# Molecular Cloning and Sequencing of a cDNA Encoding $N^{\alpha}$ -Acetyltransferase from Saccharomyces cerevisiae\*

(Received for publication, March 9, 1989)

#### Fang-Jen S. Lee<sup>‡</sup>, Lee-Wen Lin<sup>‡</sup>, and John A. Smith<sup>‡</sup>

From the Departments of \$Molecular Biology and \$Pathology, Massachusetts General Hospital and the \$Departments of Genetics and Pathology, Harvard Medical School, Boston, Massachusetts 02114

Acetylation is the most frequently occurring chemical modification of the  $\alpha$ -NH<sub>2</sub> group of eukaryotic proteins and is catalyzed by an  $N^{\alpha}$ -acetyltransferase. Recently, a eukaryotic  $N^{\alpha}$ -acetyltransferase was purified to homogeneity from Saccharomyces cerevisiae, and its substrate specificity was partially characterized (Lee, F.-J. S., Lin L.-W., and Smith, J. A. (1988) J. Biol. Chem. 263, 14948-14955). This article describes the cloning from a yeast  $\lambda gt11$  cDNA library and sequencing of a full length cDNA encoding yeast  $N^{\alpha}$ acetyltransferase. DNA blot hybridizations of genomic and chromosomal DNA reveal that the gene (so-called AAA1, amino-terminal,  $\alpha$ -amino, acetyltransferase) is present as a single copy located on chromosome IV. The use of this cDNA will allow the molecular details of the role of  $N^{\alpha}$ -acetylation in the sorting and degradation of eukaryotic proteins to be determined.

### Protein Sequence Analyses of Na-acetyltransferase

 $N^{\alpha}$ -acetyltransferase was purified from yeast as previously described (24).  $N^{\alpha}$ -acetyltransferase (3 nmoles) was reduced and alkylated, precipitated with cold chloroform/ methanol, redissolved in 0.1 M NH<sub>4</sub>HCO<sub>3</sub>, incubated with TPCK-treated trypsin (EC 3.4.21.4; Copper Biomedical, Malvern, PA) (120 pmol) for 24 hr at 37°C, recovered by lyophilization, and dissolved in 6 M guanidine hydrochloride in 0.1% CF<sub>3</sub>COOH prior to HPLC.

Tryptic peptides were separated on a Vydac phenyl (0.46 x 25 cm) HPLC column, and selected fractions were rechromatographed isocratically once or twice (25). Chromatographic peaks were detected at 214 and 280 nm, collected manually, and lyophilized. The tryptic peptides were sequenced by automated Edman degradation performed with an Applied Biosystems 470Å Protein Sequencer and an Applied Biosystems 120 Pth Analyzer (26).

## Construction and Screening of cDNA Library.

Yeast RNA was isolated as described by Sherman et al. (27).  $Poly(A)^+$  RNA was selected on oligo(dT)cellulose (28). cDNA was synthesized from 10 µg of poly(A)<sup>+</sup>RNA by the method of Okayama and Berg (29), as modified by Gubler and Hoffman (30), except that 10% of second strand was [<sup>32</sup>P]-labelled. The cDNA was prepared for ligation to  $\lambda$ gt11 arms using a method described by Aruffo and Seed (31). After the ends of the cDNA were made blunt with T4 DNA polymerase, the cDNA was ligated to adaptors consisting of two oligonucleotides: 3' CTCTAAAG 5' and 5' ACACGAGATTTC 3'. This cDNA was fractionated on a 5 to 20% linear KOAc gradient (5 ml) using a Beckman SW55 rotor centrifuged for 3 hr at 50,000 rpm at 22°C. Fractions (0.5 ml) were collected from the bottom of the tube. The cDNA was precipitated by addition of ethanol and linear polyacryamide (20 µg/ml). The size of the cDNAs in each fraction was determined on a 1% agarose gel , and the fractions containing cDNAs between 1 and 8 kb were pooled. Ten micrograms of  $\lambda$ gt11 DNA (32) was digested with EcoR1, ligated to adaptors (3' GTGTGACCAGATCTCTTAA 5' and 5' CTGGTCTAGAG 3') and precipitated with PEG8000. 600 ng of  $\lambda$ gt11 DNA bearing adaptors was ligated to 150 ng of size-selected cDNA bearing complementary adaptors in 2 µl and packaged *in vitro* (33) (Stratagene) (Jen Sheen, personal communication). *Escherichia coli* strain Y1088 was infected with recombinant phage, and the library was amplified once. The recombinant frequency was approximately 82%. Among several peptides sequences, two peptides (peptides 27-3 and 11-3-2; Fig. 2A) were chosen for constructing two oligonucleotide probes (N1 and N2) based on most probable codon usage (34). The oligonucleotide probes were synthesized with an Applied Biosystems 380A DNA synthesizer by using the silica-based solid-phase method (35) and  $\beta$ -cyanoethyl phosphoramidite method (36). The purified oligonucleotide were isolated from the crude synthetic mixtures by PAGE and labelled to a specific activity of 2-8 x 10<sup>8</sup> cpm/µg by using ( $\gamma^{32}$ P]-ATP (New England Nuclear) and T4 polynucleotide kinase (New England Biolabs) (37).

In the initial screen, 500,000 recombinant clones in  $\lambda$ gt11 yeast cDNA library were plated on *E. coli* Y1088. Duplicate transfers of the clones were made onto nitrocellulose, and the filters were prepared for hybridization (37). Afterward, the filters were washed twice at room temperature in 6xSSC (0.15 M NaCl/15 mM sodium citrate (NaCl/Cit) containing 0.1% SDS and 0.05% NaPPi ), washed once at 5°C below the minimum t<sub>d</sub> (temperature of probe dissociation based on G/C content), and exposed on x-ray film for 1 to 2 days. Maximum and minimum t<sub>d</sub> were determined for two pools of redundant oligonucleotide probes (N3 and N4) (38).

#### DNA Sequencing and Blot Analysis.

cDNA fragments were cleaved out from recombinant  $\lambda gt11$  phage DNA by EcoRI digestion. The cDNA fragments were separated by gel electrophoresis in low melting point agarose. The correct DNA band was sliced out, the gel was melted at 65 °C, and the DNA was extracted with phenol. The purified cDNA fragments were cloned into the Bluescript plasmid (Stratagene). The complete sequence of the yeast N<sup>Q</sup>-acetyltransferase cDNA was determined by exonuclease III deletion (39), the dideoxy chain termination method of Sanger (40) modified for double-stranded sequencing by Guo et al. (41), and specific priming with synthetic oligonucleotides. All restriction enzymes were purchased from New England Biolabs. RNA and DNA markers were obtained from Bethesda Research Laboratories. Biotrans nylon membrane was from ICN. Poly(A)<sup>+</sup>RNA was analyzed by RNA hybridization (42,43). Genomic DNA was isolated from yeast (27), digested with restriction enzymes, and analyzed by DNA blot hybridization (44). The chromosome bearing the AAA1 gene was identified by hybridization of labelled cDNA with a Saccharomyces chromo-di-hybridizer (Clonetech) (i.e., a yeast chromosomal agarose gel).



