Alternative CUG codon usage (Ser for Leu) in *Pichia farinosa* and the effect of a mutated killer gene in *Saccharomyces cerevisiae*

C. Suzuki, T. Kashiwagi and K. Hirayama

National Food Research Institute, 2-1-12, Kannon-dai, Tsukuba, Ibaraki 305-8642 and Central Research Laboratories, Ajinomoto Co., Inc., 1-1, Suzuki-cho, Kawasaki-ku, Kawasaki, Kanagawa 210-8681, Japan

The halotolerant yeast *Pichia farinosa* KK1 strain produces a killer toxin termed SMKT (salt-mediated killer toxin). Mass spectrometry and Edman sequencing of peptides from the mature SMKT and secreted protoxin demonstrate that positions specified by the CUG codon contain unmodified serine (Ser) in *P. farinosa*. In order to express the authentic SMKT product in *Saccharomyces cerevisiae*, which uses the universal genetic code, the three CUG codons corresponding to Ser87, Ser137 and Ser206 in the *SMK1* gene were changed to universal Ser codons by site-directed mutagenesis. The expression of the modified *SMK1* gene was lethal in wild-type cells of *S. cerevisiae* (Suzuki et al., 2000). We also showed that spf1p disruptant cells lacking a P-type ATPase Spf1p is highly resistant to both exogenous SMKT and endogenously expressed *SMK1* products (Suzuki et al., 2000). Heterologous expression with the CUG codon in *S. cerevisiae*, which uses the universal genetic code, would substitute the strongly hydrophobic Leu for polar Ser, which might perturb the three-dimensional structure of the protein or alter its activity or stability. On the other hand, CUG encoding Ser residues are found at non-essential surface sites in some cases (Cutfield et al., 2000). It is therefore necessary to evaluate whether the lethal effect of the *SMK1* product in *S. cerevisiae* is caused by substitutions of the authentic Ser residues specified by the CUG codons for Leu.

In this paper, we provide *in vivo* evidence that the CUG codon is translated as unmodified Ser in *P. farinosa*. The expression of authentic *SMK1* product by mutating the CUG codons to universal Ser codons was lethal in *S. cerevisiae* but the secretion of authentic *SMK1* product was more effective. Our results suggest that CUG encoding Ser residues in the γ region may be involved in the stability of protoxin.

Introduction

The halotolerant yeast *Pichia farinosa* KK1 strain produces a killer toxin termed SMKT (salt-mediated killer toxin). SMKT is a heterodimer consisting of α and β subunits. The amino acid sequences of each subunit have been determined and the killer gene, *SMK1*, was cloned and sequenced (Suzuki and Ninkkuni, 1994). On the other hand, *P. farinosa* KK1 strain secretes protoxin in the medium when the strain is grown in a rich medium containing high concentrations of NaCl. Protoxin gp26 was purified and the amino acid sequences of lysylendopeptidase fragments were partially determined (Suzuki, 1999). The crystal structure of SMKT was determined at 1.8 Å resolution (Kashiwagi et al., 1997). Residue 206 specified by a universal leucine (Leu) codon, CUG, was not likely to be Leu based on the shape of the electron density and was finally determined as serine (Ser) by subsequent re-examination of the amino acid sequences.

Kawaguchi et al. demonstrated that the CUG codon is translated as Ser in *Candida cylindracea* (Kawaguchi et al., 1989). To date, it has been shown by *in vitro* translation experiments that 77 species and seven varieties of *Candida* species used codon CUG as Ser, including *Candida cacaoi*, the anamorph of *P. farinosa*. Non-universal decoding is widely distributed not only in *Candida* species but also in hemiascomycetous yeasts (Sugita and Nakase, 1999). By inferring the genetic code from the frequencies of *Saccharomyces cerevisiae* amino acids corresponding to each codon obtained by the Hemiascomycete random sequence tag project, it was pointed out that the CUG codon is used to specify Ser instead of Leu in *Pichia sorbitophila*, a synonym of *P. farinosa* (Tekaia et al., 2000). However, there is no *in vivo* evidence that the CUG codon is translated as Ser in *P. farinosa*.

