

S-Adenosylmethionine-dependent Methylation in *Saccharomyces cerevisiae*

IDENTIFICATION OF A NOVEL PROTEIN ARGININE METHYLTRANSFERASE*

(Received for publication, August 28, 1998, and in revised form, October 13, 1998)

Agnieszka Niewmierzycka and Steven Clarke‡

From the Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, California 90095-1569

We used sequence motifs conserved in S-adenosylmethionine-dependent methyltransferases to identify 26 putative methyltransferases from the complete genome of the yeast *Saccharomyces cerevisiae*. Seven sequences with the best matches to the methyltransferase consensus motifs were selected for further study. We prepared yeast disruption mutants of each of the genes encoding these sequences, and we found that disruption of the *YJL125c* gene is lethal, whereas disruptions of *YCR047c* and *YDR140w* lead to slow growth phenotypes. Normal growth was observed when the *YDL201w*, *YDR465c*, *YHR209w*, and *YOR240w* genes were disrupted. Initial analysis of protein methylation patterns of all mutants by amino acid analysis revealed that the *YDR465c* mutant has a defect in the methylation of the δ -nitrogen atom of arginine residues. We propose that *YDR465c* codes for the methyltransferase responsible for this recently characterized type of protein methylation, and we designate the enzyme as Rmt2 (protein arginine methyltransferase). In addition, we show that the methylation of susceptible residues in Rmt2 substrates is likely to take place on nascent polypeptide chains and that these substrates exist in the cell as fully methylated species. Interestingly, Rmt2 has 27% sequence identity over 138 amino acids to the mammalian guanidinoacetate N-methyltransferase, an enzyme responsible for methylating the δ -nitrogen of the small molecule guanidinoacetate.

The recent explosion in the number of uncharacterized genes resulting from sequencing of entire genomes presents a new challenge in biology. Although several living organisms can now be completely defined in terms of their genetic codes, understanding the functions and interactions of macromolecules encoded by the multitude of genes is in its infancy. Methods are needed to assign function to new gene products in order to come closer to understanding what makes up a living system. With this goal in mind, we wanted to assign function to a fraction of the new sequences, specifically to S-adenosylmethionine-dependent methyltransferases in the yeast *Saccharomyces cerevisiae*, for which the complete ge-

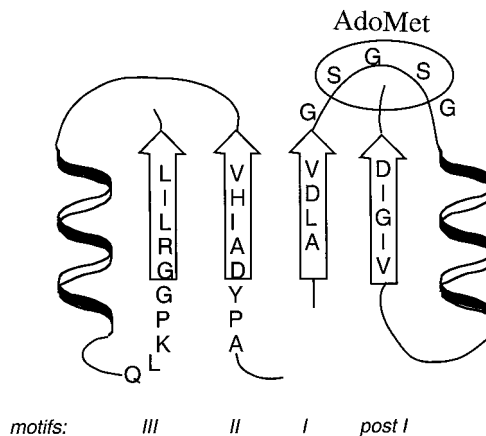


FIG. 1. Schematic representation of the conserved methyltransferase motifs in relation to S-adenosylmethionine. The conserved motifs I, post-I, II, and III are shown with the sequence of the human protein L-isoaspartate (D-aspartate) O-methyltransferase (12, 15; EC 2.1.1.77). The arrangement of motifs is predicted based on known three-dimensional structures (3–11).

nome has recently been reported (1). A large class of AdoMet¹-dependent methyltransferases have been found to share a conserved catalytic domain structure due to the interaction of the enzymes with a common cofactor, S-adenosylmethionine (2). To date, the structures of nine AdoMet-dependent methyltransferases have been solved, and they show a similar folding pattern with a central parallel β -sheet surrounded by α -helices (3–11). The common three-dimensional structure of these enzymes is reflected in sequence motifs that are conserved among a large number of AdoMet-dependent methyltransferases (12–17). The conserved regions, which we have previously designated as motifs I, post-I, II, and III (15), are always found in the same order on the polypeptide chain and are separated by comparable intervals. The three-dimensional structures of AdoMet-dependent methyltransferases have revealed that motif I and post-I interact directly with AdoMet, whereas motifs II and III interact with each other and with a portion of motif I to form the central portion of the β -sheet (Fig. 1).

By using the conserved motifs, we could theoretically identify all of the AdoMet-dependent methyltransferases in *S. cerevisiae* which belong to the major supergroup of methyltransferases that share this common structural organization. In previous studies, we have successfully predicted several open reading frames to be yeast methyltransferases based on the

* This work was supported by National Institutes of Health Grants GM26020 (to S. C.) and Medical Scientist Training Program GM08042 (to A. N.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: UCLA Molecular Biology Institute, Box 951570, Los Angeles, CA 90095-1570. Tel.: 310-825-8754; Fax: 310-825-1968; E-mail: clarke@ewald.mbi.ucla.edu.

¹ The abbreviations used are: AdoMet, S-adenosyl-L-methionine; [³H]AdoMet, S-adenosyl-L-[methyl-³H]methionine; PCR, polymerase chain reaction; ORF, open reading frame.

TABLE I
AdoMet-dependent methyltransferases in *S. cerevisiae*

Enzyme	EC number	Gene	Motifs ^a	Ref.
Gene has been identified				
mRNA cap MT ^b	2.1.1.56	<i>ABD1</i>	I, post-I, II, III	18
DHNB O-MT	2.1.1.64	<i>COQ3</i>	I, post-I, II, III	23
Coq5 C-MT		<i>COQ5</i>	I, post-I, II, III	20, 21
18S rRNA dimethyladenosine MT		<i>DIM1</i>	I, post-I, II, III	24
Δ^{24} -Sterol C-MT	2.1.1.41	<i>ERG6</i>	I, post-I, II, III	19
Protein arginine MT	2.1.1.23	<i>RMT1</i>	I, post-I, II, III	22
Probable uroporphyrin-III C-MT	2.1.1.107	<i>MET1</i>	I, post-I, II, III	25
Isoprenylcysteine O-MT	2.1.1.100	<i>STE14</i>	ND	26
Diphthamide N-MT	2.1.1.98	<i>DPH5</i>	ND	27
Phosphatidylethanolamine N-MT	2.1.1.71	<i>PEM1</i>	ND	28
Phospholipid N-MT	2.1.1.71	<i>PEM2</i>	ND	28
Mitochondrial 21S rRNA ribose MT		<i>PET56</i>	ND	29
tRNA dimethylguanine MT	2.1.1.32	<i>TRM1</i>	ND	30
Gene has not been identified				
tRNA guanine N ¹ -MT (2 activities)	2.1.1.31			31
Cytochrome <i>c</i> lysine MT	2.1.1.59			32
AdoMet: homocysteine MT	2.1.1.10			33
YL32 N-lysine MT				34
YL23 lysine MT				34
EF-1 α lysine MT				35
PP2A carboxyl-MT				36, 37
tRNA ribose MT				38
tRNA base MT				38
rRNA ribose MT				39
rRNA base MT				39
97-, 49-, 43-, 38-, 35-, 33-, 31-, 26-, 24-, 22-, and 18-kDa protein MTs				22, 40
N ^G -Monomethylarginine MT				22

^a ND, not detected. Motifs are based on consensus sequences derived by Kagan and Clarke (15).

^b MT, methyltransferase.

conserved sequence motifs. For example, YBR236c (L12000 (GB)), which we predicted to be a methyltransferase (15) was later found to be an mRNA cap methyltransferase (18). Similarly, YML008c (YYAP_Yeast (SW)) was subsequently shown to be a sterol methyltransferase (15, 19). While we were compiling a list of putative methyltransferases for the present study, one of the sequences on our list, *YML110c*, was found to code for a methyltransferase involved in the biosynthesis of ubiquinone (20, 21), and another, *YBR034c*, was shown to code for a protein arginine methyltransferase (22).

There are now a limited number of methyltransferases in *S. cerevisiae* for which the corresponding genes have been identified (Table I). About half of these methyltransferases contain identifiable sequence motifs. The enzymes lacking motifs may have three-dimensional structures distinct from the family of methyltransferases for which the crystal structures are known or may have diverged sufficiently for the motifs to become unrecognizable. In addition to the enzymes for which the gene sequence has been described, several other methyltransferases are known based on characterization of the enzyme itself or on the presence of methyl groups on a substrate molecule (Table I). Putative methyltransferase genes identified by sequence motifs can be correlated with these activities by the analysis of their disruption mutants.

In this paper, we take advantage of the completion of the yeast genome sequence to analyze directly all of the open reading frames for potential methyltransferases. We present the identification of 33 sequences that contain conserved methyltransferase motifs. Seven of the sequences represent known methyltransferases and 26 are either proteins of unknown function or proteins not previously associated with methylation. We describe the construction of yeast disruption mutants of seven genes possessing the best matches to the consensus motifs. We find that one of the yeast disruptions is lethal and two others lead to slow growth phenotypes. A fourth disruption mutant in the *YDR465c* gene shows a protein methylation defect in a recently characterized amino acid derivative where

the methyl group is present on the δ -nitrogen atom of arginine residues (41). Previously, arginine methylation was only described on the ω -guanidino nitrogens of arginine residues by enzymes such as the yeast Rmt1 methyltransferase (22, 42), although methylation of the δ -nitrogen atom on the small molecule guanidinoacetate is known to occur (43). We provide evidence that the *YDR465c* gene encodes the methyltransferase activity responsible for this modification, and we designate the gene for this novel enzyme as *RMT2* (protein arginine methyltransferase). Analyses of the other disruption strains have provided some clues to the possible function of their corresponding methyltransferases.

EXPERIMENTAL PROCEDURES

Cloning of Seven Putative Yeast Methyltransferases—The genes designated as *F1*, *F7*, *F8*, *F10*, and *F12* were subcloned into the pGEX-2T vector (Table II). Each gene was amplified by PCR from yeast genomic DNA using primers containing a *Bam*HI site in the 5' primer and *Eco*RI site in the 3' primer (Table III). PCR products and the pGEX-2T vector (Amersham Pharmacia Biotech) were digested with *Bam*HI and *Eco*RI, and gene inserts were ligated into the pGEX-2T vector in frame with the glutathione *S*-transferase coding region to allow for expression of each gene as a fusion protein. DH5 α cells were transformed with each ligation mixture, and the plasmids containing gene inserts *F1*, *F7*, *F8*, *F10*, and *F12* were designated pAN102, -100, -106, -107, and -108, respectively (Table IV).