Figure 1. HPLC separation of yeast N<sup> $\alpha$ </sup>-acetyltransferase tryptic peptides. Three nanomoles of purified N<sup> $\alpha$ </sup>-acetyltransferase was reduced, alkylated, digested with trypsin, and chromatographed on a 0.46 x 25 cm Vydac phenyl HPLC column with 0.1% CF<sub>3</sub>COOH in a linear gradient of 0-60% CH<sub>3</sub>CN over 2 hr. Numbers refer to the tryptic peptides, the sequences of which are shown in Fig. 2C.

XhoI

K J E G A S A S P I C C N V L G I Y M Pepside 27-3 TTC TAA CTT CCA CGA AGA CGA AGG GGT TAG ACG ACG GTG CGA AAC CCA TAG ATG TAC -5" People N1 . . .. . . .

415 P T L V E F Y I L K A R I L K H L G L H D T A

S V A A Y P S D Q D H D V F'G E Pepeide 11-3-2 3'- AGA CAA CGA CGA ATG GGT AGA CTG GTT CTG TTG CTG CAA AAG CCA CTT -5' Product N2 . . . . . . ...

c
---

$ \begin{array}{cccccccccccccccccccccccccccccccccccc$																				5'.	••••	TTT	CAG	GACC	CTA	NC GJ	AGT	1333	GC	: GGL G		T TT L	с с <b>и</b>	6.0	NA 66	IT AG	is ca	A CT	GA.1	TTG	CAG	GAT	: AGA 8		T ATC	-	TGT	***	ACG	: 011 V	T A
$ \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \\$																															-	-					•	-	-	-	•	-		-	-	-	-	-	•	-	
$ \begin{array}{c} 41  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  L  8  4  1  7  1  8  4  1  7  1  8  4  1  7  1  8  4  1  7  1  8  4  1  7  1  8  1  8  1  1  1  1  1  1$	22	ATC	TC	TM	GG A		GA	MGT	ACT	. **	6 00	CN	NG C	CA	GCA	GCT	~~~	ATI	• GC	T TI	G N		AA G		NAT.	GAC	CAG	1402	TM	. 11	T TT.	A AG	ic oc	T N	NC AN	-	C GM		: 000	GTG	GAA	GTC	: 000	TCC	: 011	TTC	ACC	***	-	GAT	
1) THE CIC LLE GLA THE CLAM CALL THE CLAM CALL CHAN AND CT HALL CHAN CHAN CHAN CHAN CHAN CHAN CHAN CHAN	1	M	s		ĸ	ĸ	•	s	т	ĸ		,	ĸ	•	*	*	ĸ	1		L	,	. 1	ĸ	£	M	D	Q	461	Y	F	L	R		• •		( I	D	x		۷		v		8	L		Ŧ	ĸ		0	
$ \begin{array}{c} 33  7  L  L  L  L  L  L  L  L  L$	91	тт	:	ເພ	<b>.</b>		TG	***	CTA	TA	C GA		G A	a c		TAC	-			+ c+				-	-	<b></b>												_	14				15-3	4-1 -			>1				
$ \begin{array}{c} \ 1 \ 1 \ 2 \ 1 \ 3 \ 3$	24	-	L			A .	L	ĸ	L	Y	E		5	ĸ	0	Ŷ	ĸ	. ~~~	c. s				L	L	0	~~~	1	1471	TCT	r GTT		1 66	TAT	T N	NG GM	C 11	ACN	C CT1	GTG	GAA	OCT	TCT	TOG	TTC	ATC	GTA	GAA	CAG	GCA	GN	
$ \begin{bmatrix} 0 \text{ TH} \text{ AM} \text{ AM}  AC OF ACT OF C AT TO G CT T A M TT A M TO TT A M A M TO A M A M A M A M A M A M A M A M A M A $															-				-	-								484				6				, r		r	v		*	5		r	1	•		8			
$ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1$	160	TTG	~		N GJ	IC G	57 A	GT (	CAC	GTT	GAT	TCC	C <b>TT</b>	GG	ст т	CTA I	AAG	GGT	CTT	GAT	TTA	TAT	TC	т ст	A G	GT (	GAG	1540	781		-					* *								~	-	-				_	
9 AL ALT GAL GOT GOT THE CHE OF GOT ANT SEC AND ALL ALL ALL GOT ALL ALL GOT ALT THE CALL ALL GOT ALL ALL GOT	47	L	x	x	5	,	3	ŝ	н	۷	Þ	s	L		A	L	ĸ	G	L	D	L	¥	s	٧		G	E	507	x	Ŷ		L						~~~~	~~~	L			1			L	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~			
<b>9 A A C C C C C C C C C C</b>																		1<-	·····		•••••				•					-		-								-	-	~	-	-	•	-	•	•	-	•	
$ \begin{bmatrix} 0 & K & 0 & K & A & K & K & K & K & K & K & K & K$	29	~	GAT	6	c 60	7 6	1 1	CC 1	TAC	GTG	GCT	***	r cc	C A	TC A	NGA 1	***	ATT	ŝ	GGC	GCT	TCJ	GC	A TO	A CO	CA J	A7C	1609	AGI	GA			C GA	A CI	-	T GC	-	T GAT	T ATC		GAA	-	: CAA	100	; c <del>11</del>	677	CGC	***	TAT		
$\frac{1}{1} = \frac{1}{1} = \frac{1}$	70	ĸ	Ð	Ď	,	• •	•	\$	¥	۷	*		*	1	1	R	ĸ	I	E	Ģ		\$		5		P	1	530	\$	D	ĸ	5	E				м	D	1	ĸ	×	N	0		L	v	R	ĸ	Y		
$ \begin{array}{c} 16 \text{ TC} \text{ CC} \text{ M} \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ V \ C \ C \ M \ M \ C \ M \ V \ C \ C \ M \ M \ C \ M \ M \ C \ M \ M$				- 25	1-2											21	1<						****																												
$ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0$		TGC	TGT	CA	T G1	ГА Т.	ra g	GT 1	ATC	TAC	ATG	AG			-		GAG	TAC	***	GAA	TCT	ATI		A TG	GT	rc J	NCG	1678	TTC	600	CT	5 AA	A AG	A 11	IC AA	c cc	T AT	r ccu	-	777	TAT	***	. CM	TTC	GAA	GAT	GAC	CAG	TTG	GAT	
$ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1$		c	с . 17			/ 1	•	G	1	T					т	ĸ		¥	ĸ	E	s	1	ĸ	*		r	т	553	L		ι	ĸ	R		г н		1		ĸ	Ŧ	¥	ĸ	Q			D	D	Q	L	Ð	
$ \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c}$			000					200	***									and				-					-					_		_			_														
$ \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1$	6			1				G	3	-			0			¥.			L			1						576		10	• т <b>л</b> а	TG	TAT	GAG			TAC	:	AGA	9000	TAT	CTG	GMG	ATG	TTA	GAA	TGG	GGA	ANG	OCA	
