# A conserved motif in the yeast nucleolar protein Nop2p contains an essential cysteine residue

Michelle KING, Duy TON and Kent L. REDMAN[1]

Department of Biochemistry and Molecular Biology, Indiana University School of Medicine-Fort Wayne, 2101 Coliseum Boulevard East, Fort Wayne, IN 46805, U.S.A.

Nop2p is an essential nucleolar protein in *Saccharomyces cerevisiae* that is involved in large ribosomal subunit assembly. It has substantial homology with human p120, the proliferation-associated nucleolar antigen that is overexpressed in many human cancers. A motif containing an invariant Pro–Cys dipeptide is found in Nop2p, p120 and the bacterial Fmu proteins. A total of nine conserved residues, including $Pro^{423}$ and $Cys^{424}$, were individually altered in Nop2p by site-directed mutagenesis. Nop2p function was abolished by conversion of $Cys^{424}$ into either alanine or serine. All of the other Nop2p mutations tested sustained yeast viability, including glycine replacement of $Pro^{423}$ and the conversion of a second conserved cysteine into alanine. The crucial role of $Cys^{424}$ in Nop2p is intriguing, due to the critical roles that cysteine residues adjacent to a proline have in a number of nucleotide-modifying enzymes.

Key words: p120, Fmu, ribosome assembly, DNA methyltransferases.

## INTRODUCTION

The altered nucleolar structure typical of cancer cells has stimulated interest in proteins localized to this region of the cell. Human p120, a 120 kDa proliferation-associated nucleolar antigen, and Nop2p from *Saccharomyces cerevisiae* are structurally related nucleolar proteins that were discovered independently using nuclear-targeted monoclonal antibodies [1,2]. It is likely that all eukaryotic organisms express a p120-like protein, because cDNA or genomic sequences encoding related proteins have been identified from the mouse [3], rice (Genbank #D15256) and the parasitic protozoan *Plasmodium falciparum* (Genbank #T02742).

The expression of Nop2p or p120 is rapidly induced following growth stimulation [2,4], making the overproduction of p120 in many types of human cancer potentially significant. The amount of p120 polypeptide expressed correlates with the tumour stage [5,6] and may provide a prognostic indicator for breast cancer survival [7,8]. Whole animal and cell culture experiments indicate a possible link between mammalian p120 expression and tumour growth rate [3,9,10]. Indeed, the apoptosis of HeLa cells following p120 antisense treatment suggests that p120 expression is required for growth of mammalian cells [11], and this concept is supported by the indispensability of the *NOP2* gene in *S. cerevisiae* [2]. As anticipated from their nucleolar location, Nop2p and p120 appear to be involved in the assembly of ribosomes. Human p120 expressed in insect cells is associated with pre-ribosomal particles [12]. Nop2p is required for 60 S subunit maturation in yeast, possibly having a role in one or more steps of rRNA processing [13]. No specific biochemical function has been assigned to either Nop2p or p120, although p120 has been proposed to be a rRNA methyltransferase (MTase) [14].

The adjacent *fmu* and *fmv* genes of *Escherichia coli* were initially thought to encode separate proteins, but actually represent a single open reading frame, encoding a protein that we will call Fmu. Koonin [14] observed similarity between mammalian p120 proteins and the amino acid sequence of Fmu. A conserved region that resembles an *S*-adenosylmethionine (SAM)-binding motif was noted and additional domains appeared to be conserved, even with the limited number of sequences that were available for comparison [14]. On the basis of the SAM motif, Koonin proposed that p120 is an rRNA 2′-hydroxy MTase [14]. Owing to an interest in the role that rRNA methylation plays in ribosome assembly, we began studies of yeast Nop2p. The comparison of additional Fmu sequences with Nop2p and p120 revealed stronger conservation of a second region, termed motif II, than observed for the proposed SAM-binding site. Site-directed mutagenesis of motif II resulted in the identification of an essential cysteine residue in Nop2p that is conserved in human p120 and the bacterial Fmu proteins.

## MATERIALS AND METHODS

### Materials

Sequence alignments were generated using the MACAW program version 2.0.5win16 provided by the National Center for Biological Information [15]. Synthetic oligonucleotides for mutagenesis were purchased from Integrated DNA Technologies (Coralville, IA, U.S.A.) and were purified by PAGE before use. Restriction enzymes were obtained from New England Biolabs, Inc. (Beverly, MA, U.S.A.), except for *Avr*II (*Bln*I) that was purchased from Boehringer Mannheim (Indianapolis, IN, U.S.A.). T4 DNA ligase and *Taq* polymerase were acquired from Promega. 5-Fluoro-orotic acid (5-FOA) was purchased from Toronto Research Chemicals Inc. (North York, ON, Canada). Large-scale plasmid DNA preparations utilized Wizard Plus Maxiprep columns from Promega (Madison, WI, U.S.A.), whereas small-scale plasmid isolations were done with Qiagen (Chatsworth, CA, U.S.A) QIAprep spin columns. Monoclonal antibody HA.11, prepared against the haemagglutinin epitope (CYPYDVPDYASL), was purchased from the Berkeley Antibody Company (Richmond, CA, U.S.A). Horseradish-peroxidase-coupled goat anti-mouse antibodies were obtained from Jackson ImmunoResearch Laboratories, Inc. (West Grove, PA, U.S.A.) and reconstituted by following the manufacturer's in-

Abbreviations used: ds, double-stranded; 5-FOA, 5-fluoro-orotic acid; MTase, methyltransferase; NBC, nitrogen base–casein media; SAM, *S*-adenosylmethionine; TCA, trichloroacetic acid.

[1] To whom correspondence should be addressed (e-mail redman@ipfw.edu).