The genes *F3* and *F4* were subcloned into the pCR2.1 vector (Table II) using the Original TA Cloning Kit (Invitrogen) according to manufacturer's instructions (Invitrogen). Resulting plasmids containing the gene inserts *F3* and *F4* were designated pAN109 and -110, respectively (Table IV). Both genes were also cloned into the pGEX-3X vector by releasing the genes from pAN109 and pAN110 with *Bam*HI and ligating the inserts into the *Bam*HI site of pGEX-3X. The resulting plasmids were designated pAN130 and pAN120.

The yeast expression vector containing *F3* was constructed by cutting the gene out from the vector pAN109 with *Bam*HI and ligating it into the *Bam*HI site of pCH1 (Table IV). The forward orientation of the insert was confirmed by restriction digests, and the resulting plasmid was designated pAN132. The yeast strain AN3-S was transformed with pAN132 or pCH1 as described below for yeast disruptions and selected on uracil-deficient SD plates to make the yeast strains AN-SR2 and

TABLE II
Cloning strategy for yeast disruptions of 7 putative methyltransferases

ORF	Code	Size	Cloning vector	Disruption sites ^a	Auxotrophic marker	Resulting disruption plasmid	Enzymes used to cut out construct	Resulting yeast strains
		bp						
F1	YDL201w	858	PGEX-2T	<i>Bgl</i> III	<i>HIS3</i> , <i>TRP1</i>	pAN103, pAN115	<i>Bam</i> HI, <i>Eco</i> RI	AN1-H, AN1
F3	YDR465c	1236	pCR2.1	<i>Cla</i> I	<i>TRP1</i>	pAN126	<i>Bam</i> HI, <i>Sma</i> I	AN3, AN3-S
F4	YJL125c	1149	PGEX-3X	<i>Bgl</i> III (×2)	<i>TRP1</i>	pAN122	<i>Bam</i> HI, <i>Sma</i> I	AN4
F7	YCR047c	825	PGEX-2T	<i>Bsp</i> EI	<i>HIS3</i> , <i>TRP1</i>	pAN101, pAN127	<i>Bam</i> HI, <i>Eco</i> RI	AN7-H, AN7
F8	YDR140w	663	PGEX-2T	<i>Nco</i> I	<i>TRP1</i>	pAN123	<i>Bam</i> HI, <i>Eco</i> RI	AN8
F10	YHR209w	873	PGEX-2T	<i>Cla</i> I	<i>TRP1</i>	pAN128	<i>Bam</i> HI, <i>Eco</i> RI	AN10, AN10-S
F12	YOR240w	1086	PGEX-2T	<i>Nco</i> I	<i>TRP1</i>	pAN114	<i>Bam</i> HI, <i>Eco</i> RI	AN12

^a There are two *Bgl*III sites in F4; therefore, insertion of the *TRP1* marker results in the deletion of a 218-base pair fragment. There are also two *Cla*I sites in F3, but the second site is blocked by *dam* methylation so it is not cut during linearization of the vector.

TABLE III
Primers

Primer	ORF No.	Sequence ^a	Restriction sites
Primers used for PCR amplification of putative methyltransferase genes			
D1075.1-N	(F1)	TTGGATCCATGAAAGCCAAGCCACTA	<i>Bam</i> HI
D1075.2-N	(F1)	CGAATTCCTTACAATATGGCTGGCG	<i>Eco</i> RI
GAMT.F	(F3)	CATGGATCCGTATGTCAGAATTACATGC	<i>Bam</i> HI
GAMT.R	(F3)	AATTCCTGGGCGATTGTGCAATTTACAT	<i>Sma</i> I
GCD.F	(F4)	GCAGGGATCCAAATGTCAACAAATTTGT	<i>Bam</i> HI
GCD.R	(F4)	TTTACCCGGGCGAGATTGCTTTACAT	<i>Sma</i> I
YCT.1	(F7)	CTTGGATCCATGTCACGTCCTGAGGA	<i>Bam</i> HI
YCT.2	(F7)	CTGGATCCGAATTCCTAGAACCTGTGTCTTC	<i>Bam</i> HI, <i>Eco</i> RI
YD9302.F	(F8)	ACAGGATCCGAAAGGTGAAGTATGCT	<i>Bam</i> HI
YD9302.R	(F8)	TGGAATTCCTCACCTTGTAAAGCTGTACAC	<i>Eco</i> RI
YHO9.F	(F10)	TATGGATCCCTGTCAATGGAAGAATC	<i>Bam</i> HI
YHO9.R	(F10)	CCGAATTCCTCTTTCTACATAAGTAA	<i>Eco</i> RI
YOR240.F	(F12)	CCGGGATCCAAGAAGTGAATGATTTG	<i>Bam</i> HI
YOR240.R	(F12)	GTGGAATTCGCTCTCAGAACTTCCTA	<i>Eco</i> RI
Primers used to amplify auxotrophic markers from YDp vectors ^b			
Mark B/C		GCAGTCCGGAATCGATCaGGTGATTGATTGAGC	<i>Bsp</i> EI, <i>Cla</i> I
Mark N/X		CATCTAGACCATGGATCaGGTGATTGATTGAGC	<i>Xba</i> I, <i>Nco</i> I
HIS/ <i>Bsp</i> EI.F		CGTCCGGAGTCACTGCCAGGTAT	<i>Bsp</i> EI
HIS/ <i>Bsp</i> EI.R		TATCCGGAAGCGCCCTCGT	<i>Bsp</i> EI

^a Underlined sequences show placement of restriction sites shown on the right.

^b YDp vectors are from Berben *et al.* (44).

AN-SV2, respectively (Table IV).

Yeast Chromosomal Disruptions of Seven Putative Methyltransferases—Yeast disruption was carried out using the one-step technique described previously (50). Disruption constructs were created by inserting auxotrophic markers into the middle of the coding region for each gene at a specific restriction site (Table II). The auxotrophically selectable markers for disruption were *TRP1* or *HIS3* and were obtained from YDp plasmids (44) either directly by digesting at *Bam*HI sites or by PCR amplification using the YDp vectors as templates. The primers used for amplification of auxotrophic markers contained restriction sites (*Bsp*EI, *Cla*I, or *Nco*I) required for subsequent ligation of each marker into a matching site in the gene (Tables II and III). One primer was used in each PCR reaction because the sequences flanking the auxotrophic markers on YDp vectors are identical. However, to amplify *HIS3* for F7 disruption, two primers, HIS/*Bsp*EI.1 and HIS/*Bsp*EI.2, were used (Table III). For versatility, PCR products of the markers were subcloned into either the pCR2.1 vector (original TA cloning vector, Invitrogen) to create plasmids pAN116 and pAN121 or into pBluescript to create pAN099 (Table IV). Each auxotrophic marker was cut out of its plasmid at the appropriate restriction sites as needed for disruption.

The disruption constructs were created by linearizing each of the vectors containing one of the seven genes at a restriction site in the middle of the coding region (Table II). The gene F4 contained two *Bgl*III sites, and therefore, digestion of the pAN120 vector resulted in the deletion of a 218-base pair fragment (Table II). An auxotrophic marker with matching restriction sites at each end was then obtained from the plasmids pAN116, pAN121, and pAN099 (Table IV) by digesting with the appropriate restriction enzyme and ligating the marker into one of

the linearized vectors. In cases where *Bgl*III sites were used for disruption (F1 and F4; Table II), the auxotrophic marker was obtained from one of the YDp vectors by digesting with *Bam*HI and ligating the marker directly into the *Bgl*III sites.

Disruption constructs were released from each vector using *Bam*HI and either *Eco*RI or *Sma*I (Table II). The disruption mutants of *S. cerevisiae* were created in the strain SEY6210 or in JDG9100-2, a strain defective in arginine methylation (Table IV). For diploid disruptions of F4 and F10, the strain GPY278 was used (Table IV). The transformation protocol involved a simplified lithium acetate procedure as described (51). Yeast transformants were selected on tryptophan or histidine-deficient SD plates using an amino acid dropout mix (Bio 101). The transformants were replated at least 3 times before they were confirmed by Southern blot or PCR. The yeast strains are listed in Table IV. The strain SEY6210-R was created in the strain SEY6210 as described previously (22) in order to disrupt the chromosomal *RMT1* gene and was confirmed by PCR and by amino acid analysis of methylarginines (22).

Tetrad Analysis—Sporulation and tetrad analyses were conducted as described by Sherman and Hicks (52). A haploid disruption mutant of F10 (AN10-S; Table IV) was obtained by sporulating the strain AN10 (Table IV) and isolating a spore with the genotype *MATα*, *leu2-3*, *112*, *ura3-52*, *his3-Δ200*, *trp1-Δ901*, *lys2-801*, *suc2-Δ9*, *mel*, *F10::TRP1*. The isogenic strain SEY6210 was used as a control strain for characterization of AN10-S.

