1997) or by mass spectrometry (MS-MALDI) (J.Norbeck, T.Larsson, K.-A.Karlsson and A.Blomberg, unpublished data.).

Acknowledgements

We thank R.B.Wickner for plasmids, R.Sternglanz for plasmids and useful discussions, and B.C.VanWuyckhuyse for microsequencing of some samples. This work was supported by a National Institutes of Health Grant R01 GM12702 (to F.S.) and by grants from the Swedish National Board for Natural Sciences (to A.B.).

References

Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.

Arendt,C.S. and Hochstrasser,M. (1999) Eukaryotic 20S proteasome catalytic subunit propeptides prevent active site inactivation by N-terminal acetylation and promote particle assembly. *EMBO J.*, **18**, 3575–3585.

Arfin,S.M., Kendall,R.L., Hall,L., Weaver,L.H., Stewart,A.E., Matthews,B.W. and Bradshaw,R.A. (1995) Eukaryotic methionyl aminopeptidases: two classes of cobalt-dependent enzymes. *Proc. Natl Acad. Sci. USA*, **92**, 7714–7718.

Augen,J. and Wold,F. (1986) How much sequence information is needed for the regulation of amino-terminal acetylation of eukaryotic proteins. *Trends Biochem. Sci.*, **11**, 494–497.

Baim,S.B., Pietras,D.F., Eustice,D.C. and Sherman,F. (1985) A mutation allowing an mRNA secondary structure diminishes translation of *Saccharomyces cerevisiae* iso-1-cytochrome *c*. *Mol. Cell. Biol.*, **5**, 1839–1846.

Barwick,R.C., Jones,R.T., Head,C.G., Shih,M.F-C., Prchal,J.T. and Shih,D.T-B. (1985) Hb Long Island: a hemoglobin variant with a methionyl extension at the NH_2 terminus and a prolyl substitution for the normal histidyl residue 2 of the b chain. *Proc. Natl Acad. Sci. USA*, **82**, 4602–4605.

Baudin,A., Ozier,K.O., Denouel,A., Lacroute,F. and Cullin,C. (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **21**, 3329–3330.

Blomberg,A. *et al.* (1995) Interlaboratory reproducibility of yeast protein patterns analyzed by immobilized pH gradient two-dimensional gel electrophoresis. *Electrophoresis*, **16**, 1935–1945.

Boucherie,H., Sagliocco,F., Joubert,R., Maillet,I., Labarre,J. and Perrot,M. (1996) Two-dimensional gel protein database of *Saccharomyces cerevisiae*. *Electrophoresis*, **17**, 1683–1699.

Bradshaw,R.A., Brickey,W.W. and Walker,K.W. (1998) N-terminal processing: the methionine aminopeptidase and N^{α}-acetyl transferase families. *Trends Biochem. Sci.*, **23**, 263–267.

Coleman,C.S., Huang,H. and Pegg,A.E. (1996) Structure and critical residues at the active site of spermidine/spermine-N1-acetyltransferase. *Biochem. J.*, **316**, 697–701.

Cook,R.K., Sheff,D.R. and Rubenstein,P.A. (1991) Unusual metabolism of the yeast actin amino terminus. *J. Biol. Chem.*, **266**, 16825–16833.

Corey,E.J., Matsuda,S.P.T. and Bartel,B. (1994) Molecular cloning, characterization and overexpression of *ERG7*, the *Saccharomyces cerevisiae* gene encoding lanosterol synthase. *Proc. Natl Acad. Sci. USA*, **91**, 2211–2215.

Corpet,F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.*, **16**, 10881–10890.

Dores,R.M., McDonald,L.K., Steveson,T.C. and Sei,C.A. (1990) The molecular evolution of neuropeptides: prospects for the '90s. *Brain Behavior Evol.*, **36**, 80–99.

Driessen,H.P.C., de Jong,W.W., Tesser,G.I. and Bloemendal,H. (1985) The mechanism of N-terminal acetylation of proteins. *CRC Crit. Rev. Biochem.*, **18**, 281–325.

Dutnall,R.N, Tafrov,S.T., Strenglanz,R. and Ramakrishnan,V. (1998) Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related *N*-acetyltransferase superfamily. *Cell*, **94**, 427–438.

Feoktistova,A., McCollum,D., Ohi,R. and Gould,K.L. (1999) Identification and characterization of *Schizosaccharomyces pombe asp1*$^+$, a gene that interacts with mutations in the Arp2/3 complex and actin. *Genetics*, **152**, 895–908.

Flinta,C., Persson,B., Jörnvall,H. and von Heijne,G. (1986) Sequence determinants of cytosolic N-terminal protein processing. *Eur. J. Biochem.*, **154**, 193–6.

Garrels,J.I. *et al.* (1997) Proteome studies of *Saccharomyces cerevisiae*: identification and characterization of abundant proteins. *Electrophoresis*, **18**, 1347–1360.