**Table 1 Sense strand oligonucleotides used for site-directed mutagenesis**

Sequences are shown for the sense strand oligonucleotides of each complementary pair of primers used for site-directed mutagenesis. The plasmids used for templates and the restriction enzyme used to identify mutant clones are listed for each mutant allele. Underlined bases do not match the template plasmid sequence.

| Mutant allele | Template plasmid | Restriction site | Oligonucleotide sequence |
|---|---|---|---|
| NOP2a | BpNOP2 | *Avr*II | 5′-GCT AAT ATC CAC CGC CTA GGC TGT ACC AAC-3′ |
| D416A | BptNOP2a | *Nru*I | 5′-CTA AGG TAA TTG GAG GTT TCG CGA GAA TTT TAC TGG ATG CC-3′ |
| L419A | BptNOP2a | *Sty*I[a] | 5′-GGA GGT TTT GAC AGA ATT GCC TTG GAT GCC CCA TGT T-3′ |
| D421A | BptNOP2a | *Nhe*I | 5′-GTT TTG ACA GAA TTT TGC TAG CTG CCC CAT GTT CC-3′ |
| P423G | BptNOP2a | *Nae*I | 5′-CAG AAT TTT ACT GGA TGC CGG CTG TTC CGG TAC TGG T-3′ |
| C424A | BptNOP2a | *Nhe*I | 5′-GAA TTT TAC TGG ATG CCC CAG CTA GCG GTA CTG GTG TTA TCG G-3′ |
| S425A | BpNOP2a | *Nae*I | 5′-CTG GAT GCC CCA TGT GCC GGC ACT GGT GTT ATC GGT-3′ |
| K432A | BptNOP2a | *Bcl*I[a] | 5′-CTG GTG TTA TCG GTG CTG ATC AAT CTG TCA AGG-3′ |
| K437A | BptNOP2a | *Bsm*BI | 5′-GTA AGG ATC AAT CTG TCG CCG TCT CTC GTA CCG AGA AG-3′ |
| C478A | BpNOP2a | *Bgl*I | 5′-GTG TAA TAG TAT ATT CGA CAG CCT CTG TGG CAG TGG AAG AGG ACG-3′ |
| Rev424 | BptC424A | *Nhe*I[b] | 5′-GAA TTT TAC TGG ATG CCC CAT GTA GCG GTA CTG GTG TTA TCG G-3′ |
| C424S | BptC424A | *Nhe*I[b] | 5′-TTA CTG GAT GCC CCA TCT AGC GGT ACT GGT GTT ATC-3′ |

[a] Clones were screened by direct DNA sequence analysis.
[b] Screened by loss of *Nhe*I site previously introduced.

structions. The bicinchoninic acid reagent from Pierce Chemical Company (Rockford, IL, U.S.A.) was used to determine protein concentrations. Chemiluminescence was detected using Fuji RX medical X-ray film. BA79 nitrocellulose membranes with a 0.1-$\mu$m pore size were obtained from Schleicher and Schuell (Keene, NH, U.S.A.).

### Yeast strains and plasmids

The YBH3 strain (*MATa*, *ade2*, *can1–100*, *his3–11*,*15*, *leu2–3*,*112*, *trp1*, *ura3–1*, *nop2::LEU2*) carries the plasmid pJPA40, which has a URA3 marker and a 3-kb genomic fragment containing the *NOP2* gene [13]. Plasmid WpNOP2 was generated by moving the *Kpn*I/*Sac*I fragment of pJPA40 [13] to pRS314, a TRP1 plasmid [16]. BpNOP2 is the Bluescript SK$^+$ plasmid, into which the *Eco*RI/*Kpn*I fragment of pJPA40 was inserted, following digestion with the same pair of enzymes. Plasmid WpNOP2a is identical with WpNOP2, except that the Nop2p codon for Arg$^{392}$ has been changed from CGT to CGC, generating a unique *Avr*II site (see below).

### Site-directed mutagenesis of NOP2

Plasmids were altered with the Quik-Change mutagenesis kit from Stratagene (La Jolla, CA, U.S.A) using two complementary primers. The coding-strand primer used in each case is listed in Table 1. BpNOP2 was used as template for the silent mutation that generates the *Avr*II site, thus creating BpNOP2a. Deletion of the *Sma*I/*Hin*cII fragment of BpNOP2a resulted in a truncated version of the plasmid called BptNOP2a, which seemed to reduce the toxic effect of some NOP2 constructs on *E. coli* host cells. The *Avr*II site was placed into a functional *NOP2* gene by moving the *Eco*RI/*Kpn*I fragment of BpNOP2a to WpNOP2, generating WpNOP2a. All mutations that result in amino acid changes were initially formed in BpNOP2a or BptNOP2a. Clones carrying mutated plasmids were identified by either restriction-enzyme digestion or direct DNA sequence analysis of isolated DNA, as indicated in Table 1. Expression constructs for the various *NOP2* mutant alleles were formed by replacing the *Avr*II/*Nsi*I fragment of WpNOP2a with the corresponding fragment from the altered BpNOP2a or BptNOP2a plasmid. The complete sequence of the 295 bp *Avr*II/*Nsi*I region was determined for each expression construct by DNA sequencing in both directions on an ABI Prism model 377 sequencer using fluorescent terminator chemistry. Primers employed for sequencing the *Avr*II/*Nsi*I region were NOP2sp1 (CACAGGTTGTGTTTTCGCTA) in the forward direction and NOP2sp2 (CGACTAACTTGACATTCGGTCT) in the reverse direction. The synthesis of sequencing primers and the sequence analysis were carried out in the Biochemistry Biotechnology Facility at the Indiana University School of Medicine, Indianapolis, IN, U.S.A. Plasmids encoding haemagglutinin-tagged Nop2p (WpHemNOP2) and Cys$^{424}$ → Ala Nop2p (WpHemC424A) were formed by inserting a double-stranded oligonucleotide into the *Aat*II site near the 5′-end of the NOP2-coding sequence of WpNOP2a and WpC424A, as described previously [2]. The oligonucleotides used were TACCCATATGATGTTCCTGATTACGCTCAACGT (ET-S) for the sense strand and TGAGCGTAATCAGGAACATCATATGGGTAACGT (ET-AS2) for the antisense strand. Clones with inserts in the correct orientation were identified by PCR. Bacterial extracts were amplified using ET-S as the forward primer and the antisense primer used to form the *Avr*II site as the reverse primer, thus generating a 1200 bp product only when the insertion was present and in the correct orientation.