$ \begin{array}{c} 6 \ \text{GC} \ \text{CAT} TC ANA ANT CCT TA GTT FCC GNG ANA ANA TT TGC CTT GTT THC CTT GOT THC COT GOT GOC ANC  \begin{array}{c} 6 \ \text{GC} \ \text{C} \ \text{T} \ \text{K} \ \text{K} \ \text{K} \ \text{L} \ \text{K} \ K$				-					-	-			•		-			14-			·		20-	4				5/6	-	3	f	c				G	1		*		I	L.			L.		w	G	x	x	1
$ \begin{array}{c} 3 & G & D & F & K & A & L & V & S & K & K & Y & K & A & F & L & G & Y & K & A & F & L & G & Y & K & A & F & L & S & G & C & C & A & C & A & A & F & L & C & G & K & K & L & K & C & A & S & L & T & C & M & M & C & A & C & A & A & A & F & L & L & C & M & M & C & A & C & A & A & A & A & C & A & A$	6	<b></b>	GA?		C N	u u	NT 6	<b>ст</b> :	TTA	GTG	TCC	-				TAT 1	TGG	GAA	GCA	ττc	CTT	961	TM	c ca			MC.	1616	TAT	ACC	:			G 73	T GT	τ ma	c 604			1.	-	-		<b></b>	+++		-				
$\frac{1}{1} = \frac{1}{1} = \frac{1}$	9	G	D		,	( I			L	v	\$		x		ĸ	¥	×	E			L	G	¥	R		A	N	599	Y	T	ĸ				v				ĸ	-					Y		~			-	
5 TOE ALC TAA THE GET GATE GAC CAA CAA BAT ONE ANG CAA CAA BET ATT TAAC ANT THA THE TOT CAG 2 M T S L A V A O D V H G C A CAA ANT ANT OA CAA TH ALT ALT TOT CAG 2 M T S L A V A O D V H G C A CAA ANT ATT CC AGG AAA TT ALT ALC ANT C CAA GAC ANT TA TAT TAC ANT C CAA ANT ALT CCC ANT C CAA ANT TA ANT C CAA CCC AN ANT ALT CCC ANT C CAA ANT ANT C CAA CCC ANT ANT TA ANT C CAA CCC ANT ANT ANT C CAA CCC ANT ANT ANT C CAA CCC ANT ANT ANT C CAA CCC ANT T ANT TAC ANT C CAA ANT ANT C CAA CCC ANT TAT TAC CAA CCC ANT TAT TAC ANT C CAA ANT ANT C CAA CCC ANT ANT TAT C CAA ANT ANT C CAA CCC ANT TAT TAT C CAA ANT ANT C CAA CCC ANT ANT TAT C CAA ANT ANT C CAA CCC ANT TAT TAT C CAA ANT ANT C CAA CCC ANT ANT ANT C CAA CCC ANT TAT TAT C CAA ANT ANT C CAA CCC ANA ANT ANT C CAA CCC ANA ANT TAT C CAA CCC ANT ANT TAT C CAA CCC ANT TAT TAT C CAA CCC ANT ANT TAT ANT ANT ANT ANT C CAA CCC ANT ANT TAT ANT ANT ANT ANT ANT C CAA CCC ANT ANT TAT ANT ANT ANT ANT C CAC CCC ANT ANT TAT ANT ANT ANT ANT C CAC CCC ANT ANT TAT C CAA CCC ANT TAT TAT C CAA CCC					>	1										<				28-	2			>	I I							9-2	2-2 -			>				-				14-				0-3-			
2  N  7  S  L  A   V   A  0  D   V   H   G   E   R  0  G   A   I   H   T  L  S  0  A   I   H   T  L  S  0  A   I   H   T  L  S  0  A   I   H   T  L  S  0  A   I   H   T  L  S  0  A   I   H   T  L  S  0  A   I   H   T  L  S  0  A   I   H   T   L   S  0  A   I   H   T   L   S  0  A   I   H   T   L   S  0  A   I   H   T   A   A   S   S   C   L   H   H   S  0  A   A	5	rgç	ACA	1 7C	A 71	rg gi	CT 4	TG (	GCA	CAA	GAT	GT	-	C G	GT G	ING I	AGG	CAA	CAA	GCT	ATT	***	AC	T TT	A 70	CT (	CAG	1885	CGC	TTA		AG		6 TC	ж 'са	T TC	T TT	-	GAA	-	TCA	GAT	GAA	ATC	CAA	AAT	-	000	CAA	ANT	
$ \frac{1}{1} $ $ 1$	2	w	Ť	s	I		•	v	A	Q	D	۷	×		G	E	R	Q	Q		I	н	τ	L		5	Q	622	R	L	ĸ	R	ĸ	5		s	L	D	E	N	s	D		I	Q		N	G	0		
r  tr  cont																													>1																						
$ \begin{array}{c} 3  AA ABC GAC ATT AT GAA GCT GC ATT GAT GAT GAA GAT ACC AAG CAA GAT ACT TH C GA ABC TAG ABA TA TH GG AGA ABA GCAA GAT ATT GAA GT ACT ACT ABA GCT TO C ATT GAA GAT ACC ATT ACC ABG CAT AGG GAT CAA GAT AAC GAT GAT AAC GAT AA$	•	111	GAA		A C1	C G	<b>T</b> 0	AG I	GCA	***	ATA	TCT	T GA	TT	CC Q	ING I	***	TAT	CAN	CAC	AGC	GAG	TG	7 77	N N	7G 1	TAC	1954	AGC	AGC	CN		G 🗛	* **	a ec	T AN	G AA.	G	GCA	GCC	CCT	ATG	ANC	***	COG	***	GAA	ACT	GAA	900	: N
3 AM AMC GAC ATT ATG ATT ATG ATG CAT GAC ANG THA CAM ANG THA CAM MAT GTA TTG AMA CAT TTG 8 K H D I H Y K A A S D H Q D K L Q H Y L K H L 2 AMT GAT ATC GAG CTC TTT GAT AMA TTT GAC ANG TTA CAM AMA GTA ATT ATG ATG ATG ATT TTG AMA CAT TTG 2 AMT GAT ATC GAG CCA TCC GTC TTT GAT AMA TTT GAG AGA AMA GCA ATA ATT TA GAG AGA AMA CAT ATT TA GAG AGA CAT ATT TA GAG AGA AMA CAT ATT TA GAG AGA CAG ATT ATT TA GAG AGA CAG ATT ATT TA GAG AGA GAG CAG ATT GAT ATT ATG GAG GAC TTT GAT ATT TA GAG AGA TTT GGA AGT AGT	5	7	٤	x	7		•	z	5	ĸ	I	\$	Ð	, ,	\$	E	ĸ	¥	z	Ħ	\$	I	c	L	• •		Y	645	5	\$	Q	x	ĸ	×	C A	ĸ	x	E	λ	*	Α.	м	M	ĸ	R	x	E	Ŧ			
J ADA COL AN AN AN HI AN COL AN AN ANT THE COL AND										~~	~~~						~~~		_	<i></i>		~	_				_																								
$ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1$		~	~~~					~	-	~	~	~		~~~		~		~~~		~~~~	~			~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~				2023	861	411	GCI		TIN	1 00	GNG	TGA	r CAU	GAT	AAC	GAT	GTA	TTC		GAA	ANG	77G	ATT	GAA	ACC	TOC	: M