Glembotski,C.C. (1982a) Acetylation of α-melanotropin and β-endorphin in the rat intermediate pituitary. *J. Biol. Chem.*, **257**, 10493–10500.

Glembotski,C.C. (1982b) Characterization of the peptide acetyltransferase activity in bovine and rat intermediate pituitaries responsible for the acetylation of β-endorphin and α-melanotropin. *J. Biol. Chem.*, **257**, 10501–10509.

Hickman,A.B., Namboodiri,M.A.A., Klein,D.C. and Dyda,F. (1999) The structural basis of ordered substrate binding by serotonin *N*-acetyltransferase : enzyme complex at 1.8 Å resolution with a bisubstrate analog. *Cell*, **97**, 361–369.

Hilfiker,A., Hilfiker-Kleiner,D., Pannuti,A. and Lucchesi,J.C. (1997) *mof*, a putative acetyl transferase gene related to the *Tip60* and *MOZ* human genes and to *SAS* genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.*, **16**, 2054–2060.

Huang,S. *et al.* (1987) Specificity of cotranslational amino-terminal processing of proteins in yeast. *Biochemistry*, **26**, 8242–8246.

Jörnvall,H. (1975) Acetylation of protein N-terminal amino groups: structural observations on α-amino acetylated proteins. *J. Theor. Biol.*, **55**, 1–12.

Kendall,R.L., Yamada,R. and Bradshaw,R.A. (1990) Cotranslational amino-terminal process. *Methods Enzymol.*, **185**, 398–407.

Kulkarni,M. and Sherman,F. (1994) *NAT2*, an essential gene encoding methionine N^{α}-acetyltransferase in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **269**, 13141–13147.

Kuo,M.H., Zhou,J., Jambeck,P., Churchill,M.E. and Allis,C.D. (1998) Histone acetyltransferase activity of yeast Gcn5 is required for the activation of target genes *in vivo*. *Genes Dev.*, **12**, 627–639.

Lee,F.-J.S., Lin,L.-W. and Smith,J.A. (1988) Purification and characterization of an N^{α}-acetyltransferase from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **263**, 14948–14955.

Lee,F.-J.S., Lin,L.-W. and Smith,J.A. (1989a) Molecular cloning and sequencing of a cDNA encoding N^{α}-acetyltransferase from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **264**, 12339–12343.

Lee,F.-J.S., Lin,L.-W. and Smith,J.A. (1989b) N^{α}-acetylation is required for normal growth and mating of *Saccharomyces cerevisiae*. *J. Bacteriol.*, **171**, 5795–5802.

Lee,F.-J.S., Lin,L.-W. and Smith,J.A. (1997) A N^{α}-acetyltransferase selectively transfers an acetyl group to NH_2-terminal methionine residues: purification and partial characterization. *Biochim. Biophys. Acta*, **1338**, 244–252.

Li,X. and Chang,Y.-H. (1995) Amino-terminal protein processing in *Saccharomyces cerevisiae* is an essential function that requires two distinct methionine aminopeptidases. *Proc. Natl Acad. Sci. USA*, **92**, 12357–12361.

Lin,Y., Fletcher,C.M., Zhou,J., Allis,D. and Wagner,G. (1999) Solution structure of the catalytic domain of GCN5 histone acetyltransferase bound to coenzyme A. *Nature*, **400**, 86–89.

Moerschell,R.P., Hosokawa,Y., Tsunasawa,S. and Sherman,F. (1990) The specificities of yeast methionine aminopeptidase and acetylation of amino-terminal methionine *in vivo*: processing of altered iso-1-cytochromes *c* created by oligonucleotide transformation. *J. Biol. Chem.*, **265**, 19638–19643.

Mullen,J.R., Kayne,P.S., Moerschell,R.P., Tsunasawa,S., Gribskov,M., Colavito-Shepanski,M., Grunstein,M., Sherman,F. and Sternglanz,R.