### Functional analysis

Yeast cells were transformed with plasmids using the modified lithium acetate method of Elble [17]. Following transformation of the YBH3 strain with a construct, the ability of the *NOP2* mutant allele to functionally replace the wild-type gene was tested by the plasmid-shuffling method using 5-FOA selection [18]. When growth occurred on 5-FOA medium, the expected phenotype of the selected cells (Ura$^-$, Trp$^+$, Leu$^+$, Ade$^-$, His$^-$) was confirmed by replica-plating several isolated colonies on appropriately supplemented synthetic dextrose media. Growth rates of the various strains were compared in a semi-defined yeast nitrogen base–casein media (NBC) containing 20 g of glucose, 6.7 g of yeast nitrogen base lacking amino acids, 10 g of Bactotryptone and 20 mg of adenine hemisulphate/litre. Growth was monitored by measuring $D_{600}$ of 50 ml cultures grown in 250-ml Nephlo flasks with rapid shaking.

### Analysis of Nop2p by Western blotting

To prepare protein extracts, cells were grown in Nephlo flasks containing 50 ml of NBC media to a $D_{600}$ of 0.3. Proteins were extracted from the cells in the presence of trichloroacetic acid (TCA) as described previously [2], except that the boiled SDS samples were vortex-mixed rather than sonicated. Proteins were separated by PAGE (5–15 % gradient gels) and then transferred to a nitrocellulose membrane as described previously [19]. Immunodetection of haemagglutinin-tagged Nop2p using luminol chemiluminescence was done essentially as reported previously [20], only substituting the appropriate primary and secondary antibodies for Nop2p detection. The protein concentration of the SDS samples was estimated following 25–50-fold dilution into 0.017 % (v/v) deoxycholate and precipitation with an equal volume of 10 % (v/v) TCA to remove reducing reagent that would interfere with the bicinchoninic acid protein assay. Precipitated protein was dissolved in 10 mM Tris/HCl, pH 8.6, containing 0.33 % (w/v) SDS before protein determination.

## RESULTS

### Identification of a highly conserved region

Nop2p has substantial homology with human p120 [2] and significant similarity to the *E. coli* Fmu protein [13]. When *fmu* gene sequences from *Coxiella burnetii* and *Haemophilus influenzae* became available, we initially compared the encoded protein sequences with the Fmu polypeptide from *E. coli* using the MACAW program. The proteins are approx. 33 % identical with regions of similarity covering most of their length. The proposed SAM-binding motif identified by Koonin [14] is well conserved with 13 out of 17 residues being identical in the Fmu sequences; however, the most similar region is comprised of 23 consecutive identical amino acids (Figure 1).

Comparison of Nop2p and p120 with all three Fmu sequences revealed four blocks of similar sequence. Blocks 2–4 overlapped motifs proposed by Koonin [14] and are shown in Figure 1. Although the analysis corroborated the SAM motif and a region previously termed motif 3 [13] that includes a conserved Thr–Cys–Ser tripeptide (Figure 1), the greatest similarity between the Fmu and p120 proteins corresponds to the Fmu 23-amino-acid identity region (Figure 1). Human p120 and Nop2p exactly match the Fmu sequences for 15 out of the 23 positions in the Fmu identity segment. Most of the duplication is within the first 15 amino acids, a region that corresponds to Koonin's motif II [14] (Figure 1). This impressive conservation of motif II suggests that it is functionally important, and this hypothesis was tested by site-directed mutagenesis.

### Mutation of the motif II region

The most conserved amino acids within the Fmu identity block were identified by examining additional Fmu-like sequences as they became available. Five motif II residues were found to be identical in the aligned sequences of Nop2p, p120 and seven

```
                |---------------- BLOCK 2 --------------------|
                                          |-----SAM-------|
Nop2p           EYLAGHYILQAASSFLPVIALDPHENERILDMAAAPGGKTTYISAMMKNTGCVFANDANKSR  382
p120            -------M--G---M---M--A-Q-H------CC------S-MAQL-----VIL-----AE-  421
                    G    Q          L        LD   APG KT              A D    R
E. coli         GFEDGWVTVQDASAQGCMTWLAPQNGEHILDLCAAPGGKTTHILEVAPEA QVVAVDIDEQR  283
H. influenzae   H--E-A-----L---WAA-L-E-K-E-W---A-------------L--Q- N-I-L-VESH-  295
C. burnetii     --NE-YCYI---AG-FAAYL-KLE-NQTV--A-----S--S-----N-HLKTL--I-NNK--  275

                                                |------------ BLOCK 3 ---
                                                 |--Motif II---|
Nop2p           TKSLIANIHRLGCTNTIVCNYDAREFPKVIG      GFDRILLDAPCSGTGVIGKDQSVKV  435
p120            L--VVG-L----V----ISH--G-Q----V-      ----V------------S--PA--T  474
                      N  RL                           FDR LLDAPCS TGVI      K
E. coli         LSRVYDNLKRLGMKATVKQGDGRYPSQWCGEQ     QFDRILLDAPCSATGVIRRHPDIKW  336
H. influenzae   -K--EE--E--NQQ-I-VC--ASK-DE-LA-IGKSAAK=======================-  353
C. burnetii     -N-IKE-IT---LRQEHL-CLLADV--I  D-WSSGEL====================L  332
                                                      *

                ----------|    |---------------- BLOCK 4 ----------------|
                                          |----Motif 3----|
Nop2p           SRTEKDFIQIPHLQKQLLLSAIDSVDCNSKHGGVIVYSTCSVAVEEDEAVIDYALRKR      496
p120            NKD---ILRCA----E---------NAT--T--YL--C---IT---N-W-V----K--      535
                  D                       K  G   Y TCS    E        L
E. coli         LRRDRDIPELA    QLQSEILDAIWPHLKTGGTLVYATCSVLPEENSLQIKAFLQRT      393
H. influenzae   --KET--AQ-V    E--KK--S-L-EK--PN-V-L----------CE--Q---NAH      410
C. burnetii     --QPG--SQYH    -KKLQL-N-L-TV--A--F-L-S------D--EKV-EE--STH      389
                                              *
```