2 ANT GAT ATC GAG CCA TOC GTC THT GAT ANA TTT GGT THA TTA GAG AGA ANA GCA ACT ATT TA ACT ATC ACC GAA TT TAA ANA TTA GG AGA ANA TTA GGA AGA GAA AA T I T H K 20 ANT GAT ATC GAG GCC TTC GTT GCT THT GAT ANA TTA GGA AGA AAA CAC GAA ATT TA ACT ACC ATT ATT TAG GAC GAC TTC GTT CCC ATT GTT TAT AGA ACA GAA AGA GAC GAC AAT 20 A C C A ATG GAA GGC C ATT GTT TAT AGA ACT CTG ATC CAC ATT ATT TAC AGA ATT TA AGA CAC GAA ATT TTA AGA GAC GAC ATT CCC ATT GTT TAT AGA ACA GGA GAC ACC GAC ATT GTT TAT ATT THC GAC TTT GAC ATT ACT AGA ATT GCA AGA GTT GTT TAT TAG GAC GAC ATT GCA ATT GCA AGA GAC GAC ATT GCA ATT ATT THC GAC TTT GAC ATT GCA AGA ATT GCA AGA GAC GAC ATT GCA ATT ATT THC GAC TTT GAC ATT GCA AGA ATT GCA AGA GAC GAC ATT GCT TA GCA ATT ATT TTG GAC GCA ATT ATT ATT GCA ATT GCA ATT GCA ATT ATT ATT GCA GCA ATT TTC GTA AGA ATT GCA TTG GAA GCA GAC ATT TTC TAC GAA ATT GCA TTT ATT AGA GAA ATT GCA TTT CTT ACC ATT AGC ATT ATT TTG GAC GCA ATT GCA ATT GCA ATT GCA ATT GCA ATT GCA GCA ATT TTC GAA ATT GCA TTT GCA ATT GCA GCA ATT TTC GAA ATT GCA TTT GCA GCA ATT TTC GAA ATT GCA TTT GCA ATT GCA GCA ATT TTC GAA ATT GCA TTT GCA ATT ATT GGA ACC GAA ATT GCC TTT CCA ATT ATT GGA GCA ATT GCA TTT GGA GCA ATT GCA TTT GGA GCA ATT GCA TTT GGA GCA ATT ATT TTG GAA GCA AAT ATT GCA CCA ATT TTG GAA GCA AAT GAT GCA ATT ATT ATT GCA CCA ATT ATT GGA GCA ATT TTG GAA GCA AAT ATT GCA ATT ATT GGA GCA ATT ATT GCA GCA ATT GCA ATT ATT GCA GCA ATT GCA ATT AT	•		-	•				•	•		-					*	2		-	۳	"	•	-						14	•	····	<b>^</b>					1-1-						6		ĸ	L	r	E	Ŧ	3	
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	2	лат	GAT	AT	C GA	G CO	A 1	GC (	GTC	777	GAT	~~~	N 77	7 64	GT T	-	TTA	GAG	AGA	***	GCA	ACT	AT	T TA	C A1	rg j	~~	2092	CCA	ATG	GA	. GA	C TT	c cc	T AC	C GN		-	-	MC	TAC	TOC	ATG	C24	670	ACA	(1)				
BI THG GOT CAA THG AAA GAC GCG TCC ATT GTT TAT ACA ACT CTG ATC AMG AGA AAT CCA GAT AAT THT ANG S1 L G Q L K O A S I V Y R T L I K K N P D N F K S0 TAC TAC AMA THG ANA GAC ACG TTG GAA ATT CCA ATT AAA AAA TCCA GAT AAT THT ANG S0 TAC TAC AMA TA CTA CA AS A TC CAA GAT ATT CCA TA AAA TTG AAG AMG GCT TTG TAT GAA 77 T Y K L L E V S L G I Q G D N K L K A L Y G 19 AMA CTT GAA CAA TTT AC CAA ACA ATT GCA ACG ATT AAA ATT CCA AGA ATT CCA ATT ACT ATT CCA TTA ACA ATT GCA ACG ATT ATT THG GAC CTT AAC ATT ACC ATT ACC ATT AGC ATT ATT THG GAC CTT AAC ATT ACC	1	N	D	I	Ε		•	c	v	F	D	ĸ			G	L	L	8	R	ĸ	٨	Ŧ	I	Y		H	ĸ	691	₽	н	E	D	F		Ť			¥	N	N	¥	8		~	v	-	-	-			
<b>SI</b> TH G GOT CAR THE ANA GAC CCT THE CALL AND ACT THE AND ACT THE CAR AND ATT CCA ATT ATT ATT ATT CCA ATT ATT THE CAR AND ATT CCA ATT ATT ATT CCA ATT ATT ATT THE CAR AND ATT CCA ATT ATT ATT THE CAR AND ATT CCA ATT ATT ATT ATT CCA ATT ATT ATT																																35.										-		•			-				
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	81	TT	5 66	T C		TG .	~	GAC	GCO	TO	-	<b>G</b>	<b>TT T</b>	TAT	AGA	ACT	r cre	-		GNG	а л	то	CA G	-	AAT	777	ANG	2161	-		-	: GM	c	T GA								***		-	-						
30TAC TAC AMA THA CTO GAA GTA TCC THE GGA ATC CAA GGT GAC AATA AMA THE AMG AMG GCT THE TA COT AMG AGA GTA THE CTO GAA GAC ACT ATG AGA AGA GC CTT TGA TA CAA ATT THE CTA AMA AAA THE AMG AMA GC CTA AGA ATT ATT ACT THE CTA CAA AMA AAA THE AMA AMA AGA ATT GTA CCAA ATT ATT ACT ATT CCAA THE CTA CTAA ATT ATT ACT THE CTA CTA ATT ACT ATT CCAA THE CTA THE CTA CTAA ATC ATT ACT ATT CCAA THE CTA THE CTA CTAA ATT ATT ACT THE CTA CTA ATT ATT ACT THE CTA CTA ATT ATT ACT THE CTA CTA ATC ATT ATT ACT THE CTA CTA ATT ATT ACT THE CTA CTA ATT ATT ACT THE CTA CTA ATT ATT ACT ATT CCAA ATT ATT ACT ATT CCAA ATT ATT	54	L	G		0	٤	ĸ	٥	*	s	. 1		v	۲	8	τ	L	1	ĸ			•	P	Ď	N	F	×	714	Y	1	ĩ	D	· · ·				m		1	6	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	1	~	L	100		UCT			~~~	~
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$																																		-			-		-	-		-	-	-	•	•	^		-		
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	50	TA	C TA	C A	**	TA (	CTG	GAA	GT/	10	с <b>тт</b>	rg gi	<b>GN N</b>	ATC	CAA	GGT	C K	- **	. **	A 77	G N	G A	NG C	27	TTG	TAT	GGA	2230	770	GCT	-	AG		T GG	C AC	c	G NGC	GGT	114	777	GGT	AGT	ATG	GCC	AT7	<b>677</b>	TTG	TTA	CAT	BCC.	
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	77	Y	Y	r I	ĸ	L	L	E	v	s	. 1		G	I	Q	G	D	N	ĸ	ı		¢ I	ĸ	x	L	Y	G	737	r	A	ĸ	8	F	G	• •	т	s	Ģ	L	F	G	s	м	•	' I	۷	L	L	×	٨	