Preparation of in Vivo Labeled Yeast Extracts—*S. cerevisiae* mutant and parent cells were labeled *in vivo* using [³H]AdoMet. Cells were grown to early log phase ($A_{600\text{ nm}} = 0.6-0.8$) in YPD media (1% bacto-yeast extract, 2% bacto-peptone, 2% dextrose) at 30 °C. Five $A_{600\text{ nm}}$

TABLE IV
Strains and plasmids

Name	Description	Ref.
Yeast strains		
GPY278	SEY6210xSEY6211 MAT α /MAT α , <i>leu2-3, 112/leu2-3, 112, ura3-52/ura3-52, his3-Δ200/his3-Δ200, trp1-Δ901/trp1-Δ901, lys2-801/LYS2, ade2-101/ADE2, suc2-Δ9/suc2-Δ9, mel/mel</i>	Greg Payne, UCLA
AN4	GPY278, F4/F4::TRP1	This work
AN10	GPY278, F10/F10::TRP1	This work
AN10-S	MAT α , <i>leu2-3, 112, ura3-52, his3-Δ200, trp1-Δ901, lys2-801, suc2-Δ9, mel, F10::TRP1</i>	This work
SEY6210	MAT α <i>ura3-52, leu2-3, 112, his3Δ200, trp1-Δ901, his2-801, suc2-Δ9, mel</i>	45
SEY6210-R	SEY6210, <i>rmt1::LEU2</i>	This work
CH9100-2	MAT α , <i>prc1-407, prb1-1122, pep4-3, leu2, trp1, ura3-52, ycl57wΔ::URA3</i>	47
JDG9100-2	CH9100-2, <i>rmt1::LEU2</i>	22
AN1	JDG9100-2, F1::TRP1	This work
AN1-H	SEY6210, F1::HIS3	This work
AN3	JDG9100-2, F3::TRP1	This work
AN3-S	SEY6210-R, F3::TRP1	This work
AN3-SR2	AN3-S with pAN132 vector	This work
AN3-SV2	AN3-S with pCH1 vector	This work
AN7	JDG9100-2, F7::TRP1	This work
AN7-H	SEY6210, F7::HIS3	This work
AN8	JDG9100-2, F8::TRP1	This work
AN12	JDG9100-2, F12::TRP1	This work
Plasmids		
pAN099	HIS3 flanked by BspEI in pBluescript	This work
pAN100	YCR047c (F7) in PGEX-2T	This work
pAN101	YCR047c (F7)::HIS3 in PGEX-2T	This work
pAN102	YDL201w (F1) in PGEX-2T	This work
pAN103	YDL201w (F1)::HIS3 in PGEX-2T	This work
pAN106	YDR140w (F8) in PGEX-2T	This work
pAN107	YHR209w (F10) in PGEX-2T	This work
pAN108	YOR240w (F12) in PGEX-2T	This work
pAN109	YDR465c (F3) in pCR2.1	This work
pAN110	YJL125c (GCD14) in pCR2.1	This work
pAN114	YOR240w (F12)::TRP1 in PGEX-2T	This work
pAN115	YDL201w (F1)::TRP1 in PGEX-2T	This work
pAN116	TRP1 flanked by XbaI and NcoI in pCR2.1	This work
pAN120	YJL125c (GCD14) in PGEX-3X	This work
pAN121	TRP1 flanked by BspEI and ClaI in pCR2.1	This work
pAN122	YJL125c (GCD14)::TRP1 in PGEX-3X	This work
pAN123	YDR140w (F8)::TRP1 in PGEX-2T	This work
pAN126	YDR465c (F3)::TRP1 in pCR2.1	This work
pAN127	YCR047c (F7)::TRP1 in PGEX-2T	This work
pAN128	YHR209w (F10)::TRP1 in PGEX-2T	This work
pAN130	YDR465c (F3) in PGEX-3X	This work
pAN132	YDR465c (F3) in pCH1	This work
YDp-W	Plasmid carrying TRP1	44
YDp-H	Plasmid carrying HIS3	44
pRS426	Multicopy shuttle vector with URA3	48
pCH1	pRS426 with CYC1 promoter inserted at EcoRI site	49
pGEX-2T	GST ^a expression plasmid	Amersham Pharmacia Biotech
pGEX-3X	GST ^a expression plasmid	Amersham Pharmacia Biotech
pBluescript		Stratagene
pCR2.1		Invitrogen

^a GST, glutathione S-transferase.

units of each culture were then harvested by centrifugation at 1,600 \times *g* for 6 min at 25 °C. The pelleted cells were resuspended in 820 μ l of YPD, and 180 μ l of [³H]AdoMet (NEN Life Science Products, 77.9 Ci/mmol, 550 μ Ci/ml in 10 mM H₂SO₄/ethanol (9:1, v/v)) was added. Cells were labeled for 30 min at 30 °C with shaking, pelleted at 16,000 \times *g* for 5 s at 25 °C, and washed twice with water, and pellets were saved at -80 °C. For the strains AN3-SR2 and AN3-SV2, *in vivo* labeling was performed as above except cells were grown and labeled in uracil-deficient SD medium to maintain the plasmids, and the [³H]AdoMet was at 72.4 Ci/mmol. Labeled cell extracts were prepared by resuspending each cell pellet in 50 μ l of lysis buffer (1% sodium dodecyl sulfate, 1 mM phenylmethylsulfonyl fluoride, 1 mM benzamidine, 2 μ g/ml each leupeptin, pepstatin, and aprotinin). To each mixture, 0.2 g of baked zirconium beads (Biospec Products) were added, and each sample was vortexed 1 min and placed on ice for 1 min. The vortexing and ice incubation was repeated seven times. The lysate was collected, and the beads were washed with 50 μ l 1% sodium dodecyl sulfate. The wash was combined with the lysate and debris was spun down at 16,000 \times *g* for 5 s. The cleared lysate was used for acid hydrolysis and amino acid analysis.

Determination of Protein Concentrations—Protein concentrations were determined using a Coomassie Plus Protein Assay Reagent (Pierce) against a standard of bovine serum albumin.

RESULTS

Identification of 26 Putative Methyltransferases in S. cerevisiae—We were interested in identifying new methyltransferases in the complete genome of the yeast *S. cerevisiae* based on conserved sequence motifs common to AdoMet-dependent methyltransferases. We used motifs I, post-I, II, and III, which we defined previously based on a large number of known methyltransferases (15), to search the yeast genome at the site Motifs in Protein Data Bases.² The search resulted in 33 candidate genes with identifiable motifs (Table V). Seven of the

² Available on-line at the following address: <http://alces.med.umn.edu/dbmotif.html>.

TABLE V
Open reading frames identified using conserved methyltransferase sequence motifs

Consensus for conserved motifs was derived by Kagan and Clarke (15) based on a large number of known methyltransferases. The conserved motifs are always found in the same order on the polypeptide as shown above. The intervals between motifs are also conserved, with the average interval length shown above. The first consensus row immediately above the sequences represents the most common residue for that position. The second, third, and fourth most common residues at each position are in the rows above the first row. h, hydrophobic residue; X, any residue.

Code	Gene or ORF	Motif I L IVEV C P VLDIGGGTG	Interval 13	Post-I E hhXhD	Interval 17-71	Motif II A Y L GTY VIV PQFDAIFC	Interval 22	Motif III V L I K IIFL LLRPGGRLLI
Known methyltransferases								
YBR236c	<i>ABD1</i>	VLELGCGKG	13	FIGID	46	FPCDIVST	23	SLKIGGHFFG
YOL096c	<i>COQ3</i>	VLDVCGGGG	14	VQGID	34	GQFDIITC	20	LNPEKGILFL
YML110c	<i>COQ5</i>	FIDVAGGSG	13	FGDTE	37	DSKDIYTV	19	VLKPGGIFYC
YPL266w	<i>DIM1</i>	VLEVPGPTG	12	VVAVE	32	PYFDICIS	36	LARPGDSLVC
YML008c	<i>ERG6</i>	VLDVCGVVG	13	VIGLN	36	NTFDKVVA	19	VLKPGGTFVA
YBR034c	<i>RMT1</i>	VLDVCGGTG	13	VIGVD	35	PKVDIIIS	22	YLVEGGLIFP
YKR069w	<i>MET1</i>	ISLVGSGPG	17	IILAD	51	KQGDPYIF	17	VVLPGISSSL
Good matches								
YDL201w	<i>F1</i>	IADIGCGFG	14	ILGME	54	CFPDPHFK	19	VLKEGGVVYT
YDR316w	<i>F2</i>	VLEVSCGTG	13	ITFLD	51	VKYDTIVE	19	LLKPDGRILL
YDR465c	<i>F3 (RMT2)</i>	ILNIGFGMG	13	HYICE	56	VFFDGIYY	19	LKPEGVVFSF
YJL125c	<i>F4 (GCD14)</i>	VIEAGTGSG	12	LFSFE				
YPL157w	<i>F5</i>	ILDVFCGGG	12	VYVVD	40	IKYDCVFG	19	HLKPMGITKM
YML014w	<i>F6</i>	GIDVCGNG	10	IIGSD	30	ETFDFAIS	22	KLRQGGQALI
YCR047c	<i>F7</i>	ILDIGCGSG	13	WCGLD	31	GSFDDAIS	30	ALKKGGKFVA
YDR140w	<i>F8</i>	VCEIGSGSG	18	HLAVD	36	NQVDVLIF	25	LALLGGKDGDM
YER175c	<i>F9</i>	LVDVCGGPG	15	IIGSD	43	QKIDMITA	18	LRKDGTFIAIW
YHR209w	<i>F10</i>	LVDIGCGTG	12	VIGID	37	ESVDMVIS	18	LRSDGTFAFW
YIL064w	<i>F11</i>	VVDLGTGNG	14	LVGID	45	GTLDAISL	21	ILKKGIFLI
YOR240w	<i>F12</i>	IFEIGCGAG	16	IIAAD	40	HSVDIAVM	21	ILKPGGKIIF
YDL014w	<i>F13 (NOP1)</i>	VLYLGAASG	15	VYAVE	36	GMVDCVFA	17	FLKDGQGVVI
Marginal matches								
YAL061w	<i>F14</i>	VLEVPGVVK	13	GTCRD				
YBR261c	<i>F15</i>	AVDIGAGIG	13	IDLVE	32	WTPDAGKY	26	GLQPNGTIVV
YBR271w	<i>F16</i>	VLELGAGTG	12	LYGTE	49	NEFDVILI	19	FLAASGTCHL
YDR083w	<i>F17</i>	IADMGCGEA						
YIL110w	<i>F18</i>	VVEIGCGTA	21	FVLTD				
YJR072c	<i>F19</i>	NGDNGLGSG	16	YSQLD	23	DEYDQYYK	24	LMKDLGLNEK
YJR129c	<i>F20</i>	VLEVAGTG	16	MYVTD				
YLR137w	<i>F21</i>	ILELGAGIS	13	YVSTD				
YLR285w	<i>F22</i>	VLELGAAAA	13	VVSTD	50	GKFDLIIIL	19	LLAEKGQALV
YNL092w	<i>F23</i>	ILIPGCGTG	12	CEGNE	98	NSKDVVVT	19	VLKPGGIWCN
YNR029c	<i>F24</i>	ILDSGSGNG	15	PLKNE				
	<i>F25 (COX1)</i>	LVESGAGTG	14	GPSVD				
YER095w	<i>F26 (RAD51)</i>	PLDIGGEGE	21	RPGLD				