We have recently demonstrated that the expression of *SMK1* under the control of a galactose-inducible promoter is lethal in wild-type cells of *S. cerevisiae* (Suzuki et al., 2000). It is therefore necessary to evaluate whether the lethal effect of the *SMK1* product in *S. cerevisiae* is caused by substitutions of the authentic Ser residues specified by the CUG codons for Leu.

In this paper, we provide *in vivo* evidence that the CUG codon is translated as unmodified Ser in *P. farinosa*. The expression of authentic *SMK1* product by mutating the CUG codons to universal Ser codons was lethal in *S. cerevisiae* but the secretion of authentic *SMK1* product was more effective. Our results suggest that CUG encoding Ser residues in the γ region may be involved in the stability of protoxin.

Materials and methods

Yeast strain and media

*P. farinosa* (NFRI 3621) (Suzuki and Ninkkuni, 1989), *Saccharomyces cerevisiae* W303–1A (MATa ade2-1 his3-4, 15 leu2-3, 112 trpl-1 ura3-1 can1-100) and CS701C (spf1::LEU2 in W303A-1A background) (Suzuki and Shimma, 1999) were used. YPD and YPGal/MB selective media were as described previously (Suzuki et al., 2000).

Purification of SMKT and protoxin

SMKT was purified as described previously (Suzuki and Ninkkuni, 1994). SMKT concentrations were determined from the UV absorption at 280 nm with ε<sub>280</sub> = 2.8×10<sup>4</sup> M<sup>–1</sup> cm<sup>–1</sup>. Protoxin gp26 was purified as described previously (Suzuki, 1999).

Amino acid sequence analysis of lysylendopeptidase fragments of SMKT and protoxin was performed as described previously (Suzuki, 1999).

Mass spectrometry

The electrospray ionization (ESI) mass spectrum of the native SMKT solution containing 0.5 μM protein, 1.0% (w/v) acetic acid sequences of each subunit have been determined and the *P. farinosa* mutagenesis. The expression of the modified gene were changed to universal Ser codons by site-directed corresponding to Ser87, Ser137 and Ser206 in the *SMK1* gene were changed to universal Ser codons by site-directed mutagenesis. The expression of the modified *SMK1* gene with universal Ser codons was lethal in *S. cerevisiae*, as well as that of the unmodified *SMK1* gene with the CUG codons. The secretion of protoxin with the authentic amino acid sequence from the modified *SMK1* was significantly increased, whereas the transcription level of *SMK1* was not affected in the presence or absence of CUG codon. Our results provide the first *in vivo* evidence that non-universal decoding of CUG is used in a hemiascomycetous yeast, *P. farinosa*.

**Keywords:** codon usage/CUG codon/killer toxin /*Pichia farinosa*/*Pichia sorbitophila*

© Oxford University Press
acid and 50.0% (w/v) methanol was acquired using a TSQ700 triple-stage quadrupole mass spectrometer equipped with an electrospray ion source (Finnigan, San Jose, CA). The measurement conditions of the ESI mass spectra were almost the same as described previously (Kashwagi et al., 2000).