19 MAR CIT GAR CAR TOC ANA TOC ATT COL MAR TOC ATT CAN THA ACT THE CIT CAN GAR ANA CHT CAN CHA ANA CHAC CAC ACT COL ATT CAR TOC ANA CHAC ANA CHAC CAC ACT COL ATT CAR CAN ATT COL ANA TOC ATT CAN ANA CHAC CHAC ANA CHAC CAC ACT COL ATT CAR CAN ATT COL ANA TOC ATT CAN ANA CHAC CHAC ANA CHAC CAC ACT COL ATT CAR CAN ATT COL ANA TOC ATT CAN ANA CHAC CHAC ANA CHAC CHAC ACC CCC TTT CAR CAN ATT COL ANA TOC ATT CAN THAC CHA ATT CAN THAC CHA ANA CHAC CHAC CHAC CHAC CHAC ANA CHAC CHAC																																																			
$ \begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 $	19	~	A CT	T G	AA (		111	TAT	CC	AG	A 10	CG	•••	CA .	ccc	~~~	. 171	T AT	r cc	A 17	A A(		TC C	, TT (	~	GAC		2299	AGA	AAC	GAG	. NCO	c cc	ст	T GA	T CC		TTG	AAG	***	GTA	GTC	ACG	AAG	AGC	CTT	GAA	***	GAG	TAT	т
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	00	ĸ	L					1	۳ 	R	, 0			P	r	×		1	P				2.1.		9-1			760	R	ж	D	Ŧ	P	,	D	P	1	L	ĸ	ĸ	v	۷	τ	ĸ	\$	L	E	ĸ	8	Y	5
$ \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} $		Ca.		6.0	TC .				TTO				AT 0		TTG	CCT	1 CAU		. GA	6 00	ic of	T G		CA (	GCA	ACT	-	2360	~		-													_							
$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 $	23	E			L	5	×	ĸ	L	R			Y V	v	L	P	0	L	E				v	P		T	F	2368			110		TI TI	~ ^^	C GA	A AT/	701	AAC	AAT	AGC	TTC	GAT	TGG	CTG	AAT	TTC	TAC	CAA	GAA	***	T
57 TCC AAC GTG AAA CCC CTT TAC CAA MGA AGG AAG GTT TCA CCA CTA TTG GAG AAA ATT GTC CTT       2437 GGT AAG AAT GAT ATA AAT GGC CTG CTA TTT CTG TAT CGC TAT CGC GAT GAT GAT GTT CCG AAT GAT GTT ATT AGG AGG CTG CAT TTG GAG GC CAA GAT GAT ATT GG AGG CCT CAC GAC GAA ATT GT CTG GAT GAT GTT GT GC GAT GAT GAT GTT GT GCG GAT GAT GAT GTT GTG GAT GAT GTT ATT AGG CCG GAT GAT GAT GTT ATT AGG AGG CCT CAC GAC GAA ATT GT CTG GAT GAT GAT GAT ATT GT GC GAT							>1		-			<			1	15-1-	-2 -			>	1 1	<								-	•		-			1	3			•	,	D	•	L	N	,	¥	٩	z	ĸ	
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	57	TCO	:	c G	TG A		cec	CTT	TAC	: ca	A AG		GG A	MG -	TCC	AAG	GT	тса	CC	A CT	A 11	IG G	AG A		ATT	GTC	ctt	2437	GGT	AAG		GAT			TGG	c c70	5 CT-		CTG	TAT	car	TA7	566	647	G	-					
15-3-1       15-3-1	46	5	N		v	ĸ	P	L	Y	Q	) P	۱ ۱	R	ĸ	s	ĸ	v	s	P		. 1		3	ĸ	I	v	L	806	G	ĸ	N	D	1	N		L	L		L	Y	8	¥.	8	-	JAI D	v11		ATC	GA	AGC	TC.
26 GAT TAT THE TEC GGA THA GAT CET ACE CAG GAT CEA ATT CET TH THE GA ACE AAT TAT THE TEC THAT THE TEC GAA GAT CEA ATT CET ATT GAT CET CAG GAT CAAT TAT THE CET THE CET AGE GAT CEA ATT CET ATT CET THE CET AGE GAT CEA ATT CET ATT CET THE CET AGE GAT CEA ATT CET ATT CET AGE GAT CEA ATT CE CAG GAT CEA ATT CEA GAT CEC CAG GAT CEA THAT THE GAT CET CEC CEA GAT CAT THE ATT CET THE CET THE CET AGE GAT CEA ATT CET ATT CET THE CET AGE GAT CEA ATT CEA GAT CEC CAG GAT CAT THE ATT ATT THAT CEA GAT CEC CAG GAT CAT THE CET THE CET THE CET AGE GAT CEA ATT CEA GAT CEC CAG GAT CEC CAG CAA CEC CAG AAC CEA ATT CE CET THE CET THE CET THE CET AGE GAT CEC AAT TAT THE GAT CET CEC CET CAC CEA CAC CEA CEC CAG CAA CEC CAG CAC CEC CA				15-	3-1						>	1																							-				_	_		-		-	-	•	·	·		3	•
69 DYLSGLDPTODPTFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	26	GA	1 18	T T	TG 1	rcc i	GGA	TTA	GA	r cc	TAC	G C	AG G	GAT :	CCA	ATT	1 001	1 11	T AT	7 70	GN	CC A	AT 1	TAT '	TAC	ŤŤG	TCT	2506	AAT	176	~~	GN		C AT	T AT	T 860	AG7	CTT	TCT	ccc	TTG	GAG	сст	CAC	TCC	CAG	AAC	GAA		CTA	
95 CAA CAT TTC CTT ANG GAT TTT CCG ANA GCC CAA GAA TAT ATT GAT GCT GCC CTT GAC CAC ACC 2535 TAT TAC TTG TMG CCTGCAACTCCTCAATGAGTCTCAATGAGTCTTAATGTATAGTATAGTATGTATGTATAGTATGTATGTATGTATGTATATGTATATGTATATGTATATGTATAGTATGTATGTATATGTATAGTATGTATGTATAGTATGTATGTATGTATAGTATGTATGTATGGTATATGTATGTATGTATGTATATGTATAGTATGTATGTATGTATGTATGTATAGTATGGTATGTATGGTATGGTATGTATGGTATGTATGGTATGGTATGTATGGTATGGTATGGTATGGTATGGTAGTA	69	D	Ŷ		L	s	G	L	D	P			0	D	P	I	Р	F	1	•		r	N	¥	Y	L	s	829	N	L	ĸ	ε	M	I	1	s	s	L	s	P	L	ε	P	в	s	0	N	t	I	1	6
								_												-	• •																														
92 O H F L F L F D F F F A O E Y I D A A L D M T - 852 Y Y F +	95	cN	CA	тт	TC C	TT	TTC	CTT	***	GA	T TT	TT C	CG A		GCC	cm.	ິ	N TA	T AT	τG	TGO	TG	cc c	. 11	GAC	CAC	: ACC	2575	TAT	TAC	TTO	TAC	CC1	TGCA	ACTO	CTCN	TGTO	TCAA	TTAN	TCT	TACT	TAAT	TATO	TAT	ATATI		ATGT	-	IGC TI	TATA'	TGC
ra y n su	92	Q	н		F	L	F	L	ĸ	D				ĸ	•	Q	E	Y	1	c	, ,	•	^	-	D	н	4	852	¥	Y	L	•																			
															~~~					c		~ ~		TC -	CAC	***	crc																								