**Figure 1 Nop2p, p120 and Fmu amino acid sequence comparisons**

Sequences of the proteins in three of four blocks (2–4) of similarity identified by MACAW analysis are shown. The first line corresponds to Nop2p and the second line shows the amino acid sequence of human p120, where dashes indicate residues that are identical with Nop2p. The lower three lines represent the Fmu protein sequences from *E. coli*, *H. influenzae* and *C. burnetii*, with dashes indicating identity with the Fmu protein from *E. coli*. The 23-amino-acid identity in the bacterial proteins is indicated by dotted shading. Gaps were inserted to maximize the alignments. Bold letters between the sequences are amino acids common to all five proteins. The locations of the previously proposed SAM motif [14], motif II [14] and motif 3 [13] are shown above the sequences, as are similarity blocks 2–4 identified by the MACAW program. Numbers to the right correspond to the sequence position of the last amino acid on each line; asterisks mark the locations of the conserved cysteine residues 424 and 478 (Nop2p numbering).

```
      Motif II        |-------------|
FMU Identity Block    FDRILLDAPCSATGVIRRHPDIK
      P72943          VDRALLDVPCSGLGTLHRNPDLR
      P76273          FDAILLDAPCSGEGVVRKDPDAL
      Q60343          FDKILLDAPCSGNIIKDKNRNVS
      P71675          FDRVLVDAPCTGLGALRRRPEAR
Mutagenesis Targets    D   L  D  PC*       *     *
      Human p120       FDRVLLDAPCSGTGVISKDPAVK 476
      Nop2p            FDRILLDAPCSGTGVIGKDQSVK
                       ^     ^     ^     ^     ^
                      415   420   425   430   435
```

**Figure 2 Targets for site-directed mutagenesis**

The 23-residue identity block shared by Fmu amino acid sequences from *E. coli*, *C. burnetii* and *H. influenzae* is shown on the first line. Comparisons with four additional hypothetical Fmu sequences, identified by Swiss Protein Database accession numbers, revealed five residues that are conserved in all seven Fmu sequences, as well as in p120 and Nop2p. These residues were selected as targets for site-directed mutagenesis, along with three less well-conserved amino acids indicated by asterisks. Numbers below the sequences refer to the amino acid position in Nop2p, and the number on the right is the position number for the last amino acid in the p120 sequence shown. The region that corresponds to Koonin's motif II is indicated above the sequences.

**Table 2 Nop2p mutations and plasmids**

Plasmids used to express the *NOP2* mutant alleles. Functional constructs are those that can replace a plasmid that carries a wild-type *NOP2* gene (pJPA40) in plasmid-shuffling experiments. The *NOP2a* allele carries only a silent mutation in $Arg^{392}$ that generates a unique *Avr*II site. All of the other constructs contain the *Avr*II site in addition to the indicated mutation that alters the encoded protein. Rev424 contains an altered codon for $Ser^{425}$, but encodes wild-type Nop2p.

| Mutant allele | Expression plasmid | Functional construct |
|---|---|---|
| NOP2a | WpNop2a | Yes |
| D416A | WpD416A | Yes |
| L419A | WpL419A | Yes |
| D421A | WpD421A | Yes |
| P423G | WpP423G | Yes |
| C424A | WpC424A | No |
| C424S | WpC424S | No |
| S425A | WpS425A | Yes |
| K432A | WpK432A | Yes |
| K437A | WpK437A | Yes |
| C478A | WpC478A | Yes |
| Rev424 | WpRev424 | Yes |

Fmu-related proteins (Figure 2). These residues were chosen as targets for site-directed mutagenesis, along with three other amino acids retained in most of the sequences (Figure 2). Plasmids encoding alanine or glycine replacements were constructed and sequenced as described in the Materials and methods section. When plasmids encoding these *NOP2* alleles were transformed into YBH3 cells, seven out of eight were able to replace pJPA40 in plasmid-shuffle experiments, demonstrating that the altered proteins are functional (Table 2). However, the WpC424A plasmid could not replace a plasmid carrying the wild-type gene, suggesting that conversion of $Cys^{424}$ into alanine results in a non-functional form of Nop2p (Table 2 and Figure 3, row b).

Growth rates of the viable motif II mutants were compared with those of cells expressing the wild-type protein, but no significant differences were noted at 30 °C (results not shown). Incubation at either 38 °C or 18 °C resulted in some deviation in growth rates from wild-type, but all of the strains were able to grow at either temperature extreme (results not shown). The most notable response to temperature was shown by the strain expressing $Asp^{416} \rightarrow$ Ala Nop2p. When compared with cells expressing wild-type Nop2p, cells forming $Asp^{416} \rightarrow$ Ala Nop2p grew slower at 38 °C, but faster at 18 °C (Figure 4). Therefore $Asp^{416}$ is not critical for Nop2p activity, but it does influence some aspect of Nop2p function.