(1989) Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. *EMBO J.*, **8**, 2067–2075.
Neuwald,A.F. and Landsman,D. (1997) GCN5-related histone *N*-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem. Sci.*, **22**, 154–155.
Norbeck,J. and Blomberg,A. (1997) Two-dimensional electrophoretic separation of yeast proteins using a non-linear wide range (pH 3–10) immobilized pH gradient in the first dimension: reproducibility and evidence for isoelectric focusing of alkaline proteins. *Yeast*, **13**, 1519–1534.
Park,E.-C. and Szostak,J.W. (1992) *ARD1* and *NAT1* proteins form a complex that has N-terminal acetyltransferase activity. *EMBO J.*, **11**, 2087–2093.
Persson,B., Flinta,C., von Heijne,G. and Jörnvall,H. (1985) Structures of N-terminally acetylated proteins. *Eur. J. Biochem.*, **152**, 523–527.
Prchal,J.T., Cashman,D.P. and Kan,Y.W. (1986) Hemoglobin Long Island is caused by a single mutation (adenine to cytosine) resulting in a failure to cleave amino-terminal methionine. *Proc. Natl Acad. Sci. USA*, **83**, 24–27.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schagger,H. and von Jagob,G. (1987) Tricine-sodium dodecyl sulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*, **166**, 368–379.
Sherman,F. (1991) Getting started with yeast. *Methods Enzymol.*, **194**, 3–21.
Sherman,F. and Slonimski,P.P. (1964) Respiration-deficient mutants of yeast. *Biochim. Biophys. Acta*, **90**, 1–15.
Sherman,F., Stewart,J.W. and Tsunasawa,S. (1985) Methionine or not methionine at the beginning of a protein. *BioEssays*, **3**, 27–31.
Sherman,F., Moerschell,R.P., Tsunasawa,S. and Sternglanz,R. (1993) N-terminal acetylation of mutationally altered form of iso-1-cytochromes *c* in normal and *nat1* strains deficient in the major N-terminal acetyl transferase of the yeast *Saccharomyces cerevisiae*. In Imahori,K. and Sakiyama,F. (eds), *Methods in Protein Sequence Analysis*. Plenum, New York, NY, pp. 173–181.
Smith,E.R., Eisen,A., Gu,W., Sattah,M., Pannuti,A., Zhou,J.,Cook,R.G., Lucchesi,J.C. and Allis,C.D. (1998) ESA1 is a histone acetyltransferase that is essential for growth in yeast. *Proc. Natl Acad. Sci. USA*, **95**, 3561–3565.
Smith,T.F. and Waterman,M.S. (1981) Comparison of bio-sequences. *Adv. Appl. Math.*, **2**, 482–489.
Symth,D.G. and Zakarian,S. (1980) Selective processing of β-endorphin in regions of porcine pituitary. *Nature*, **288**, 613–615.
Symth,D.G., Massey,D.E., Zakarian,S. and Finnie,.D.A. (1979) Endorphins are stored in biologically active and inactive forms: isolation of α-*N*-acetyl peptides. *Nature*, **279**, 252–254.
Takakura,H., Tsunasawa,S., Miyagi,M. and Warner,J. (1992) NH_2-terminal acetylation of ribosomal proteins of *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **267**, 5442–5445.
Tercero,J.C. and Wickner,R.B. (1992) *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**, 20277–20281.
Tercero,J.C., Riles,L.E. and Wickner,R.B. (1992) Localized mutagenesis and evidence for post-transcriptional regulation of *MAK3*. A putative *N*-acetyltransferase required for double-stranded RNA virus propagation in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **267**, 20270–20276.
Tercero,J.C., Dinman,J.D. and Wickner,R.B. (1993) Specificity of the yeast *MAK3 N*-acetyltransferase that modifies gag of the L-A dsRNA virus. *J. Bacteriol.*, **175**, 3192–3194.
Tribioli,C., Mancini,M., Plassart,E., Bione,S., Rivella,S., Sala,C., Torri,G. and Toniolo,D. (1994) 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**, 1061–1067.
Tsunasawa,S., Stewart,J.W. and Sherman,F. (1985) Amino-terminal processing of mutant forms of yeast iso-1-cytochrome *c*: the specificities of methionine aminopeptides and acetyltransferase. *J. Biol. Chem.*, **260**, 5382–5391.
Wach,A., Brachat,A., Pohlmann,R. and Philippsen,P. (1994) New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast*, **10**, 1793–1808.
Wang,X., Dumont,M.E. and Sherman,F. (1996) Sequence requirements for mitochondrial import of yeast cytochrome *c*. *J. Biol. Chem.*, **271**, 6594–6604.
Wertman,K.F., Drubin,D.G. and Bostein,D. (1992) Systematic mutational analysis of the yeast *ACT1* gene. *Genetics*, **132**, 337–350.
Wolf,E., Vassilev,A., Makino,Y., Sali,A., Nakatani,Y. and Burley,S.K. (1998) Crystal structure of a GCN5-related *N*-acetyltransferase: *Serratia marcescens* aminoglycoside 3-*N*-acetyltransferase. *Cell*, **94**, 439–449.
Yamada,R. and Bradshaw,R.A. (1991a) Rat liver polysome N^{α}-acetyltranferase: isolation and characterization. *Biochemistry*, **30**, 1010–1016.
Yamada,R. and Bradshaw,R.A. (1991b) Rat liver polysome N^{α}-acetyltranferase: substrate specificity. *Biochemistry*, **30**, 1017–1021.
Yamamoto,T., Moerschell,R.P., Wakem,L.P., Ferguson,D. and Sherman,F. (1992) Parameters affecting the frequencies of transformation and co-transformation with synthetic oligonucleotides in yeast. *Yeast*, **8**, 935–948.

Received August 11, 1999; revised September 10, 1999; accepted September 13, 1999