3290–3294 Nucleic Acids Research, 1995, Vol. 23, No. 16

Mapping of the 13 pseudouridine residues in *Saccharomyces cerevisiae* small subunit ribosomal RNA to nucleotide resolution

Andrey Bakin and James Ofengand*

Roche Institute of Molecular Biology, Roche Research Center, Nutley, NJ 07110, USA

Received April 3, 1995; Revised and Accepted July 12, 1995

ABSTRACT

The number and location of all of the pseudouridine (Ψ) residues in *Saccharomyces cerevisiae* small subunit (SSU) ribosomal RNA have been determined by a reverse transcriptase sequencing method [Bakin, A. and Ofengand, J., 1993, *Biochemistry*, 32, 9754–9762]. Thirteen residues were found in addition to the previously described m^{1}acp^{3}Ψ1189. The residues were scattered throughout the molecule with three being in expansion segments. No Ψ was found in the three highly conserved single-stranded sequence elements common to all SSU RNAs. Specifically, Ψ563, the analog of Ψ516 (*Escherichia coli*) and Ψ517 (*Bacillus subtilis*) were not found. Eight of the Ψ were located identically to those in mammalian SSU RNA and three were near to mammalian Ψ residues in the secondary structure. There was no discernible correlation between the sites for Ψ and the known locations of the methylated nucleosides as exists in large subunit (LSU) RNAs. Comparison of the structural context in which Ψ was found in SSU RNA with that in LSU RNA showed a differential bias suggestive of possible different roles for Ψ in the two rRNAs. This work also identified the locations of three putative new modified bases in SSU rRNA, and revealed 15 sequence differences between the yeast strain used here and the reported sequence (GenBank, Release 88.0, April 1995, accession. no. J01353).

INTRODUCTION

Although pseudouridine (Ψ; 5-β-D-ribofuranosyluracil) has been known for over 35 years (1–4) and is commonly found in both small subunit (SSU) and large subunit (LSU) ribosomal RNA of prokaryotes and eukaryotes (5), its role in rRNA structure and/or function is still not understood. This is undoubtedly due to the fact that until recently the position of most Ψ residues was unknown because Ψ could only be located by labor-intensive conventional sequencing procedures. This problem was solved by the development of a rapid primed reverse transcriptase sequencing method (6). Using this approach, we recently reported the precise location of all of the Ψ residues in *Escherichia coli* ribosomal RNA (6–8). This information has indicated the potential importance of Ψ in rRNA function. The single Ψ in *E.coli* SSU RNA was found in the '525' loop (7) known to be important for codon recognition (9, and references therein), and the eight Ψ plus one modified Ψ in *E.coli* LSU RNA were found to be clustered in three areas, all of which were at or near the peptidyl transferase center (6,8). We also located the 30 Ψ in *Saccharomyces cerevisiae* LSU RNA and found them to also be clustered at the peptidyl transferase center (8).

In this work, we have completed the analysis of *S.cerevisiae* rRNA by determining the sites of all of the Ψ in the SSU rRNA. We found 13 Ψ residues in this RNA and report here their precise location.

MATERIALS AND METHODS

Materials

S.cerevisiae strain D273-10B/A1 was the gift of Prof. Alexander Tzagoloff, Columbia University. Total yeast RNA was prepared from an overnight culture by the procedure of Fitch, *et al.* (10). Primers and other reagents used for sequencing were as described (6).

Methods

Sequence localization of pseudouridine residues was carried out by chemical modification using the N-cyclohexyl-*N'*-β-(4-methylmorpholinium)-ethylcarbodiimide *p*-tosylate (CMC)-alkali (OH) method and reverse transcription as described previously (6) except that the time of CMC treatment was decreased to 15 min (11). Primers complementary to *S.cerevisiae* SSU RNA residues 186–203, 402–420, 624–641, 854–870, 1083–1099, 1226–1243, 1326–1342, 1522–1538 and 1763–1779 were used.

RESULTS

Exact localization of the 13 Ψ residues

Application of the CMC-OH method to the localization of all of the Ψ residues in *S.cerevisiae* SSU RNA was straightforward.

* To whom correspondence should be addressed

The entire molecule was screened except for the 3′-43 residues. With the modified procedure, no doublet bands were encountered, and all bands could be unequivocally assigned by the CMC-OH procedure alone. The locations of the Ψ residues are listed in Table 1 and shown in Figure 1. Table 1 also indicates the structural context (8) in which each Ψ occurs.