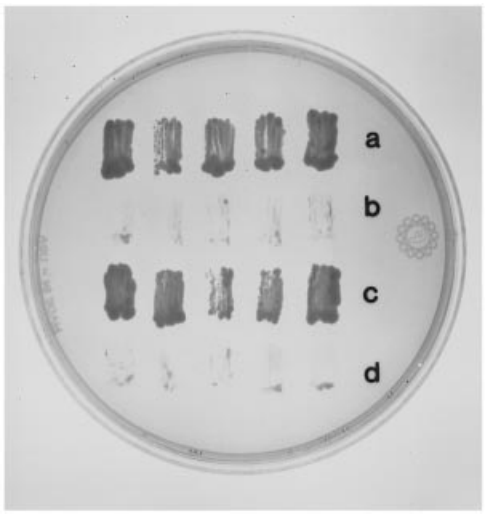


**Figure 3 Plasmid shuffling by 5-FOA selection**

YBH3 cells were transformed with the TRP1 plasmids listed below and patched on to synthetic dextrose media supplemented with adenine, histidine and uracil, in addition to 1 mg/ml 5-FOA. Each row contains cells transformed with a different plasmid and every patch contains cells from a single colony. Cells were transformed with plasmids encoding wild-type Nop2p, WpNOP2a (**a**); $Cys^{424} \rightarrow$ Ala Nop2p, WpC424A (**b**); the reversion mutant containing a low-frequency serine codon, WpRev424 (**c**); or $Cys^{424} \rightarrow$ Ser Nop2p, WpC424S (**d**). Growth occurred where the URA3 plasmid (pJPA40) was lost and Nop2p function was replaced by the TRP1 plasmid. The plate was photographed following incubation at 30 °C for 72 h.

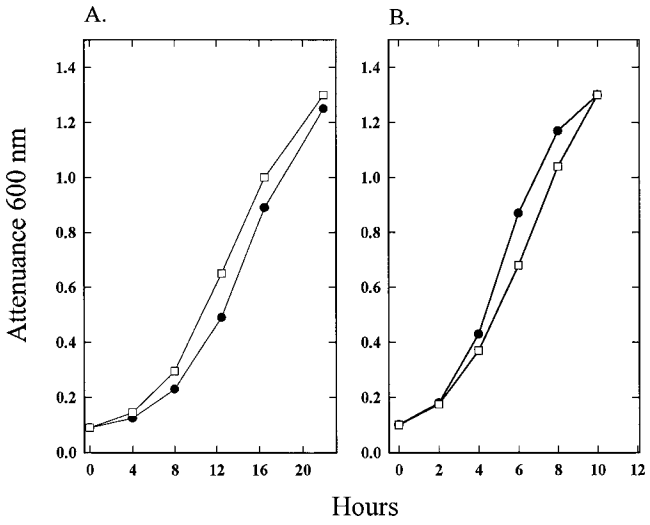


**Figure 4 Growth of cells expressing $Asp^{416} \rightarrow$ Ala Nop2p**

The growth rate of cells carrying the WpD416A plasmid (□) was compared with that of cells expressing wild-type Nop2p from the WpNOP2a plasmid (●). Cells were grown in liquid media at 18 °C (**A**) or 38 °C (**B**), as described in the Materials and methods section.

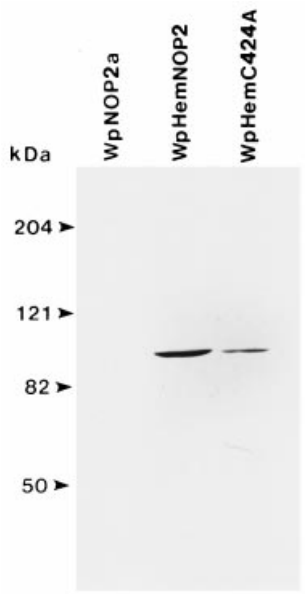


**Figure 5 Expression of haemagglutinin-tagged Cys$^{424}$ → Ala Nop2p**

YBH3 cells were transformed with plasmids that encode the haemagglutinin-tagged wild-type Nop2p (WpHemNOP2), the tagged Cys$^{424}$ → Ala mutant (WpHemC424A) or the wild-type Nop2p lacking a tag (WpNOP2a). Starter cultures were grown in synthetic dextrose media without uracil and tryptophan. For Western blot analysis, cells were grown in NBC media to a $D_{600}$ of 0.3 and the proteins were extracted, separated and detected as described in the Materials and methods section. Each gel lane contained approx. 180 $\mu$g of total protein. The haemagglutinin-tagged proteins were detected using the HA.11 monoclonal antibody at a dilution of 1:5000, and horseradish-peroxidase-coupled secondary antibodies at 1:2000. Luminol chemiluminescence was detected by a 12-s exposure of the membrane to X-ray film. Arrowheads indicate the positions of molecular-mass markers of the sizes shown.

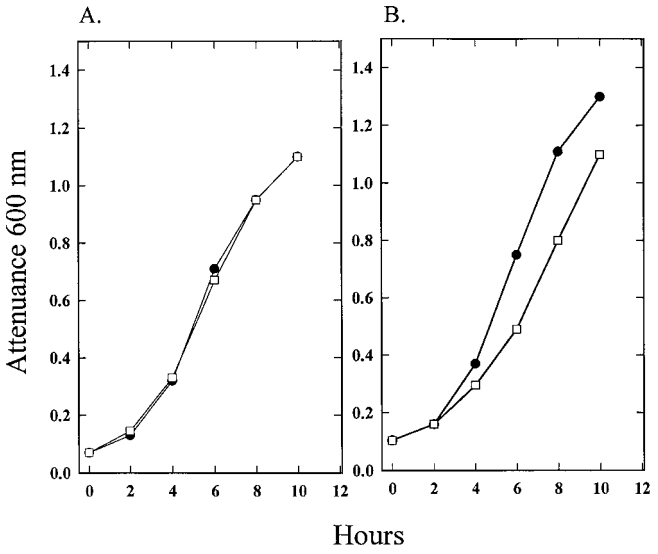


**Figure 6 Temperature sensitivity of cells expressing Cys$^{478}$ → Ala Nop2p**

The growth rate of cells carrying the WpC478A plasmid (□) was compared with that of cells expressing wild-type Nop2p from the WpNOP2a plasmid (●). Cells were grown in liquid media as described in the Materials and methods section at either 30 °C (**A**) or 38 °C (**B**).