BetXI

MindIII HindIII BatXI

0.2 Kb

EcoBV.

EcoRV SacI BanHI

- 1

FIG. 2. Cloning and sequencing of the cDNA encoding yeast  $N^{\alpha}$ -acetyltransferase. A, oligonucleotide probes used for initially screening the  $\lambda$ gt11 library. The amino acid sequences of two tryptic peptides were used to construct the codon-usage frequency based oligonucleotide probes. The nucleotide positions indicated by the asterisks differ from the actual DNA sequence shown in C. The numbering of the tryptic peptides is as follows: the first number refers to the corresponding peak in Fig. 1, the second number refers to the peak in the first isocratic HPLC separation (data not shown), and the third number refers to the peak in the second isocratic HPLC separation (data not shown). B. restriction map and DNA sequencing strategy for the cDNA clones. The arrows indicate the direction and extent of sequence determination for each fragment after exonuclease III deletion. C. nucleotide and deduced amino acid sequence of  $N^{\alpha}$ -acetyltransferase cDNA clones. The amino acid sequences of HPLC-purified tryptic peptides determined by automated protein sequence analysis are also shown. The protein sequence analyses were completed with repetitive yields between 87 and 93% for 100-200 pmol of each peptide.

FIG. 2. Cloning and sequencing of the cDNA encoding yeast  $N^{\alpha}$ -acetyltransferase. A, oligonucleotide probes used for initially screening the  $\lambda$ gt11 library. The amino acid sequences of two tryptic peptides were used to construct the codon-usage frequency based oligonucleotide probes. The nucleotide positions indicated by the *asterisks* differ from the actual DNA sequence shown in C. The numbering of the tryptic peptides is as follows: the *first number* refers to the corresponding peak in Fig. 1, the second number refers to the peak in the first isocratic HPLC separation (data not shown), and the *third number* refers to the peak in the second isocratic HPLC separation (data not shown). B, restriction map and DNA sequencing strategy for the cDNA clones. The *arrows* indicate the direction and extent of sequence determination for each fragment after exonuclease III deletion. C, nucleotide and deduced amino acid sequence of  $N^{\alpha}$ -acetyltransferase cDNA clones. The amino acid sequences of HPLC-purified tryptic peptides determined by automated protein sequence analysis are also shown. The protein sequence analyses were completed with repetitive yields between 87 and 93% for 100-200 pmol of each peptide. A

Sequence data GI Y N EGASASPIC v L Pepide 27-3 С H K I 31- TTO TAN OTT CON OGN NGN CGN AGG GGT TAG ACG ACG GTG CGN NAC CON TAG ATG TAC -51 Probe N1 \* ٠ \*\* ٠ . . . ٠ • G E v F D Pepdde 11-3-2 S D Q D N P S v λ A Y 3'- AGA CAA CGA CGA ATG GGT AGA CTG GTT CTG TTG CTG CAA AAG CCA CTT -5' Probe N2 . . . . . ٠ ٠ ۰. ۰ .