Table 1. Location of the pseudouridine residues in yeast SSU ribosomal RNA

Position numbers		
106 (e)	466 (e)	1000 (d)
120 (d)	632 (b)	1179 (e)
211 (c)	759 (?)	1185 (b)
302 (a)	766 (?)	1289 (a)
		1414 (b)

The location of the Ψ residues was determined as described in Materials and Methods. The entire molecule was screened except for the 3′-terminal 43 nucleotides. The letters in parentheses identify the structural class (8) in which each Ψ is found, according to the secondary structure of Figure 1. (?), residues located in a region for which the secondary structure has not been determined.

Corrections to the primary sequence of *S cerevisiae* SSU RNA

In the course of these experiments, several differences from the primary sequence reported in the literature (12,19) were noted. They are listed in Table 2. Note that although the sequence reported by Gutell (12) is a modified version of (19), we found that our sequence sometimes agreed with the Gutell sequence, sometimes with the Mankin, *et al.* (19) sequence, and sometimes with neither. These changes may be true errors in the sequence compilation or they may be due to strain differences. Nucleotide numbering in this article is based on the revised sequence (see Fig. 1).

Other putative modified bases

As was the case with *E.coli* LSU RNA (6), the reverse transcriptase technique used here could also detect other bases modified such that base-pairing was blocked and a reverse transcriptase stop was produced. These putative modified bases are listed in Table 3. The position of the stop is shown as well as the deduced location of the base which should be one base 3′ to the stop. The known modified bases, $m^1acp^3\Psi$1189 and m^7G1574 (13) were detected. Three additional stops were found as indicated in the Table. Note that the stop at G1268 could correspond either to a modified U, such as m^3U, at 1267 or to a m^1G1268, and likewise the band at G1533 could be either a modified C1532 or m^1G1533 since we previously observed that for unknown reasons the reverse transcriptase stop at m^1G does not occur one base 3′ to the G, but at the m^1G itself. A modified C1532 could be ac^4C as this base has been reported to occur in yeast SSU RNA (14). Although ac^4C is alkali-labile, the alkaline conditions used here, 37°C for 4 h at pH 10.4, would only be expected to result in hydrolysis of <6% of the ac^4C based on the reported half-life of 5 min in 0.5 M KOH (14).

Table 2. Corrections to the primary sequence of *Saccharomyces cerevisiae* SSU ribosomal RNA[a]

GenBank[a]	Gutell (12)	This work[b]
C150	C150	U150
G235	C235	G235
	–	A884
	–	A899
A943	G943	A945
A962	G962	A964
A982	G982	A984
G983	A983	G985
U995	C995	U997
U1000	U1000	A1002
G1002	A1002	A1004
	G1040	G1042
	G1041	G1043
U1159	C1161	C1163
A1163	C1165	A1167
A1168	A1170	C1172
G1169	G1171	A1173
C1170	C1172	–
C1171	C1173	–
U1172	U1174	–
	U1181	U1180
A1444	A1447	–
A1742	G1745	A1743

[a]GenBank, Release 88.0, April 1995, accession no. J01353 (19). The same sequence, except for the replacement of G235 by C, but based on the same references as J01353, is also listed as no. V01335.
[b]Our revision of the GenBank 18S rRNA sequence (Release 88.0, April 1995, accession no. J01353). This strain is D273-10B/A1. Residues 2–360, 390–535, 570–838, 840–974, 977–1186, 1190–1508, 1530–1622, 1625–1636, 1641–1752 have been sequenced, accounting for a total of 1644 nucleotides (91%) of the 18S rRNA sequence.

Table 3 Reverse transcriptase stops indicative of base-modified nucleotides in yeast SSU ribosomal RNA

Stop	Modified base	Literature
C976	A975	–
C1190	U1189	$m^1acp^3\Psi$
G1268	U1267 or m^1G1268	–[b]
G1533	C1532 or m^1G1533	–
A1575[a]	G1574	m^7G

[a]Revealed only after the alkaline treatment used in the Ψ detection assay.
[b]Corresponds closely in the secondary structure to m^2G1207 in *E.coli*.