### Analysis of the C424A mutation

Creation of an *Nhe*I site allowed rapid screening of clones with the Cys$^{424}$ → Ala mutation; however, formation of the restriction site required conversion of the codon for Ser$^{425}$ from a common sequence (TCC) into an infrequently used degenerate codon (AGC). The intended base changes were confirmed by DNA sequencing, but there remained the minor concern that the less frequently used serine codon (AGC) might limit the rate of Nop2p synthesis. To determine if this codon did restrict Nop2p synthesis, codon 424 was reverted to cysteine, while leaving the Ser$^{425}$ codon as AGC. This mutant (Rev424) was readily able to replace the wild-type *NOP2* gene in plasmid-shuffle experiments (Figure 3, row c), and cells containing the WpRev424 construct as the only source of Nop2p grew at the same rate as cells expressing the wild-type gene (results not shown). Therefore the low-frequency serine codon does not explain the failure of the WpC424A construct to functionally replace the wild-type *NOP2* gene in plasmid-shuffle experiments.

### Confirmation of Cys$^{424}$ → Ala Nop2p expression

We sought to demonstrate the formation of Cys$^{424}$ → Ala Nop2p directly, even though the low-frequency codon for Ser$^{425}$ does not appear to limit Nop2p synthesis. Antibodies raised against Nop2p would not be expected to discriminate between the wild-type and Cys$^{424}$ → Ala forms of Nop2p in strains producing both proteins, but cells forming only the altered protein are not viable. To allow specific detection of Cys$^{424}$ → Ala Nop2p, we generated haemagglutinin-tagged forms of Nop2p that are regulated by the natural promoter, as described in the Materials and methods section. Following transformation of YBH3 cells with plasmids that encode tagged wild-type or tagged Cys$^{424}$ → Ala Nop2p, cultures were maintained on minimal media lacking uracil and tryptophan to force retention of pJPA40 and the TRP1 plasmid carrying the tagged construct. Selection was removed only during the few hours of growth in NBC media before harvesting the cells. Expressed proteins were extracted as indicated in the Materials and methods section and separated on a gradient polyacrylamide gel containing SDS. Haemagglutinin-specific antibodies readily detected the tagged Cys$^{424}$ → Ala and wild-type Nop2p proteins, but did not reveal any protein bands in the extract from cells that expressed the WpNOP2a construct, which lacks a haemagglutinin tag (Figure 5). The tagged Cys$^{424}$ → Ala and wild-type Nop2p proteins are identical in size; however, the amount of tagged Cys$^{424}$ → Ala Nop2p detected was lower than the amount of the tagged wild-type Nop2p (Figure 5). Equal loading of total protein was confirmed by staining the lower half of the blot (stained blot not shown). Therefore the inability of WpC424A to replace pJPA40 in plasmid-shuffle experiments might result from either the functional inadequacy of Cys$^{424}$ → Ala Nop2p or the reduced steady-state level of the altered protein.

### Serine will not replace Cys$^{424}$

The cysteine-to-alanine change could cause a critical structural alteration in Cys$^{424}$ → Ala Nop2p. Although improbable, this possibility was addressed by making the structurally conservative mutation of Cys$^{424}$ → Ser. The plasmid encoding this mutation (WpC424S) would not functionally replace pJPA40 during plasmid-shuffle experiments (Table 2 and Figure 3, row d). The inability of either alanine or serine to replace Cys$^{424}$ indicates a role for the thiol group of this cysteine in some aspect of Nop2p structure or function. Therefore Cys$^{424}$ might function as a catalytic residue, in disulphide-bond formation or in metal-ion binding.

### A second conserved cysteine is not critical for Nop2p function

Nop2p contains seven cysteines, two of which are conserved in human p120. The analysis shown in Figure 1 revealed that $Cys^{424}$ and $Cys^{478}$ of Nop2p are conserved not only in p120, but also in the Fmu proteins. Therefore $Cys^{478}$ could also be functionally important, potentially forming an intramolecular disulphide bond with $Cys^{424}$ or assisting in metal-ion binding. To examine its importance, $Cys^{478}$ was mutated to alanine. $Cys^{478} \rightarrow Ala$ Nop2p is functional, as shown by the ability of the WpC478A plasmid to replace the wild-type gene in plasmid-shuffle experiments (Table 2). At 30 °C, cells expressing $Cys^{478} \rightarrow Ala$ Nop2p grow at the same rate as cells expressing wild-type Nop2p from the WpNOP2a plasmid; however, $Cys^{478} \rightarrow Ala$ Nop2p cells grow slower than control cells at 38 °C (Figure 6). Therefore $Cys^{478}$ is not crucial for Nop2p function, but its conversion into alanine appears to impair either Nop2p function or stability at higher temperatures.

| | | | |
|---|---|---|---|
| Human p120 | **K**V**V**G**V**F**D**R**V**L**L**DA**PC**S**G**TGVIS**K** | 471 | X55504 |
| Nop2p | **K**V**I**G**G**F**D**R**I**L**L**D**A**P**C**S**G**TGVI**G**K** | 432 | U12141 |
| *S. pombe* | **K**K**I**L**K**F**D**R**I**L**A**D**V**P**C**S**G**DGTF**R**K** | 281 | Z99753 |
| HhaI DNA MTase | **K**T**I**PDH**D**ILC**A**GF**PC**QAFSIS**GK** | 89 | J02677 |
| EcoRII DNA MTase | EH**V**PDH**D**VL**LA**GF**PC**QPFSLA**G**V | 194 | X05050 |
| Mouse DNA MTase | PQK**G**DVEMLCGGP**PC**Q**G**FSGMNR | 1119 | X14805 |
| | * | | |
| Shared Residues | K IG D LA **PC** G GK<br>V | | |

**Figure 7 Region of similarity between Nop2p, p120 and the $m^5C$ DNA MTases**

The sequences of p120, Nop2p and a related hypothetical protein from *Schizosaccharomyces pombe* were compared with three $m^5C$ DNA MTase sequences using the MACAW program. A region of similarity was identified that contains the active site cysteine of the $m^5C$ DNA MTase enzymes (*). The similarity block is essentially the same size as the motif IV region shared by the DNA MTases [21], but it is shifted toward the N-terminus by five residues. Numbers following each sequence correspond to the position number of the last amino acid shown and, on the far right, the Genbank database identification of the complete sequence. Shared residues (indicated in bold) are amino acids found in one or more members of both protein families.