TABLE VI
BLAST-P searches of putative methyltransferases against GenBank™

ORF	Closest methyltransferase match for each yeast ORF	E value ^a closest MT match	E value ^b closest human EST match
<i>F1</i>	Protein similar to RNA MT (<i>Bacillus subtilis</i>)	0.47	6×10^{-59}
<i>F2</i>	Ubiquinone/menaquinone biosynthesis MT (<i>Archeoglobus fulgidus</i>)	3×10^{-7}	None
<i>F3</i>	Guanidinoacetate N-MT (human)	2×10^{-6}	None ^c
<i>F4</i>	β -Aspartate MT (<i>Mycobacterium leprae</i>)	10^{-8}	2×10^{-25}
<i>F5</i>	Hypothetical RNA methyltransferase (<i>Escherichia coli</i>)	0.42	5×10^{-9}
<i>F6</i>	2-hexaprenyl-1, 4-naphthoquinone MT (<i>Micrococcus luteus</i>)	0.014	2×10^{-8}
<i>F7</i>	Glycine N-MT (<i>Sus scrofa</i>)	0.013	9×10^{-38}
<i>F8</i>	Hypothetical adenine-specific MT (<i>Haemophilus influenzae</i>)	2×10^{-6}	8×10^{-10}
<i>F9</i>	2-hexaprenyl-1, 4-naphthoquinone MT (<i>M. luteus</i>)	1.5	None
<i>F10</i>	Cyclopropane fatty acid synthase (<i>Helicobacter pylori</i>)	4×10^{-4}	None
<i>F11</i>	EryG protein (<i>Saccharopolyspora erythraea</i>)	0.016	10^{-8}
<i>F12</i>	Ubiquinone biosynthesis MT (<i>Schizosaccharomyces pombe</i>)	0.072	3×10^{-30}
<i>F13</i>	Protein L-isoaspartate O-MT (<i>Methanobacterium thermoautotrophicum</i>)	0.11	1×10^{-72}
<i>F15</i>	Carminomycin 4-O-MT (<i>Streptomyces peucetius</i>)	0.22	4×10^{-16}
<i>F16</i>	Protein-L-isoaspartate O-MT (<i>Arabidopsis thaliana</i>)	0.75	10^{-6}
<i>F20</i>	Ribosomal protein L11 MT (<i>H. pylori</i>)	1.0	None

^a The E (Expect) value of the closest known methyltransferase (MT) match to each yeast ORF in the database of all non-redundant GenBank™ CDS translations + PDB + SwissProt + PIR obtained with BLAST version 2.0.

^b The E value of the closest human match to each yeast ORF from the non-redundant GenBank™ database of human EST entries. *F13* codes for yeast fibrillar, and the E value for the human match listed represents human fibrillar.

^c The closest human match to *F3* in the EST database is guanidinoacetate N-methyltransferase with an E value of 0.001.

genes turned out to be known methyltransferases responsible for modifying a variety of substrates including ubiquinone precursors, rRNA, the mRNA cap, an ergosterol precursor, uropor-

phyrin, and arginine residues on proteins (Tables I and V). We grouped the remaining 26 genes in two categories, "good matches" and "marginal matches" based on the number of

FIG. 2. Alignment of YDR465c (F3; Rmt2) and human guanidinoacetate N-methyltransferase (guanidinoacetate N-methyltransferase (GAMT); Ref. 43). The alignment was made using BLAST version 2.0. The conserved motifs I, post-I, II, and III are underlined (15). The deduced sequence of YDR465c is labeled RMT2. The guanidinoacetate N-methyltransferase sequence is GenBank™ accession number 2948404. Amino acid residues that are similar between the two sequences are marked by a +.

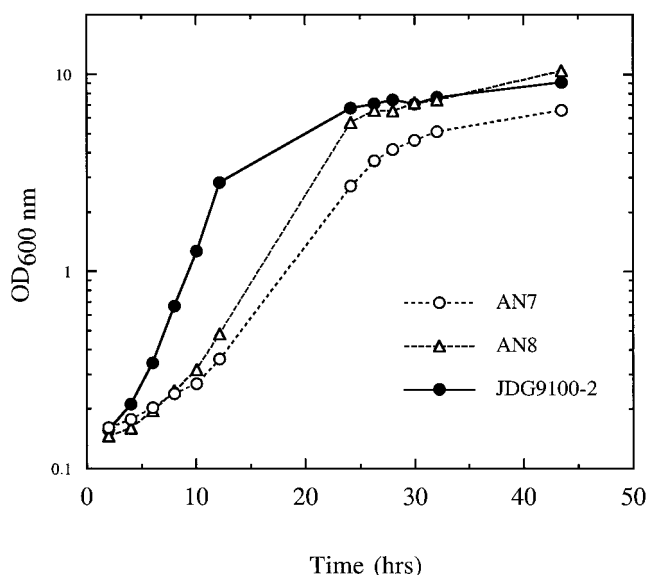
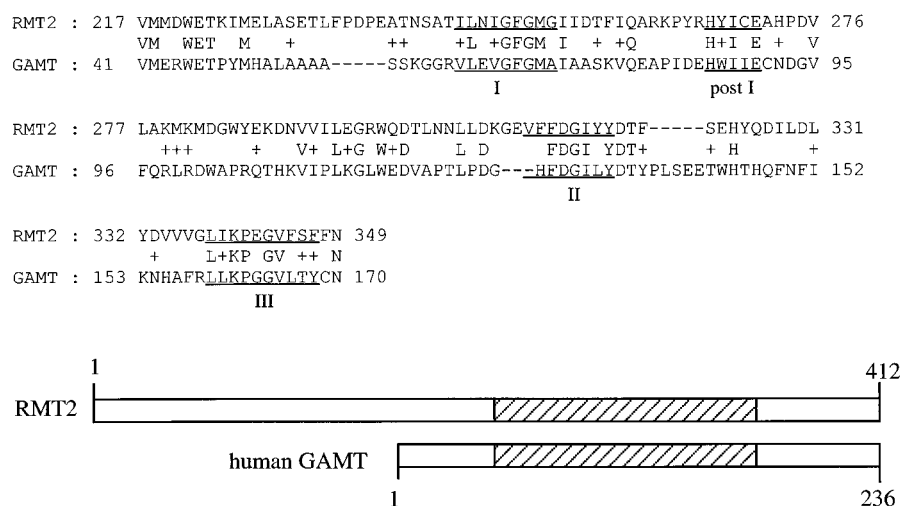


FIG. 3. Disruption of yeast *YCR047c* and *YDR140w* genes leads to slow growth of the resulting yeast strains AN7 and AN8. The parent strain JDG9100-2 and mutant strains AN7 and AN8 were grown in YPD-rich media at 30 °C, and cell numbers were estimated by the optical density of the culture at 600 nm.

conserved regions present in each sequence and on how well the conserved regions, and the intervals between the regions, matched the consensus sequences. We designated these genes as *F1* to *F26* (Table V). Among the new sequences are four genes (*GCD14*, *NOPI1*, *RAD51*, and *COX1*) which have previously been characterized but have not been associated with methylation. Interestingly, Nop1, a fibrillar homologue, is known to be important for proper methylation of pre-rRNA (53), and it associates with small nucleolar RNAs to allow small nucleolar RNA to properly guide 2'-*O*-methylation (54). It is possible that fibrillar may itself be the rRNA ribose *O*-methyltransferase. However, its direct role as a methyltransferase has not been demonstrated. The remaining sequences were previously uncharacterized open reading frames.

Sequence Analysis of the Putative Methyltransferases—In order to identify likely substrates for each putative methyltransferase, we carried out BLAST 2.0 searches of the GenBank™ data base with the translated yeast ORF sequences (Table VI). Many of the putative methyltransferases displayed significant amino acid sequence similarities to known AdoMet-dependent methyltransferases in the data base. However, most of the observed similarities were found in the protein segments span-

ning the conserved methyltransferase motifs, and the overall similarity scores were low. For example, the F7 putative methyltransferase could be aligned to five different classes of known methyltransferases, but even the best alignment resulted in the relatively high expect (E) value of 0.013 (Table VI). This observation suggested that although the *F7* ORF likely encodes a methyltransferase, the degree of sequence similarity to known methyltransferases was insufficient to determine its specific substrate. The low overall similarity scores most likely prevented the sequences identified in this study from being recognized as putative methyltransferases previously. In fact, only one (*F7*) of the 26 putative methyltransferases was identified as such in the GenBank™ data base.

There were several sequences, however, that displayed additional amino acid similarities outside of the conserved methyltransferase motifs. For example, the F3 coding sequence was aligned with an E value of 2×10^{-6} to the human guanidinoacetate N-methyltransferase (GenBank™ accession number 2498404, Table VI) and showed sequence similarity in regions preceding motifs I and II (Fig. 2). These similarities suggested that the substrates of F3 and guanidinoacetate N-methyltransferase may share common features (see below). Other notable alignments included the alignment of Gcd14 to a putative β -aspartate methyltransferase from *Mycobacterium leprae* with an E value of 10^{-8} (GenBank™ accession number 2145817, Table VI), the alignment of F8 to several adenine methyltransferases, and the alignment of F2 to a ubiquinone/menaquinone biosynthesis methyltransferase from *Archaeoglobus fulgidus* with an E value of 3×10^{-7} (GenBank™ accession number 2650497, Table VI). The sequence similarity between F2 and the ubiquinone/menaquinone biosynthesis methyltransferase is consistent with the predicted location of F2 in the mitochondria³ and suggests that F2 may methylate a ubiquinone precursor or a similar molecule. Two of the gene products, F9 and F10, were aligned with an E value of 3×10^{-28} with each other, indicating that they may share similar functions. In addition, for a majority of the putative methyltransferases, a human sequence of unknown function with high similarity to each yeast ORF was found in the EST data base (Table VI).

Yeast Disruption Mutants of 7 Putative Methyltransferase Genes Reveal One Lethal and Two Slow Growth Phenotypes—As a first step to characterize the putative methyltransferases, we decided to make yeast disruption mutants of several genes which we considered good matches. Our goal was to

³ PSORT II available at the following address: <http://cookie.imcb.osaka-u.ac.jp/nakai/psort.html>.