DISCUSSION

In this work, we describe the location of all 13 of the Ψ residues in *S.cerevisiae* ribosomal SSU RNA. In conjunction with our previous localization of all 30 Ψ in the LSU rRNA of this organism (8), this report provides a complete description of the

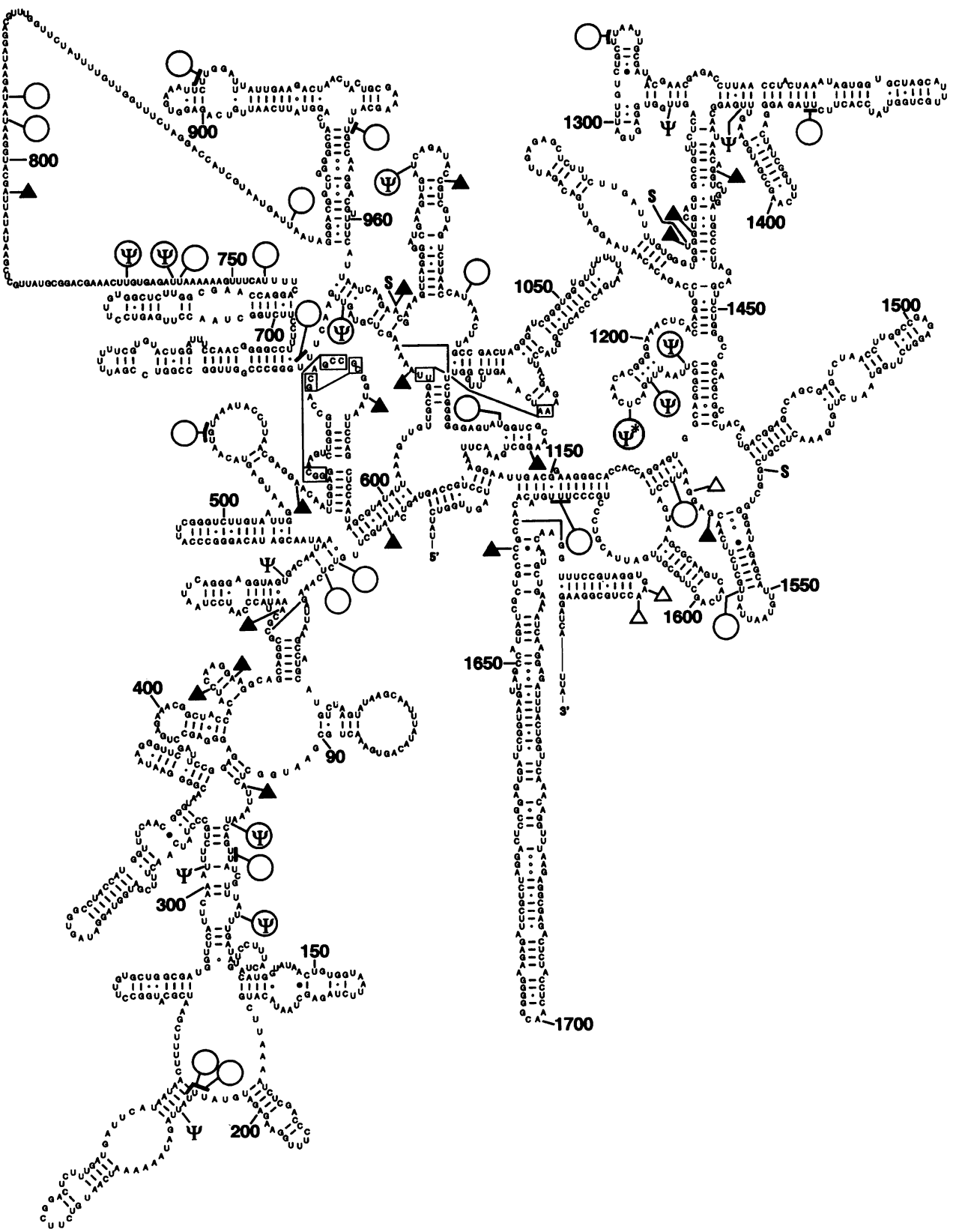

Figure 1. Positions of the pseudouridines and methylated residues in yeast SSU RNA, and the pseudouridines in human SSU RNA. The secondary structure is modified from Gutell (12) to include the corrections in the sequence and secondary structure in Table 2. Ψ, pseudouridines in *S.cerevisiae*; ○, mammalian pseudouridine positions (5); Ⓨ, pseudouridines at the same site in both *S.cerevisiae* and mammals; Ψ*, $m^1acp^3\Psi$; Δ, base-methylated, and ▲, 2′-O-methyl, nucleosides in *S.cerevisiae* (5); S, putative modified base identified by reverse transcriptase stop (Table 3).

location of all 43 Ψ residues in the large rRNAs of *S.cerevisiae*. This is the first eukaryotic species for which this has been done. An equivalent complete analysis of the nine Ψ plus one modified Ψ present in the prokaryote, *E.coli*, was reported by us earlier (6–8).