## DISCUSSION

The substantial homology between Nop2p and human p120, combined with the ability to readily manipulate the yeast system, provides a sound rationale for using Nop2p as the experimental model for this protein family. The Fmu–p120 similarity noted by Koonin [14] extends to both Nop2p [13] and the Fmu-like proteins of additional bacterial species (Figure 1). Koonin noted a conserved region that resembles a SAM-binding motif, and suggested that further motifs were conserved [14]. Our analysis of additional sequences revealed that Koonin's motif II is more highly conserved than the SAM motif observed initially and we tested the importance of several amino acids within the motif II region. Eight residues in motif II, or immediately adjacent to it, were converted into alanine or glycine, but only mutation of $Cys^{424}$ seriously disrupted the function of Nop2p. The NOP2 mutations that support cell viability allowed cells to grow at the same rate or slightly slower than cells expressing wild-type Nop2p at all of the temperatures tested, with the exception of the $Asp^{416} \rightarrow Ala$ mutation. Relative to wild-type controls, cells expressing the $Asp^{416} \rightarrow Ala$ allele grew slightly slower at 38 °C, at the same rate at 30 °C and slightly faster at 18 °C. These findings do not support any specific conclusions about the role of $Asp^{416}$ in the wild-type protein, although they suggest that $Asp^{416}$ has a negative effect on wild-type Nop2p function at low temperatures, while facilitating Nop2p activity at higher temperatures.

Control experiments revealed that the inability of WpC424A to functionally replace pJPA40 in plasmid-shuffle experiments does not result from changes in the $Ser^{425}$ codon. Expression of $Cys^{424} \rightarrow Ala$ Nop2p was demonstrated using a haemagglutinin-tagged construct, confirming that a full-sized protein is encoded by the modified plasmid. However, the steady-state level of tagged $Cys^{424} \rightarrow Ala$ Nop2p was consistently found to be lower than the wild-type control. It is likely that the reduced level of $Cys^{424} \rightarrow Ala$ Nop2p results from instability of the protein, since overexposure of the Western blot shown in Figure 5 revealed a ladder of bands above the full-sized $Cys^{424} \rightarrow Ala$ Nop2p that are only slightly visible with the tagged wild-type protein (results not shown). Although confirmatory evidence is needed, this suggests that the altered Nop2p might be a target for ubiquitin-mediated degradation. It should be noted that the haemagglutinin tag does not prevent Nop2p function. The plasmid encoding tagged, but otherwise normal, Nop2p can functionally replace the wild-type gene, whereas the tagged $Cys^{424} \rightarrow Ala$ construct cannot (results not shown). Finally, the conservative substitution of $Cys^{424} \rightarrow Ser$ does not generate a protein that will support yeast viability. Our conclusion from this series of experiments is that the thiol group of $Cys^{424}$ is functionally or structurally critical for Nop2p activity.

The crucial function of $Cys^{424}$ may have a structural basis, such as metal-ion binding or intermolecular disulphide-bond formation. $Cys^{478}$ of Nop2p lies outside of motif II in a region that has been designated motif 3 [13]. This residue is in the centre of a Thr–Cys–Ser tripeptide that is conserved in all of the sequences analysed in Figure 1. The conservation of a second cysteine suggested the possibility of an intramolecular disulphide bond between $Cys^{424}$ and $Cys^{478}$ or a metal-ion co-ordination site, but mutagenesis of $Cys^{478}$ revealed that it is not essential for Nop2p function. Therefore the formation of an intramolecular disulphide link between $Cys^{424}$ and $Cys^{478}$ is improbable, otherwise both residues would be critical for Nop2p function.

Another possible role for $Cys^{424}$ is to function as an active site residue. A limited, but intriguing, similarity was found between Nop2p and the MTases that methylate cytosine at the C-5 position ($m^5C$ DNA MTases; Figure 7). These enzymes typically share ten regions of similarity [21]. Comparison of Nop2p-related proteins with a subset of sequenced $m^5C$ DNA MTases revealed only one similar region, but it overlaps with $m^5C$ DNA MTase region IV, which contains the Pro–Cys active site dipeptide (Figure 7). The MTase active site cysteine forms a covalent bond with C-6 of the pyrimidine ring during DNA methylation [22] by a mechanism similar to that of thymidylate synthase, which also has a cysteine in a Pro–Cys dipeptide at its active site [23]. The requirement for Nop2p in 25 S rRNA processing and the known presence of 5-methylcytosine only in the large subunit of rRNA suggested to us that Nop2p could act as an rRNA-specific $m^5C$ MTase, an attractive alternative hypothesis to its role as a 2′-hydroxy MTase, proposed by Koonin [14]. On the basis of the $m^5C$ DNA MTase model, we predicted that stable covalent complexes would form between Nop2p and rRNA containing 5-fluoropyrimidines. We have looked for such complexes under conditions that prevent rRNA degradation, but have found no evidence to support the formation of Nop2p–rRNA cross-links following treatment of yeast with either 5-fluorouracil or 5-fluorocytosine. Therefore the active site hypothesis for $Cys^{424}$ remains speculative.

In the case of *Eco*RII $m^5C$ DNA MTase, it has been reported that converting the proline of the active-site Pro–Cys dipeptide into serine reduced enzyme activity, and it was proposed that proline limits the rotational freedom of the active site cysteine [24]. For Nop2p, the conversion of $Pro^{423} \rightarrow Gly$ had no de-

tectable effect on Nop2p function, as assessed by the obviously crude assay of cellular growth. Owing to the rotational freedom allowed by glycine, any important function of $Pro^{423}$ in Nop2p would not appear to involve limiting the positions available to the adjacent cysteine.