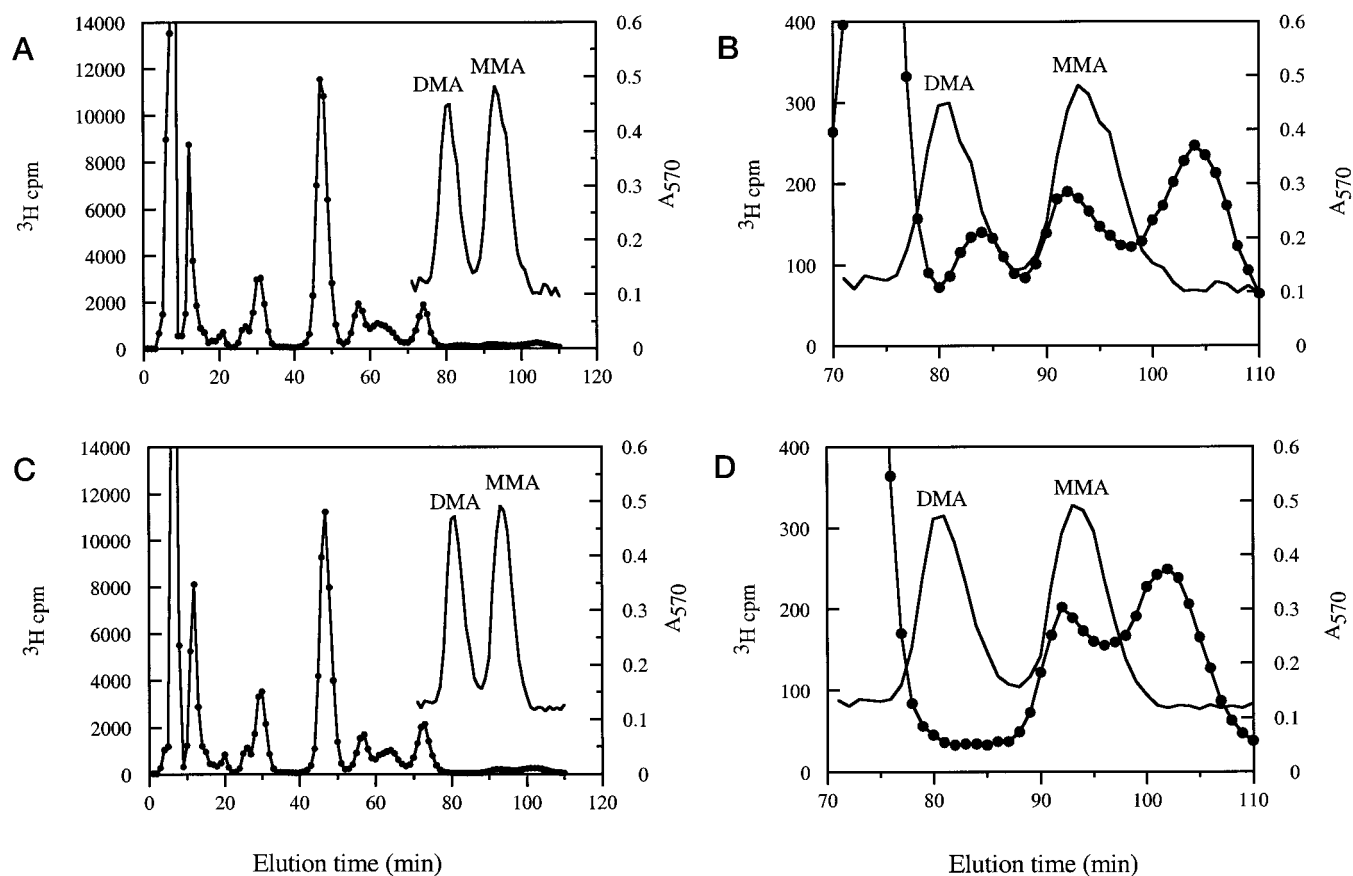


FIG. 4. Disruption of the *YDR465c/RMT2* gene product leads to a methylation defect of a novel type of arginine derivative. *A* and *B*, control JDG9100-2 strain. *C* and *D*, AN3 mutant strain. *B* and *D* are enlargements of *A* and *C*, respectively, in the region where methylated arginine derivatives elute. The yeast disruption mutant AN3 and its parent strain JDG9100-2 were labeled in YPD *in vivo* with [^3H]AdoMet, and extracts were prepared as described under "Experimental Procedures." 30 μl of cleared *in vivo* labeled lysate was combined with an equal volume of 25% trichloroacetic acid in a 6 \times 50-mm glass vial and incubated at 25 $^{\circ}\text{C}$ for 20 min. Each sample was spun down at 4000 $\times g$ for 20 min at 25 $^{\circ}\text{C}$. The pellets were then washed once with acetone at -20°C and acid-hydrolyzed with 200 μl 6 M HCl at 110 $^{\circ}\text{C}$ for 20 h in a Waters Pico-Tag vapor-phase apparatus. The hydrolyzed samples were suspended in 200 μl water and an equal volume of citrate dilution buffer (0.2 N in Na^+ containing 2% thiodiglycol and 0.1% phenol at pH 2.2; Pierce). The samples were mixed with 1 μmol of each of the non-isotopically labeled standards, N^G,N^G -dimethylated (asymmetric) arginine (Sigma) and N^G -monomethylated arginine (Calbiochem). The entire sample was loaded on a Beckman AA-15 sulfonated polystyrene column, 0.9 cm diameter \times 11 cm height. The sample was eluted at 1 ml/min with sodium citrate buffer (0.35 N in Na^+ , pH 5.14) at 55 $^{\circ}\text{C}$. After each run, the column was washed with 0.2 N NaOH for 20 min prior to the next run. One-minute fractions were collected, and ^3H radioactivity was determined by liquid scintillation counting by dissolving a 200- μl aliquot of each fraction in 500 μl of water and placing it in 5 ml of fluor (Safety-Solve, Research Products International). An additional 100 μl of some fractions was analyzed using the ninhydrin method to detect the non-radiolabeled standards. 100 μl of each fraction was combined with 300 μl of ninhydrin reagent (2% (w/v) ninhydrin and 3 mg/ml hydrindantin in a solvent of 75% (v/v) dimethyl sulfoxide and 25% (v/v) 4 M lithium acetate at pH 4.2). The mixture was heated for 15 min at 100 $^{\circ}\text{C}$, and the absorbance at 570 nm was measured. DMA, N^G,N^G -dimethylarginine (asymmetric) standard; MMA, N^G -monomethylarginine standard.

find the methylation substrate for each new methyltransferase on our list, and the availability of yeast mutants facilitates identification of methylation defects. Disruptions were made in the genes *F1*, *F3*, *F4*, *F7*, *F8*, *F10*, and *F12* by inserting either a *TRP1* or a *HIS3* cassette at a specific restriction site in each gene as described under "Experimental Procedures."

We found that *F4* (*GCD14*) is an essential gene. Dissection of six individual tetrads from the diploid disruption mutant AN4 resulted in 2:2 segregation of the lethal phenotype (data not shown). In addition, all of the viable spores of *F4* were auxotrophic for tryptophan (data not shown), reflecting the absence of the *TRP1* insert in the disrupted *F4* gene. In contrast, disruption of the genes *F1*, *F3*, *F7*, *F8*, *F10*, and *F12* resulted in viable colonies. In addition, we observed that the disruption of two of the genes, *F7* and *F8*, leads to slow growth of the resulting yeast mutants. The doubling times for the strains AN7 and AN8 in rich media during exponential growth phase were 4.2 and 3.4 h respectively, compared with 2.1 h for the parent strain (Fig. 3). Similarly, a disruption of *F7* in a SEY6210 yeast background resulted in a doubling time of 3 h for the AN7-H

mutant compared with 1.5 h for the parent strain during exponential growth phase (data not shown). The remaining yeast mutants exhibited doubling times comparable to the corresponding parent strains (data not shown). We did not observe any additional growth defects of the mutants on glycerol, on minimal media, at 37 or 14 $^{\circ}\text{C}$ (data not shown).

Disruption of *F3* Leads to a Methylation Defect of a Novel Type of Arginine Derivative—As a first step to identify the substrate for each new methyltransferase on our list, we decided to examine the disruption mutants for defects in protein methylation. These mutants were made in a strain lacking the Rmt1 protein arginine methyltransferase to reduce the level of *N*- ω -methylated arginine species that could obscure other methylated derivatives (22). Several different protein methyltransferases, for which the corresponding genes have not yet been identified, are predicted to exist in yeast (Table I). In addition, the sequence similarity between one of the sequences in this study, *F3*, and guanidinoacetate *N*-methyltransferase (Fig. 2) led us to suspect that *F3* may be a methyltransferase responsible for modifying a molecule similar to guanidinoac-