																	5'	1	TTC	AGG	ACCC	TAAC	GAN	π	133	3 0	1 G			•G G				~ ~					6 61	7 10	-	17	<b>.</b> .				~	-	
																									43		1 0	G	1	L	6	E	6	R	0	L	<b>D</b> 1			, ,			1	*	c	ĸ	•	v	T
22	ATG	тст	AGG	***	AGA	NGT	ACT	ANG	ccc	ANG	CCA	GCA	GCT		<b>A TA</b>	GCT	770			A GA		тGJ	AC C	NG	140	2 71		** *		~ ~				-			~ ~			~ ~	~ -	~ ~		-	~ .				
1	M	s	R	ĸ	R	s	Ŧ	ĸ	P	ĸ	P	A	A	ĸ	1	A	L	ĸ	ĸ	E			5	9	46	1 1		r 1	L	R .	A .	8 I	AT A N	I	b i	K J	N 1	/ 16	1 U	. u	1 1	5 L	ur r Li i	T	сс <i>і</i> Т	K K	an a	0	D
•••			~~~				-																													1	<			- 15	-1-1			;	>1				
91	TTC	CTC	GAG	GCA	TIG	~~~	CTA	TAC	GAA	GGG	AAG	CAA	TAC	:		TC1	· CTC		; c <b>t</b>	GCT	X GI	C GO	CA A	TT .	147	1 TO	TG	FT A	AT G	GT A	TT A	AG G	AC T	TA C	AC C	TT G	IG G	<b>v</b> 00	TT	T TG	IG 11	IC AT	ic c	TA G	**	CAG O	X.A	GAA	9000
~	•	2	Ŀ	î	2	•	2	•	-	•	ĥ	v	I	•		3		ĸ	Ľ	Ľ			•	Ľ	48	4 5			N (	6	1	ĸ	D	L	В	L 1	1			T		r 1	1 1	v	E	9	*	8	A
160	TTG	***	***	GAC	GGT J	AGT (	CAC	GTT (	GAT	TCC -	TTG	GCT	TTA	AAG	GGT	CTT	GAT	TTA	TAT	тст	GTA	GGT	GA	•	154	ο τι	T T/	AT A	GAC	TA T.	AC C	TG G	AT A	GA A		AG A	**		c da		<b>a</b> a		<u>с</u>	** *			28.0	<b>677</b>	-
47	L	x	ĸ	D	G	ŝ	н	v	Þ	S	L	*	L	к	G	L	D	L	¥	s	v	Ģ	E		50	7 1	1 3	¥	R	L	¥	L	D	R	ĸ	ĸ	ĸ	6 0	) (	1	. ,	<b>\</b> 5	<u>،</u>	L	ĸ	ĸ	1	v	
		~~~	~~~			<b>.</b>									1<	~~~~																																	
70	~		wal i		*	8	,	. ore		~		1	-	~~~	<b>X</b> 1-1			GCT .	7CA	uc.	TCA E	LLA D		•	160	9 M	TG	AT A	** *	GC G		** *	TT G	CG A	AT G	AT A	TC N	N GA	~ ~	6 63	A TO	56 C1	11 G	<b>TT</b> C	GC /		TAT	***	GGT
			25-2	·			·	·	·				->1	1<-											53	0 5			ĸ	S I		Q .	I .		M	D	I I				1 1	f 1		Υ	R	ĸ	Y	ĸ	G
298	TGC	TGT	CAT	GTA	па с	GGT J	ATC	TAC	ATG .	AGA	AAC	ACC		GAG	TAC	~~	GAA	TCT	ATT	***	TGG	TTC	ACC		167		~ ~		~ .		~ +	-		~													_		
93	с	с	8	v	L	G	I	¥	н	R	M	T	ĸ	C	Y	ĸ	E	s	I	ĸ	W	F	т		55	3 L				~ ~					1 0						<u> </u>			ATG	ALC C			GAT	TTC -
		27-	)						·>1																			-				• •		~	•		•	•				-			•		-		Ŧ
367	GCA	GCT	TTG A	AAC I	NAT O	GGG 1	FCC .	ACT I	AAC	ANG (	CAA	ATA	TAT	AGA	GAC	TTA	GCA	ACT	TTG	CAA	TCA	CAA	AT		174	7 CA	TTO	ж тл	NC TO	-	TG A	GA N	AA G	GT N	cc c		SA GO	х та	т ст	G	-G AT	G TT	A G	-	66 0		UNG (		CTT
116	A		L	N	M	G	5	Ŧ	M	ĸ	9	I	Y	R	D	L	A	τ	L	Q	S	۵	1		57	6 1	5		r (		H	R I	ĸ	6	T			Y	1			1 1		8		G	x		L
															<~-					20-4																	I	<			;	27-1					>1		1<
436	GGC	GAT	77C /		UAT C	GCT 1	TTA (	STG 1	rcc	NGG	***	***	TAT	TGC	GAN	GCA	TTC	CTT	007	TAC	CGT	900	***		161	6 ТА	TAC	n n	NA CI	C A1	TG T	AT G	<b>TT</b> C	oc o	-	TG AJ	le cu	A GC	A TO	X XX	G CT	-	C T	77 C	***	TG C	TAT	GAT	GAT
139	G	D	r	ĸ	N	*	L	v	\$	R	ĸ	ĸ	¥	W	E	*		L	G	¥	R	A	N		59	9 Y	7	r 1	K I		M .	¥ 1	<b>v</b> 1	R J	N 1	( I	c 1		. 5	ĸ	L	. ч	/ <b>F</b>	F (	Q	N	Ħ	D	Ð
				•>।		-	~			<b>686</b>		007	>	200		~~~	28*. 				->								- 9-	2-2			>	•1					_		1.	<			- 10	-3-1			
162	100	*	E .	1.		v 914	~	~		. uto	~~~	6		-	~~~	~	GCT .	TT		<b>7</b>	114	7C1		•	160	5 66	ст	(A A)	~ ~	5A A	AG T	യ	AT T	CT T	TA G	NT GJ	u n	TTC	X GA	TGA	A AT	C CA	עג	AT A	AT C	C C		AAT	AGT
	-	•		-	~	•	~	•	•	•	-	•	•	ſ	•	•	•	•	-	•		2	•		62	'〔					K :	5 1		S I			•	S	D	F	1	Q	1 9		M	G	Q	ĸ	\$
574	TTT	GAA		CTC	ст (	GAG	GGA		ATA	TCT (	GAT	rcc	GAG	***	TAT	GNA	CAC	AGC	GAG	ŤGŤ	TTA	ATG	TA		195		с <b>м</b> а	r c					-					~ ~	*		~	. ~	~ •			-			
185	7	Ł	ĸ	r	x	E	6	ĸ	T	\$	D	\$	E	ĸ	¥	t	11	8	E	с	L	H	Y		64	5 8	3	5 0		~ ~	K 1	к 7		к 7	ж. —			~ ~			~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~			*	~ ^	-		. 200	ANG
																													-																-	•	-	•	•
643	~~~	AAC	GAC /	NTT I	NTG 1	TAT /	~~	GCT (	ecc .	AGT	GAT	AAC	CAA	GAC	ANG	TTA	CAA	NAT	GTA	TTG	***	CAT	TTO		202	3 AG	T GT	T G	CT GO	T T	AT C	CG AC	T G	AT C	<b>M</b> (1)	NT N	c ca	T GT	A TT	c 😔	CGA	* **	6 T	TG A	TT G	***	<b>cc</b> :	TOC	ACT
208	ĸ	N	Ð	I	м	Y	ĸ	*		5	D	N	õ	D	ĸ	L	0	N	¥	L	ĸ	и	L		66	B S	v	, I		• •	Y I	P 5	s I	D (		<b>,</b>		v		G	E	K	. 1	L	I	E	T	5	Ŧ
		~		-	~		-											~~.							200	1								11-3	-2 -						>	1		<					
231	~~:	0	<b>1</b>	E	P	с .	v		D	ĸ	÷.	<u> </u>	t.	L.		8	×~~	*		7,7	Y		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	•	209		^ ^î												стс		- CA	AGT	с <b>м</b> с		AA G	AC G		NGG	GAT
		•	-	2	-	•	•	•	-		•	-	-	-	-	•			•	•	•				••	•••		•••		· ·				- 1							•	•		ĸ	5	D	E	R	D