Recently, it has been recognized that other nucleotide-modifying enzymes contain regions that resemble motif IV of the $m^5C$ DNA MTases. The $N^6$-adenine and $N^4$-cytosine DNA MTases have motif IV-like regions; however, the Pro–Cys dipeptide essential for $m^5C$ DNA MTase activity is not conserved in the N-MTases [25]. Pro–Cys dipeptides are conserved in the $m^5C$ DNA MTase motif IV-like regions found in double-stranded (ds)RNA adenosine deaminases [26]. No specific function has been demonstrated for the Pro–Cys dipeptide in this group of deaminases, but mutagenesis of the corresponding cysteine (residue 966) in the human dsRNA adenosine deaminase results in loss of the RNA-editing activity [27].

Additional enzymes contain functionally important Pro–Cys dipeptides that are not in $m^5C$ DNA MTase motif IV-like regions. A motif conserved between cytosine deaminases and the RNA deaminase responsible for apolipoprotein-B mRNA editing contains a Pro–Cys dipeptide that is probably involved in metal binding [28]. Both active site cysteines of the Ada suicide DNA-repair protein from *E. coli* follow proline residues, and the dipeptide is conserved in related proteins from other bacteria, yeast and man [29]. Clearly, cysteines found within Pro–Cys dipeptides are involved in a variety of functions. It will be intriguing to see if the Nop2p/p120/Fmu protein group falls into one of the known enzyme classes that utilize a Pro–Cys dipeptide, or whether these proteins form yet another protein family in which the cysteine of a Pro–Cys dipeptide is functionally important.

We are indebted to John Aris for providing the yeast YBH3 strain, the pJPA40 plasmid and helpful discussions at the start of the project. We also thank Ron Wek for his advice on matters related to *S. cerevisiae* and for his reading of the manuscript. This work was supported by American Cancer IRG #161-K.

## REFERENCES

1 Busch, H. (1990) Cancer Res. **50**, 4830–4838
2 de Beus, E., Brockenbrough, J. S., Hong, B. and Aris, J. P. (1994) J. Cell Biol. **127**, 1799–1813
3 Valdez, B. C., Perlaky, L., Saijo, Y., Henning, D., Zhu, C., Busch, R. K., Zhang, W.-W. and Busch, H. (1992) Cancer Res. **52**, 5681–5686
4 Freeman, J. W., Busch, R. K., Gyorkey, F., Gyorkey, P., Ross, B. E. and Busch, H. (1988) Cancer Res. **48**, 1244–1251
5 McGrath, P. C., Holley, D. T., Hamby, L. S., Mattingly, C. A. and Freeman, J. W. (1994) Surg. Oncol. **3**, 69–77
6 Ueki, T., Nakayama, Y., Sugao, Y., Kohno, K., Matsuo, K., Sugimoto, Y., Yamada, Y., Kuwano, M. and Tsuneyoshi, M. (1997) Hum. Pathol. **28**, 74–79
7 Freeman, J. W., McGrath, P., Bondada, V., Selliah, N., Ownby, H., Maloney, T., Busch, R. K. and Busch, H. (1991) Cancer Res. **51**, 1973–1978
8 McGrath, P. C., Holley, D. T., Hamby, L. S., Powell, D. E., Mattingly, C. and Freeman, J. W. (1994) Surgery **116**, 616–620
9 Fonagy, A., Swiderski, C., Dunn, M. and Freeman, J. W. (1992) Cancer Res. **52**, 5250–5256
10 Fonagy, A., Swiderski, C., Ostrosky, A. M., Bolton, W. E. and Freeman, J. W. (1994) Cancer Res. **54**, 1859–1864
11 Busch, R. K., Perlaky, L., Valdez, B. C., Henning, D. and Busch, H. (1994) Cancer Lett. **86**, 151–157
12 Ren, Y., Busch, R., Durban, E., Taylor, C., Gustafson, W. C., Valdez, B., Li, Y.-P., Smetana, K. and Busch, H. (1996) Prot. Expr. Purif. **7**, 212–219
13 Hong, B., Brockenbrough, J. S., Wu, P. and Aris, J. P. (1997) Mol. Cell. Biol. **17**, 378–388
14 Koonin, E. V. (1994) Nucleic Acids Res. **22**, 2476–2478
15 Schuler, G. D., Altschul, S. F. and Lipman, D. J. (1991) Proteins **9**, 180–190
16 Sikorski, R. S. and Hieter, P. (1989) Genetics **122**, 19–27
17 Elble, R. (1992) BioTechniques **13**, 18–20
18 Sikorski, R. S. and Boeke, J. D. (1991) Methods Enzymol. **194**, 302–318
19 Redman, K. L. and Rechsteiner, M. (1988) J. Biol. Chem. **263**, 4926–4931
20 Redman, K. L. (1994) Insect Biochem. Mol. Biol. **24**, 191–201
21 Posfai, J., Bhagwat, A. S., Posfai, G. and Roberts, R. J. (1989) Nucleic Acids Res. **17**, 2421–2435
22 Chen, L., MacMillan, A. M., Chang, W., Ezaz-Nikpay, K., Lane, W. S. and Verdine, G. L. (1991) Biochemistry **30**, 11018–11025
23 Santi, D. V., McHenry, C. S., Raines, R. T. and Ivanetich, K. M. (1987) Biochemistry **26**, 8606–8613
24 Kossykh, V. G., Schlagman, S. L. and Hattman, S. (1995) Gene **157**, 125–126
25 Malone, T., Blumenthal, R. M. and Cheng, X. (1995) J. Mol. Biol. **253**, 618–632
26 Hough, R. F. and Bass, B. L. (1997) RNA **3**, 356–370
27 Lai, F., Drakas, R. and Nishikura, K. (1995) J. Biol. Chem. **270**, 17098–17105
28 Navaratnam, N., Bhattacharya, S., Fujino, T., Patel, D., Jarmuz, A. L. and Scott, J. (1995) Cell **81**, 187–195
29 Friedberg, E. C., Walker, G. C. and Siede, W. (1995) DNA Repair and Mutagenesis, pp. 91–133, American Society for Microbiology, Washington D. C.

Received 8 June 1998/28 September 1998; accepted 16 October 1998