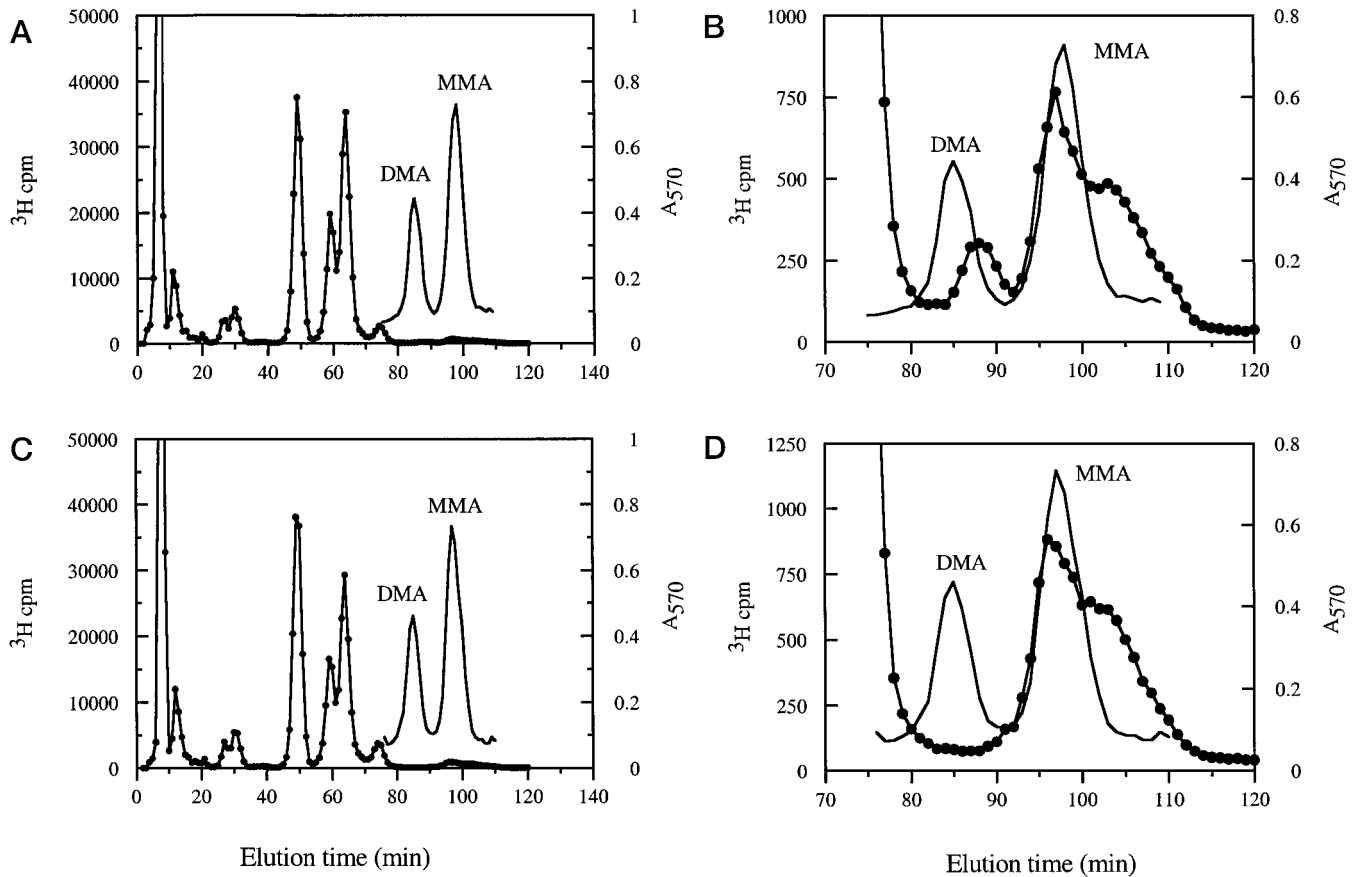


FIG. 5. **Constitutive expression of the *RMT2* gene product on a multicopy plasmid restores the defect in arginine methylation.** A and B, AN3-SR2 strain carrying the *RMT2* expression plasmid pAN132. C and D, AN3-SV2 strain carrying the control plasmid pCH1. B and D are enlargements of A and C, respectively, in the region where methylated arginine derivatives elute. The yeast strains AN3-SR2 and AN3-SV2 (Table IV) were labeled in uracil-deficient defined medium *in vivo* with [^3H]AdoMet, and extracts were prepared as described under "Experimental Procedures." The hydrolysates were prepared and analyzed as described in Fig. 4 except 75 μl of each yeast extract was used for acid hydrolysis and amino acid analysis. DMA, N^G,N^G -dimethylarginine (asymmetric) standard; MMA, N^G -monomethylarginine standard.

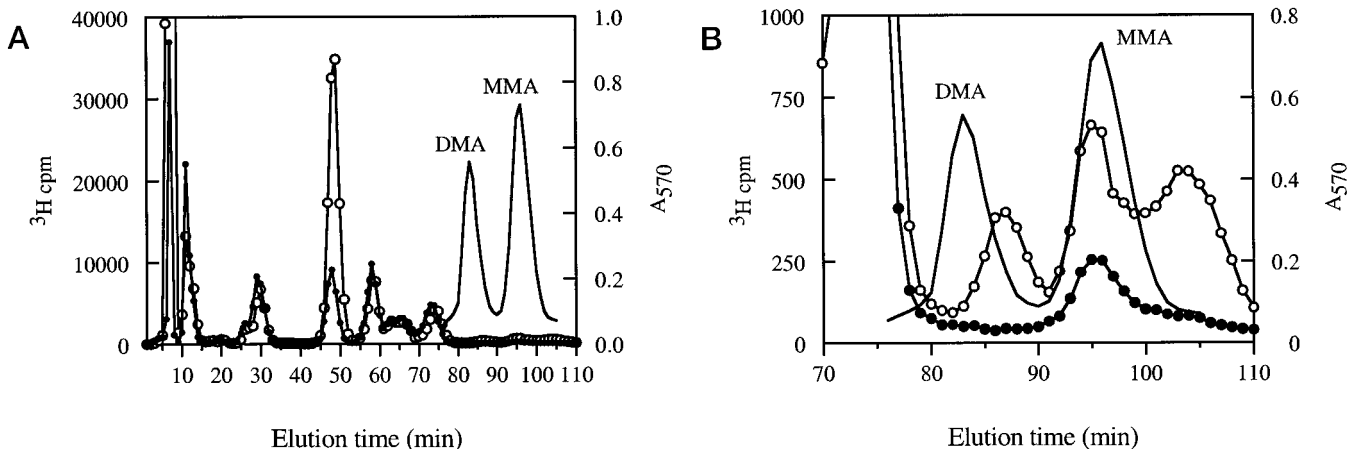


FIG. 6. **Inhibition of protein synthesis prevents *in vivo* methylation of the Rmt2 substrate.** B is an enlargement of A in the region where methylated arginine derivatives elute. JDG9100-2 cells were grown to $A_{600\text{ nm}} = 0.637$ in YPD media at 30 $^{\circ}\text{C}$. Five $A_{600\text{ nm}}$ units of each culture were then harvested by centrifugation at $1,600 \times g$ for 6 min at 25 $^{\circ}\text{C}$. The pelleted cells were resuspended in 894 μl of YPD and 10 μl of either cycloheximide (Sigma; 10 mg/ml stock in ethanol) or ethanol and incubated at room temperature for 10 min. 106 μl of [^3H]AdoMet (Amersham Pharmacia Biotech, 76.0 Ci/mmol, 1.0 mCi/ml in dilute HCl/ethanol 9:1) was added to give a final [^3H]AdoMet concentration of 1.4 μM . The final concentration of cycloheximide during labeling was 100 $\mu\text{g}/\text{ml}$. Cells were labeled for 20 min at 30 $^{\circ}\text{C}$ with shaking, pelleted at $16,000 \times g$ for 5 s at 25 $^{\circ}\text{C}$, and washed twice with water, and pellets were saved at -80°C . Hydrolysates were prepared and analyzed as described in Fig. 4. Open symbols, control JDG9100-2 cells; closed symbols, cycloheximide-treated JDG9100-2 cells. DMA, N^G,N^G -dimethylarginine (asymmetric) standard; MMA, N^G -monomethylarginine standard.

etate, such as an arginine residue of a protein.

We examined *in vivo* labeled extracts from all mutants for defects in protein methylation using amino acid analysis. Yeast proteins, methylated *in vivo* using [^3H]AdoMet, were collected

by trichloroacetic acid precipitation and acid-hydrolyzed. The methylated derivatives were analyzed using cation exchange liquid chromatography. The resulting pattern of amino acid methylation for the parent strain, JDG9100-2, is shown in Fig.

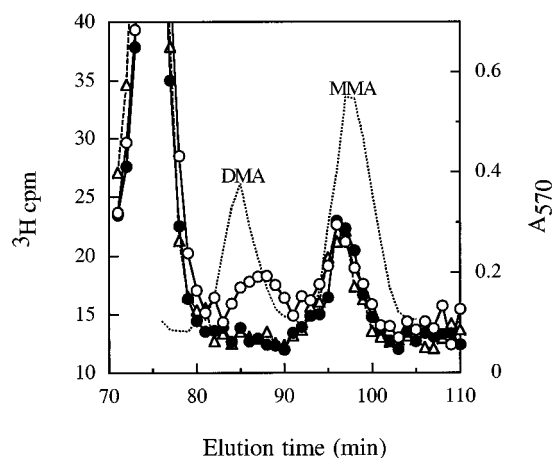


Fig. 7. *In vitro* formation of *N*- δ -methylarginine is limited by substrate availability. Yeast extracts from the strains AN3 and JDG9100-2 were prepared by growing a 100-ml culture of yeast cells to $A_{600\text{ nm}} = 0.616$ (JDG9100-2) or $A_{600\text{ nm}} = 0.737$ (AN3). Cells were spun down at $16,000 \times g$ for 5 s, washed twice with water, and resuspended in 150 μl of lysis buffer (250 mM Tris-HCl, pH 7.5, 1 mM sodium EDTA, 1 mM sodium EGTA, 0.3 mM phenylmethylsulfonyl fluoride, 1 mM benzamidine, 2 $\mu\text{g}/\mu\text{l}$ each leupeptin, pepstatin, and aprotinin). To each mixture, 0.6 g of baked zirconium beads (Biospec Products) were added, and each sample was repeatedly vortexed as described under "Experimental Procedures." The lysate was collected, and debris was spun down at $16,000 \times g$ for 5 s. Extracts from the JDG9100-2 and AN3 strains were assayed separately by combining 75 μl of each extract (945 μg of protein) with 1 μl of [^3H]AdoMet (Amersham Pharmacia Biotech, 76.0 Ci/mmol, 1.0 mCi/ml in dilute HCl/ethanol 9:1), or together by combining 37.5 μl of extract from each strain (945 μg of total protein) with 1 μl of [^3H]AdoMet. The reactions were incubated at 30 $^\circ\text{C}$ for 30 min. The reactions were stopped by adding an equal volume of 25% trichloroacetic acid, and hydrolysates were prepared as described in Fig. 4 except that the total volume of each methylation reaction (76 μl) was used for acid hydrolysis and amino acid analysis. *Open circles*, combined extracts of AN3 and JDG9100-2; *closed circles*, AN3 extract; *triangles*, JDG9100-2 extract. *DMA*, N^G,N^G -dimethylarginine (asymmetric) standard; *MMA*, N^G -monomethylarginine standard.