Previous work on the SSU RNA of *S.carlsbergensis* by Brand, *et al.* (15) estimated the number of Ψ residues at 11, and provided some surrounding nucleotide sequence information for four of them. We found the number of Ψ in *S.cerevisiae* to be 13 by sequencing the entire RNA. From the sequence data provided by Brand, *et al.*, it was possible to unequivocally confirm the identity and location of the Ψ residues described by them. Spot 46′ is Ψ120, 52′ is Ψ211, 60′ is Ψ759 and 74a′ is Ψ302 (there are seven U residues in this oligonucleotide in *S.cerevisiae* instead of the six reported in *S.carlsbergensis*).

In the LSU RNA of *S.cerevisiae*, only one of the 30, or 3%, of the Ψ residues was found in an expansion segment (8), the definition of expansion segment being that of Raué, *et al.* (13,16). By contrast, in the SSU RNA, three out of 13, or 23%, of the Ψ residues were found in expansion segments. Occurrence in an expansion segment, a region which varies from organism to organism, is generally taken to indicate that the element in question is not necessary for a function common to all ribosomes and therefore is probably not of general significance. The fact that 23% of the Ψ in SSU RNA fall in this category compared to 3% of the LSU RNA Ψ suggests that at least some of the Ψ in SSU RNA may be there either adventitiously or for a specialized purpose. An adventitious occurrence could arise if a Ψ synthase specific for a site in another class of RNA would also recognize a site in ribosomal RNA by virtue of its similarity or identity to the 'true' site.

Likewise, the structural environment in which the Ψ residues were found appears to be biased differently in *S.cerevisiae* SSU RNA than in the LSU RNA. When the number of Ψ residues in the different structural classes were compared for SSU RNA and LSU RNA (Table 4), it appears that LSU RNA favors Ψ–Pu as the terminal base pair of a helix (class A) whereas SSU RNA has twice as much single-stranded Ψ adjacent to a stem (class E) as LSU RNA. Ψ–Pu base pairs at least 3 bp into a stem (class C) are similarly rare in both RNAs, and Ψ next to the terminal base pair (class B) or in a loop (class D) are similarly prominent.

Table 4. Distribution of SSU and LSU RNA Ψ residues into structural classes

Structure class[a]	SSU RNA Number[b]	%	LSU RNA Number[c]	%
A	2	17	8	27
B	3	25	9	30
C	1	8	2	7
D	3	25	7	23
E	3	25	4	13

[a]Structural classes of Ψ sites are as follows (8). A, Pu–Ψ base-pair closing a loop or bulge. B, Pu–Ψ base-pair in a helix one base-pair removed from a loop or bulge. C, Pu–Ψ base-pair in a helix surrounded by at least two base pairs on either side. D, Ψ in loops or single-stranded regions but not adjacent to a base pair. E, Ψ in loops or single-stranded regions adjacent to a base pair.
[b]The 12 Ψ and modified Ψ of yeast found in structured regions (Fig. 1).
[c]The 30 Ψ of yeast (8).

The correlation of the positions of the Ψ residues in *S.cerevisiae*, a single-cell eukaryote, with those of mammals, a multi-cell eukaryote, is strong. As shown in Figure 1, eight sites are identical and three more yeast Ψ are nearby in the secondary structure to those of mammalian Ψ sites. Only two, Ψ1289 and Ψ1414, are not obviously near Ψ sites in mammalian SSU RNA. However, since approximately eight Ψ sites in mammals remain to be determined (5), the correlation could be even greater.

On the other hand, there is no correlation between eukaryotes and eubacteria. Despite the fact that the only Ψ in *E.coli* SSU RNA, Ψ516, is also found in *B.subtilis* (17), there is no Ψ at the equivalent 563 position in *S.cerevisiae* and probably none in mammals either. In addition, no Ψ at this position was found in *H.halobium* SSU RNA. About 70% of the RNA, residues 1–292, 393–611, 913–1220 and 1456–1728 (yeast SSU RNA numbering, see Fig. 1) were screened for Ψ, but none was found (A. Bakin and J. Ofengand, unpublished results). Thus, the Ψ516 site is one of those elements specific to eubacterial RNAs. It clearly is not common to all ribosomal RNAs.

We did, however, find a stop at U911 (*H.halobium* SSU RNA numbering) in the absence of CMC/OH treatment. This indicates the presence of a base modification at U910. U910 corresponds in position to the hypermodified $m^1acp^3Ψ1189$ in yeast, which has also been found in several other eukaryotic SSU RNAs (5). It will be interesting to know the structure of this modified uridine. No other strong stops to reverse transcription were detected in the ~70% of the *H.halobium* SSU RNA that was screened.