781	TTO	GGT	CAA	TTG	***	GAC	GCG	TCC	ATT	GTT	TA1		ACT	T CT	G ATC	:	. NGI		T CC	A GA	TN	.т. т.	TT 8	AG	216				- 35	-2,	39-1						•••••			>	1								
254	Ł	6	0	L	×	0	*	s	1	v	Ţ	8	τ.	L	1	ĸ	,		 P	D			F	ĸ	210		T AT		rg GJ		rr G.			AC TA		5A A1	TT GO		G TT	A GC	TTT	G TG	с ті	TT G	ст т	CT C	TA I	NAC 1	***
			-																							• •	•			, <b>,</b>									. 1	^	L	. с	,		A	S	Ł	N	x
850	τM	TAC	***	TTA	CTG	GAA	GTA	tec	TTG	GCA	ATC	CN	GG	GM			TTC		G AA	GGC	1 11	GT	AT G	G <b>A</b>	223		c cc	TN	G AG	-	<b>FT</b> G6	GC AC		cg ស	sc Ga	it 11	-	T GG	T AG	-	e ec	-	7 9	<del>,,</del> ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	TG T	TA C	AT (	ecc .	ACA
277	Y	Y	ĸ	L	L	E	v	S	L	G	I	Q	G	Þ	N	ĸ	L	ĸ	ĸ	*	. 1	- 1	Y	G	73	7 .		1	<b>c</b> 1	. ,	F (	G 1	1	т	5 (	; 1	. ,	G	S	м	A	í I	,	V I	L	L	я	A	Ŧ
919		CTT		-	***	-			TOO	. 688				а тт:						C CT	тс		AC A																										
300	×	L.	E	0	F	Y	P		с.			 P	ĸ	F	1	P	<i></i>	Ť	F	L		) 1	D	ĸ	229	7 AG	~ ^^	ыс қа ,		e ee		rr GJ	NT CO		rr 17		IG AJ	A GT	A GT	C AC	5 AA	GAG	c ci	TT G	** *	AA G	AG :	TAT	TCT
		14-			9-1-	-1		>1	-	-	+	-		i<					;	2-1-2	, 29	-1 -						•••	· ·									v	•	τ	ĸ	\$		L 1	2	ĸ	£	Y	s
988	GN	GAG	CTC	AGC	***	***	τTG	AGA	GAA	-	GTT	1 110	; cc1	r cau	-	GAG	c c c c c	GG	T GT	t cc	A GO		ст т	TT	236	B GA		T T	c co	A 11	-	LC GJ		TA 70	7 4		-	с <b>т</b> т	c 64		c	~							
323	E	E	L	\$	ĸ	ĸ	L	R	E	¥	v	L	P	Q	L	E	Ŗ	G	۷	P			т	F	78	3 6	N						5 1	1 1	5 1		5		n J		1				NC C	<b>AA</b> G			TTC
					->1				1<				15-1	-2 -		•	>1	1<			••••	•		• •								-			-	-			-	-	-		,		•	¥	•	*	'
1057	TCC	AAC	GTG	***	ccc	CTT	TAC	CAA	AGA	AGG	. ***	TCC	ANC	GT	T TC	CCI	1 CT	TT	G GA	G 🗚	A A1	T G	rc c	TT	243	7 GG	T AA	GN	T GJ	-	-	NT GG	ic cr	rg ct	-	T CT	G TA	T CG	C TA	t co	C GA	T GA	.T G1	rt co	CG A	TC G	GA J	NGC -	TC 7
346	5	N	v	ĸ	P	L	¥	Q	R	R	ĸ	Ş	K	v	S	P	L	L	Ε	ĸ		1	v	L	80	6 G	ĸ	. )	• •	1	1 1	• •	; 1	LI	. 1	. 1	. ч		Y	R	D	D		vi	•	r	G	5	5
		;	15-3-	-1					>1															-																								-	-
1126	GAI	TAT	11G	TCC	GGA	TTA	CAT	CCT	ACG	: CAG	GAT	1 001	ν λΤ1	1 00	T TT1	· AT:	TGO	3 AC		T TA	T T/	NC T	TGT	CT e	250	6 AA	דד ד	G N	VA GJ	A A1	IG AT	וא דו	TT AC	SC AC	7 C1	7 70	T CC	C TT	G GA	cc.	L CY	C TC	c ca	AG AJ	AC G	-	77 (	CTA (	CAG
369	D	Ŷ	L	5	G	L	D	P	т	0	D	P,	1	P		1		T		1			-	3	82	9 N	Ľ	•	( E		•	r 1	1 5	5 5	5 1	, 5	P	L	E	P	к	S	0	2,	N	8	I	L	Q
1195	CAA	CAT	TTC	CTT	ттс	CTT	AAG	GAT	TTT	ccc	; ,,,,	GCC	C C AU	GA	-	TA 1	GA	r GC	T GO	c c7	ΤG	kc c	AC A	cc	257	5 TA	T TA	ст	G T	G CC	TGC	ACTO	cte	ATG	GTC	ATTA	ACTO	TTAC	****	TTA	TGTA	-			ю <b>т</b> ь	-	· · · ·		~~
													-			-								-																							~ * * *		

852 Y Y L +

392 Q H E L F L K D F P K A Q E Y I D A A L D H T

1264 CCA ACT TTA GTT GAG TTT TAC ATC CTC AAG GCA CGT ATC CTG AAG CAC TTA GGC CTA ATG GAC ACA GCG 415 P T L V E F Y I L K A R I L K H L G L H D T A

С

Figure 3. Autoradiogram of Northern blot analysis of yeast  $poly(A)^+$  RNA. Yeast  $poly(A)^+$  RNA (10 µg) was electrophoresed on a 1.2% agarose/formaldehyde gel (33). The mRNA was transferred onto a nylon membrane and hybridized with random primed, [<sup>32</sup>P]-cDNA (derived from pBN1) for 24 hr and washed (42,43). The gel lane containing the RNA markers was sliced out, visualized by staining with ethidium bromide, and used for determining the molecular size of the yeast poly(A)<sup>+</sup> RNA.



Figure 4. Autoradiogram of Southern blot analysis of restriction fragments of yeast genomic DNA. Yeast DNA (10 µg) was digested with indicated restriction enzymes. The restriction fragments were electrophoresed on a 0.8% agarose gel in Tris-borate buffer. The DNA was transferred onto a nylon membrane and hybridized with random primed, [<sup>32</sup>P]-cDNA (derived from pBN1) for 24 hr and washed (44).



Figure 5. Chromosome identification of yeast N<sup> $\alpha$ </sup>-acetyltransferase gene. An agarose gel of yeast chromosomal DNA was hybridized with random primed, [<sup>32</sup>P]-cDNA (derived from pBN1) for 24 hr and washed according to the manufacturer's recommendations. The position of the individual yeast chromosomes on the gel is indicated by the numbers.



Figure 6. Hydrophobicity plot of yeast N $\alpha$ -acetyltransferase. The plot was calculated by the method of Kyte and Doolittle (51) with a window of 9.