4A. The major peak of methylation corresponds to mono-, di-, and trimethyllysine eluting at 47 min. Since this parent strain is defective in N^G -mono- and N^G,N^G -asymmetric dimethylation of arginine residues (22), there are no large peaks of radioactivity which migrate with the methylarginine standards (Fig. 4A). We found, however, no large differences in the methylation pattern of AN3 (Fig. 4C) or of the AN1, AN7, AN8, AN10-S, and AN12 mutants compared with the parent strains (data not shown).

However, closer analysis of the methylarginine region from the AN3 mutant acid hydrolysate reveals a defect in methylation of an amino acid derivative eluting at about 84 min (Fig. 4D), which is clearly present in similarly prepared hydrolysates of the parent strain (Fig. 4B). This derivative elutes in a position distinct from that of the N^G,N^G - ω -dimethylarginine or N^G - ω -monomethylarginine standards. The sequence similarity between F3 and guanidinoacetate *N*-methyltransferase (Fig. 2), which methylates this small molecule substrate on the δ -nitrogen atom of its guanidino group, suggested that the substrate of F3 may also be methylated on a δ -nitrogen of a guanidino group but on an arginine residue. In fact, we have recently demonstrated that yeast proteins contain *N*- δ -methylarginine residues (41). Direct chemical analysis of the acid hydrolysis product eluting at 84 min revealed that this derivative corresponds to free *N*- δ -methylarginine (41). Three other amino acid derivatives which elute in the methylarginine region at 74, 93, and 103 min are unaffected by the *F3* disruption. The identity of these derivatives is unknown, but the peak eluting at about 92 min elutes with the N^G - ω -monomethylargi-

nine standard and suggests the existence of an additional type of N^G -arginine methyltransferase.

The methylation defect at 84 min in extracts from the AN3 strain is also observed in the AN3-S strain (Table IV) which has an *F3* disruption in a different *rmt1* background (data not shown). In order to confirm that the methylation defect in the AN3 mutants is due to the absence of *F3* activity, we reintroduced the *F3* gene into the AN3-S strain on a multicopy plasmid as described under "Experimental Procedures." The *F3* expression plasmid restores the methylation defect in the AN3-S mutant (Fig. 5B). In contrast, a vector-only control fails to restore methylation of the new derivative (Fig. 5D). In additional control experiments, we analyzed the pattern of methylated molecules from extracts of AN3 and JDG9100-2 strains which had not been precipitated with trichloroacetic acid prior to acid hydrolysis. The pattern of methylated derivatives obtained was very similar to the pattern seen in Fig. 4, showing that the methylated derivative is not due to a small molecule contaminant.

The absence of *N*- δ -methylarginine in acid hydrolysates of extracts of the *F3* disruption mutant coupled with the sequence similarity between *F3* and a known δ -nitrogen methyltransferase clearly indicates that *F3* is responsible for the formation of this novel arginine derivative. In contrast to guanidinoacetate methyltransferase, the *F3* enzyme modifies an amino acid residue of a protein. The 176 amino acid extension at the N-terminal of *F3*, which is absent in the small molecule methyltransferase (Fig. 2), may represent a substrate binding domain which allows *F3* to accommodate a much larger protein substrate. We designate the new enzyme as *Rmt2*.

Protein Synthesis Is Required for *In Vitro* Methylation of the *Rmt2* Substrate—We wanted to find out whether formation of the δ -methylarginine derivative can occur *in vitro*. However, incubation of yeast extracts with [^3H]AdoMet resulted in the formation of only a minimal level of the *N*- δ -methylarginine derivative (data not shown). We reasoned that this type of methylation may only occur in a short window of time immediately following translation, before the substrate becomes folded. This is the case for methylation of lysine residues of cytochrome *c* in yeast (55). We therefore decided to test whether inhibition of protein synthesis during *in vivo* labeling will prevent the new derivative from being formed. We labeled the JDG9100-2 strain in the presence or absence of cycloheximide and compared the resulting patterns of methylated amino acid residues (Fig. 6). We find that the *N*- δ -methylarginine peak at 87 min is completely absent upon cycloheximide treatment (Fig. 6B). Interestingly, there is also a significant reduction in the levels of methyllysine at 48 min (Fig. 6A), N^G -monomethylarginine at 95 min (Fig. 6B), and the unknown derivative at 104 min (Fig. 6B). In contrast, the levels of the other methylated derivatives are unaffected by cycloheximide treatment.

The absence of *N*- δ -methylarginine under conditions where protein synthesis is inhibited suggests that this novel type of modification may be cotranslational or may in some other way be dependent on protein synthesis. For example, the *Rmt2* substrate could have a high turnover rate or it could exist only in its fully methylated form. In order to test these possibilities, we carried out *in vitro* methylation reactions in which extracts from AN3 (which would be expected to contain unmethylated substrates) and JDG9100-2 (which would be expected to contain the *Rmt2* activity) were mixed. Although methylation of each extract individually did not result in formation of *N*- δ -methylarginine, combining the extracts led to the formation of the new derivative (Fig. 7). This result suggests both that the

substrate of Rmt2 is stable when it is unmethylated and that it normally exists in its fully methylated form.

DISCUSSION

Since the completion of the DNA sequence for the entire genome of the yeast *S. cerevisiae*, there has been a focused worldwide effort to uncover the function of each gene product in this organism (56). This effort has included systematic disruption of individual genes and genome-wide phenotypic assays of the resulting disruption mutants (Ref. 56 and references within; see also Ref. 57). It is expected that these genome-wide approaches to functional analysis will contribute tremendously to the understanding of how an organism works. However, direct biochemical analysis will remain a powerful approach to the study of many classes of proteins.

In order to identify and characterize new AdoMet-dependent methyltransferases, we combined a genome-wide approach with biochemical analysis. The presence of conserved motifs in this class of enzymes allowed us to identify 26 potential methyltransferases in the yeast genome that were not detected based on traditional sequence comparisons. Among the putative methyltransferases, we have identified a previously unknown enzymatic activity which catalyzes the unusual δ -nitrogen methylation of arginine residues. A direct biochemical approach was especially useful for characterizing this new methyltransferase because the disruption mutant does not have an obvious phenotype, and its amino acid sequence does not have high similarity to any known proteins. The low degree of sequence similarity of this Rmt2 methyltransferase to guanidinoacetate *N*-methyltransferase, an enzyme which also methylates the δ -nitrogen atom of a guanidino derivative, provided a clue about the substrate, but only in the context of what was already known about protein methylation. The recent disruption of the major yeast protein arginine methyltransferase gene *RMT1* (22) revealed the presence of a new arginine derivative and thus pointed to a possible connection between this derivative and the *RMT2* gene product. It is important to note that our identification of the activity of the Rmt2 enzyme was dependent upon having cells disrupted in two genes, a situation which would not occur when single mutants are analyzed en masse. We have therefore shown the added power of combining a direct biochemical approach with genome-wide approaches to elucidate the functions of new gene products.

The function of the new type of protein methylation by Rmt2 is unknown. Since the absence of Rmt2 does not affect general yeast viability or growth, *N*- δ -arginine methylation may be important under certain specific conditions or it may fine-tune the function of its protein substrate. Interestingly, in a recent study of yeast gene expression using DNA microarray technology, *RMT2* expression was found to be down-regulated after liquid cultures reached an $A_{600\text{ nm}}$ of 0.8, decreasing 3.6-fold at an $A_{600\text{ nm}}$ of 7.3 (58). The repression of *RMT2* implies that *N*- δ -arginine methylation may be important during logarithmic growth and may be dispensable or even deleterious during post-exponential growth. Homologs of Rmt2 are present in *Schizosaccharomyces pombe* (GenBankTM accession number 1723243; $E = 6 \times 10^{-55}$) and *Arabidopsis thaliana* (GenBankTM accession number 2827702; $E = 7 \times 10^{-42}$), but so far, it has not been found in prokaryotes or mammals. Finding the substrate for Rmt2 will be the next step in identifying the function of this new type of arginine methylation.

Amino acid sequence similarities, as in the case of Rmt2 and guanidinoacetate *N*-methyltransferase, may provide valuable starting points for elucidating the function of uncharacterized gene products. For example, the gene product of *F4/GCD14*, found in this study to be essential, has been recently identified as a translational repressor of *GCN4* (59). Gcn4 is a transcrip-

tion factor that activates at least 40 different genes encoding amino acid biosynthetic enzymes in response to amino acid or purine starvation (60, 61). Multiple translational repressors of *GCN4* (GCD proteins) are required for efficient repression of *GCN4* mRNA translation under non-starvation conditions (62). All of the GCD proteins characterized thus far have general roles in protein synthesis initiation, in addition to *GCN4* repression and deletions of many of these genes are lethal (Ref. 59 and references within). Several GCD proteins function to repress *GCN4* by reducing the activity of the translation initiation factor 2 (59). The mechanism of *GCN4* repression by F4/Gcd14 is not known, but our results suggest that it may involve a methylation reaction. In fact, F4/Gcd14 has sequence similarity to a family of putative protein carboxyl methyltransferases from Eubacteria and Archaea (for example, *M. leprae*, GenBankTM accession number 2145817, aligns with an E value of 10^{-8} with F4/Gcd14). Carboxyl methylation can be reversible and is known to play a regulatory role in several systems (63).

Slow growth, as seen in the mutants AN7 and AN8, indicates a general metabolic defect (64), and there are various ways in which methylation may be involved in this phenotype. For example, disruption of the Erg6 methyltransferase, which is involved in the biosynthesis of ergosterol, leads to slow growth of the resulting yeast mutants (65, 66). Although the reason for slow growth is not known, it can be hypothesized that the altered sterols that are formed in the absence of the Erg6 methyltransferase may be unable to fulfill the normal cellular functions of ergosterol, such as its role in membrane structure (67). Interestingly, the gene disrupted in the slow-growing AN8 mutant (*F8*) has the sequence NPPY immediately following motif II which matches the signature motif of adenine methyltransferases (16, 68). It is conceivable that a defect in RNA methylation in the AN8 mutant could affect protein synthesis and lead to slow growth.