Unlike the location of Ψ in yeast LSU RNA which clustered around the peptidyl transferase center (8), the SSU RNA Ψ were neither clustered nor concentrated around the three highly conserved and functionally important elements of SSU RNA, the '525', '1400' and '1500' regions (18). In fact, the Ψ residues were conspicuous by their absence from these three regions. Moreover, while the LSU Ψ positions were correlated with the positions of the methyl groups (8), in SSU RNA the distribution of methyl groups was not detectably correlated with that of the Ψ residues, at least in the secondary structure display.

Overall, it appears that there is no specific function that can yet be attributed to the Ψ residues in SSU rRNA. A discussion of potential roles for Ψ residues in ribosomal RNA has been presented elsewhere (6–8,20–22). Nevertheless, the obtained sequence information has shown that the presence of Ψ in the '525' loop is not universal. Moreover, by mapping all of the Ψ residues to nucleotide resolution, this work has created a basis for studying the formation and function of Ψ in SSU RNA.

ACKNOWLEDGEMENT

We thank Prof. B. E. H. Maden, University of Liverpool, for a helpful discussion of this work and the prior literature.

REFERENCES

1 Yu, C-T. and Allen, F.W. (1959) *Biochim. Biophys. Acta* **32**, 393–406.
2 Scannell, J.P., Crestfield, A.M. and Allen, F.W. (1959) *Biochim. Biophys. Acta* **32**, 406–412.
3 Cohn, W.E. (1959) *Biochim. Biophys. Acta* **32**, 569–571.
4 Cohn, W.E. (1960) *J. Biol. Chem.* **235**, 1488–1498.
5 Maden, B.E.H. (1990) *Progr. Nucl. Acids Res. Mol. Biol.* **39**, 241–303.
6 Bakin, A. and Ofengand, J. (1993) *Biochemistry* **32**, 9754–9762.
7 Bakin, A., Kowalak, J.A., McCloskey, J.A. and Ofengand, J. (1994) *Nucleic Acids Res.* **22**, 3681–3684.

8 Bakin, A.V., Lane, B.G., and Ofengand, J. (1994) *Biochemistry* **33**, 13475–13483.
9 Santer, M., Santer, U., Nurse, K., Bakin, A., Cunningham, P., Zain, M., O'Connell, D. and Ofengand, J. (1993) *Biochemistry* **32**, 5539–5547.
10 Fitch, I., Dahmann, C., Surana, U., Amon, A., Nasmyth, K., Goetsch, L., Byers, B., and Futcher, B. (1992) *Mol. Biol. Cell* **3**, 805–818.
11 Bakin, A. and Ofengand, J. (1995) In, R.Martin (ed.), *Protein Synthesis: Methods and Protocols*. Humana Press, in press.
12 Gutell,R.R. (1993) *Nucleic Acids Res.* **21**, 3051–3054.
13 Raué, H.A., Klootwijk, J. and Musters, W. (1988) *Prog. Biophys. Mol. Biol.* **51**, 77–129.
14 Thomas, G., Gordon, J., and Rogg, H. (1978) *J. Biol. Chem.* **253**, 1101–1105.
15 Brand, R.C., Klootwijk, J., Sibum, C.P. and Planta, R.J. (1979) *Nucleic Acids Res.* **7**, 121–134.
16 Raué, H.A., Musters, W., Rutgers, C.A., Van'T Riet, J. and Planta, R.J. (1990) In: *The Ribosome. Structure, Function, and Evolution.* (eds) Hill, W.E., Dahlberg, A., Garrett, R.A., Moore, P.B., Schlessinger, D. and Warner, J.R., American Society for Microbiology, Washington, D.C., pp. 217–235.
17 Wrzesinski, J., Bakin, A., Nurse, K., Lane, B.G., and Ofengand, J. (1995) *Biochemistry* **34**, 8904–8913.
18 Ofengand,J., Bakin, A., and Nurse, K. (1993) In Nierhaus, K.H., Franceschi, F., Subramanian, A.R., Erdmann, V.A., and Wittman-Liebold, B., (eds), *The Translational Apparatus*. Plenum Press, New York NY, pp 489–500.
19 Mankin, A. S., Skryabin, K.G. and Rubtsov, P.M. (1986) *Gene* **44**, 143–145.
20 Lane, B.G., Ofengand, J. and Gray, M.W. (1992) *FEBS Lett.* **302**, 1–4.
21 Lane, B.G., Ofengand, J. and Gray, M.W. (1995) *Biochimie* **77**, 7–15.
22 Ofengand, J., Bakin, A., Wrzesinski, J., Nurse, K., and Lane, B.G. (1995) *Biochem. Cell Biol.* in press.