This initial study of seven of the putative methyltransferases has revealed one essential gene, two genes whose disruption leads to slow growth, and a novel arginine methyltransferase. There are a number of methyltransferases in yeast predicted to exist based on the presence of methyl groups on various substrates or on specific enzymatic activities that have been demonstrated (Table I). We hope to be able to match the other putative methyltransferases in this study with these activities. Studies are currently under way in our laboratory to analyze the pattern of methylation of proteins, nucleic acids, lipids, and small molecules of the remaining mutants described in this work.

Acknowledgments—We thank Dr. Ron Kagan for helpful comments on the manuscript and for sharing his expertise during this study. We thank Dr. Catherine Clarke for providing the pCH1 plasmid and Dr. Gregory Payne for providing the GPY278 yeast strain.

REFERENCES

- Goffeau, A. *et al.* (1997) *Nature* **387**, (suppl.) 1–105
- Schluckebier, G., O'Gara, M., Saenger, W., and Cheng, X. (1995) *J. Mol. Biol.* **247**, 16–20
- Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., and Roberts, R. J. (1993) *Cell* **74**, 299–307
- Labahn, J., Granzin, J., Schluckebier, G., Robinson, D. P., Jack, W. E., Schildkraut, I., and Saenger, W. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10957–10961
- Vidgren, J., Svensson, L. A., and Liljas, A. (1994) *Nature* **368**, 354–358
- Reinisch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. (1995) *Cell* **82**, 143–153
- Fu, Z., Hu, Y., Konishi, K., Takata, Y., Ogawa, H., Gomi, T., Fujioka, M., and Takusagawa, F. (1996) *Biochemistry* **35**, 11985–11993
- Hodel, A. E., Gershon, P. D., Shi, X., and Quiocho, F. A. (1996) *Cell* **85**, 247–256
- Yu, L., Petros, A. M., Schnuchel, A., Zhong, P., Severin, J. M., Walter, K., Holzman, T. F., and Fesik, S. W. (1997) *Nat. Struct. Biol.* **4**, 483–489
- Bussiere, D. E., Muchmore, S. W., Dealwis, C. G., Schluckebier, G., Nienaber, V. L., Edalji, R. P., Walter, K. A., Ladrer, U. S., Holzman, T. F., and Abad-Zapatero, C. (1998) *Biochemistry* **37**, 7103–7112
- Djordjevic, S., and Stock, A. M. (1998) *Nat. Struct. Biol.* **5**, 446–450
- Ingrosso, D., Fowler, A. V., Bleibaum, J., and Clarke, S. (1989) *J. Biol. Chem.*

- 264, 20131–20139
13. Lauster, R., Trautner, T. A., and Noyer-Weidner, M. (1989) *J. Mol. Biol.* **206**, 305–312
 14. Posfai, J., Bhagwat, A. S., Posfai, G., and Roberts, R. J. (1989) *Nucleic Acids Res.* **17**, 2421–2435
 15. Kagan, R. M., and Clarke, S. (1994) *Arch. Biochem. Biophys.* **310**, 417–427
 16. Malone, T., Blumenthal, R. M., and Cheng, X. (1995) *J. Mol. Biol.* **253**, 618–632
 17. Hamahata, A., Takata, Y., Gomi, T., and Fujioka, M. (1996) *Biochem. J.* **317**, 141–145
 18. Mao, X., Schwer, B., and Shuman, S. (1995) *Mol. Cell. Biol.* **15**, 4167–4174
 19. Hardwick, K. G., and Pelham, H. R. (1994) *Yeast* **10**, 265–269
 20. Barkovich, R. J., Shtanko, A., Shepherd, J. A., Lee, P. T., Myles, D. C., Tzagoloff, A., and Clarke, C. F. (1997) *J. Biol. Chem.* **272**, 9182–9188
 21. Dibrov, E., Robinson, K. M., and Lemire, B. D. (1997) *J. Biol. Chem.* **272**, 9175–9181
 22. Gary, J. D., Lin, W.-J., Yang, M. C., Herschman, H. R., and Clarke, S. (1996) *J. Biol. Chem.* **271**, 12585–12594
 23. Clarke, C. F., Williams, W., and Teruya, J. H. (1991) *J. Biol. Chem.* **266**, 16636–16644
 24. Lafontaine, D., Delcour, J., Glasser, A. L., Desgres, J., and Vandenhoute, J. (1995) *J. Mol. Biol.* **241**, 492–497
 25. Hansen, J., Muldbjerg, M., Cherest, H., Surdin-Kerjan, Y. (1997) *FEBS Lett.* **401**, 20–24
 26. Hrycyna, C. A., Sapperstein, S. K., Clarke, S., and Michaelis, S. (1991) *EMBO J.* **10**, 1699–1709
 27. Mattheakis, L. C., Shen, W. H., and Collier, R. J. (1992) *Mol. Cell. Biol.* **12**, 4026–4037
 28. Kodaki, T., and Yamashita, S. (1987) *J. Biol. Chem.* **262**, 15428–15435
 29. Sirum-Connolly, K., and Mason, T. L. (1993) *Science* **262**, 1886–1889
 30. Ellis, S. R., Morales, M. J., Li, J. M., Hopper, A. K., Martin, N. C. (1986) *J. Biol. Chem.* **261**, 9703–9709
 31. Sindhuphak, T., Hellman, U., and Svensson, I. (1985) *Biochim. Biophys. Acta* **824**, 66–73
 32. Liao, H. N., and Sherman, F. (1979) *J. Bacteriol.* **138**, 853–860
 33. Shapiro, S. K., Almenas, A., and Thomson, J. F. (1965) *J. Biol. Chem.* **240**, 2512–2518
 34. Lobet, Y., Lhoest, J., and Colson, C. (1989) *Biochim. Biophys. Acta* **997**, 224–231
 35. Cavallius, J., Zoll, W., Chakraborty, K., and Merrick, W. C. (1993) *Biochim. Biophys. Acta* **1163**, 75–80
 36. Xie, H., and Clarke, S. (1993) *J. Biol. Chem.* **268**, 13364–13371
 37. Xie, H., and Clarke, S. (1994) *J. Biol. Chem.* **269**, 1981–1984
 38. Gehrke, C. W., and Kuo, K. C. (1989) *J. Chromatogr.* **471**, 3–36
 39. Klootwijk, J., and Planta, R. J. (1973) *Eur. J. Biochem.* **39**, 325–333
 40. Hrycyna, C. A., Yang, M. C., and Clarke, S. (1994) *Biochemistry* **33**, 9806–9812
 41. Zobel-Thropp, P., Gary, J. D., and Clarke, S. (1998) *J. Biol. Chem.* **273**, 29283–29286
 42. Gary, J. D., and Clarke, S. (1998) *Prog. Nucleic Acids Res. Mol. Biol.* **61**, 65–131
 43. Isbrandt, D., and von Figura, K. (1995) *Biochim. Biophys. Acta* **1264**, 265–267
 44. Berben, G., Dumont, J., Gilliquet, V., Bolle, P., and Hilger, F. (1991) *Yeast* **7**, 475–477
 45. Lussier, M., Camirand, A., Sdicu, A. M., and Bussey, H. (1993) *Yeast* **9**, 1057–1063
 46. Robinson, J. S., Klionsky, D. J., Banta, L. M., and Emr, S. D. (1988) *Mol. Cell. Biol.* **8**, 4936–4948
 47. Hrycyna, C. A., and Clarke, S. (1993) *Biochemistry* **32**, 11293–11301
 48. Christianson, T. W., Sikorski, R. S., Dante, M., Shero, J. H., and Hieter, P. (1992) *Gene (Amst.)* **110**, 119–122
 49. Hsu, A. Y., Poon, W. W., Shepherd, J. A., Myles, D. C., and Clarke, C. F. (1996) *Biochemistry* **35**, 9797–9806
 50. Rothstein, R. J. (1983) *Methods Enzymol.* **101**, 202–211
 51. Elble, R. (1992) *BioTechniques* **13**, 18–20
 52. Sherman, F., and Hicks, J. (1991) *Methods Enzymol.* **194**, 21–37
 53. Tollervey, D., Lehtonen, H., Jansen, R., Kern, H., and Hurt, E. C. (1993) *Cell* **72**, 443–457
 54. Kiss-Laszlo, Z., Henry, Y., Bachelier, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996) *Cell* **85**, 1077–1088
 55. Farooqui, J., Kim, S., Paik, and W. K. (1980) *J. Biol. Chem.* **255**, 4468–4473
 56. Dujon, B. (1998) *Electrophoresis* **19**, 617–624
 57. Rieger, K. J., Kaniak, A., Coppee, J. Y., Aljinovic, G., Baudin-Baillieu, A., Orłowska, G., Gromadka, R., Groudinsky, O., Di Rago, J. P., and Slonimski, P. P. (1997) *Yeast* **13**, 1547–1562
 58. DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997) *Science* **278**, 680–686
 59. Cuesta, R., Hinnebusch, A. G., and Tamame, M. (1998) *Genetics* **148**, 1007–1020
 60. Hinnebusch, A. G. (1988) *Microbiol. Rev.* **52**, 248–273
 61. Rolfes, R. J., and Hinnebusch, A. G. (1993) *Mol. Cell. Biol.* **13**, 5099–5111
 62. Harashima, S., and Hinnebusch, A. G. (1986) *Mol. Cell. Biol.* **6**, 3990–3998
 63. Clarke, S. (1993) *Curr. Opin. Cell Biol.* **5**, 977–983
 64. Hampsey, M. (1997) *Yeast* **13**, 1099–1133
 65. Welihinda, A. A., Beavis, A. D., and Trumbly, R. J. (1994) *Biochim. Biophys. Acta* **1193**, 107–117
 66. Palermo, L. M., Leak, F. W., Tove, S., and Parks, L. W. (1997) *Curr. Genet.* **32**, 93–99
 67. Parks, L. W., Bottema, C. D., Rodriguez, R. J., and Lewis, T. A. (1985) *Methods Enzymol.* **111**, 333–346
 68. Klimauskas, S., Timinskas, A., Menkevicius, S., Butkiene, D., Butkus, V., and Janulaitis, A. (1989) *Nucleic Acids Res.* **17**, 9823–9832