In order to identify residue 206, peptide mapping of the reduced, carboxymethylated and proteolytic digested β subunit was carried out by liquid chromatography–electrospray ionization ion trap mass spectrometry–mass spectrometry (LC–ESI IT-MS–MS). The Cys residues of SMKT were reduced by incubating 100 µg of SMKT in 100 µl of 0.5 M Tris–HCl buffer (pH 8.1) containing 6 M guanidine, 2 mM ethylenediaminetetraacetic acid (EDTA) and 15 mM dithiothreitol (DTT) at 50°C for 3 h and then at room temperature for 10 h. Next, 10 µl of 0.5 M Tris–HCl buffer (pH 8.1) containing 0.1 M iodoacetic acid was added to this solution. After 30 min, the reduced and 5-carboxymethylated (RCM) β subunit was isolated by reverse-phase liquid chromatography using a Vydac (San Francisco, CA) C4 capillary column packed with LC packing and lyophilized. The RCM-β subunit was digested by lysylendopeptidase (Wako Pure Chemical Industries, Osaka, Japan), under the conditions of a 200:1 molar ratio of substrate to enzyme in 20 µl of 50 mM ammonium bicarbonate buffer (pH 9.0) at 35°C for 8 h. Digestion was stopped by the addition of 30 µl of 0.1% trifluoroacetic acid (TFA). Next, 10 µl of a 5-fold dilution of this solution was subjected to LC–ESI IT-MS–MS. The LC solvents consisted of (A) 0.1% aqueous TFA and (B) acetonitrile–water (9:1) containing 0.095% TFA. The digested peptides of the RCM-β subunit were isolated by 10 min isocratic elution with 98% solvent A, gradient elution from 98% to 2% A (10–50 min) and 10 min isocratic elution with 2% solvent A. ESI IT-MS–MS was performed on a Finnigan LCQ instrument equipped with a Finnigan electrospray ion source. The measurement conditions for LC–ESI IT-MS–MS were almost the same as described previously (Hirayama et al., 1998).

**Construction of plasmids**

Standard molecular manipulations were similar to those described by Sambrook et al. (Sambrook et al., 1989).

Three CTG codons in the ORF of SMK1 subcloned into pUC118 (pCS106) were serially changed to TCG by site-directed mutagenesis using a QuikChange pUC118 (pCS106) were serially changed to TCG or TTC respectively. The introduced mutations were constructed into yeast strains as described previously (Elble, 1992).

**SMK1 in pCS224 was substituted for that containing mutated SMK1 in pCS224 was substituted for that containing mutated SMK1, resulting in pCS262. The plasmids were transformed into yeast strains as described previously (Elble, 1992).**

**Reverse transcriptase polymerase chain reaction (RT-PCR)**

Total RNA was prepared according to the method of Köhrer and Domdey (Köhler and Domdey, 1991). After DNase I digestion, RT-PCR was performed as described previously (Suzuki, 1999).

**Other methods**

Preparation of cell lysates, precipitation of secreted proteins by TCA, electrophoresis and immunoblotting were performed as described previously (Suzuki, 1999). Protein A purified rabbit antiserum against the α subunit, that against the β subunit and horseradish peroxidase-linked anti-rabbit immunoglobulin (Bio-Rad Technologies, Richmond, CA) were used at dilutions of 1:5000, 1:1000 and 1:3000, respectively. For detection, the ECL Plus Western Blotting Detection Reagents (Amersham Pharmacia Biotech, Piscataway, NJ) were used.

**Results and discussion**

To identify the amino acid specified by the CUG codon in *P. farinosa*, the secreted protoxin gp26 were digested by lysylendopeptidase and the purified fragments were examined by Edman sequence analysis. Two peptides that contained the amino acid residue specified by the CUG codon were sequenced as SSDSFSIVHNQGE (residues 83–96) and FSLAGGSNPDGSPCSDD (residues 205–222). Residue 87 in the γ peptide region and residue 206 in the β subunit were identified as Ser. Residue 137 in the γ peptide region could not be identified owing to the low signal.

The native SMKT was also analyzed by ESI-MS. Under the conditions described in Materials and methods, SMKT was completely dissociated into two subunits and the mass values of the α and β subunits were estimated as 6341.9 and 7847.0 by ESI-MS. The calculated molecular masses of the α subunits, β subunit with Leu206 and β subunit with Ser206 were 6342.17, 7872.54 and 7846.46, respectively. The observed molecular mass of the β subunit indicates that the 206th residue should be Ser.

For further analysis of residue 206, the β subunit was reduced, carboxymethylated and digested with lysylendopeptidase. The resultant peptide fragments were analyzed by LC–ESI IT-MS–MS (Figure 1). Each peak was clearly identified as the expected peptide fragments. Peak 7 was identified as the C-terminal peptide including carboxymethylated Cys199 and Ser206. The product ion spectrum obtained by LC–ESI IT-MS–MS measurement of peak 7 is shown in Figure 1C. Fragmentation of the peptide residues 205–222 indicated that residue 206 is unmodified Ser. Based on these data, we concluded that the CUG codon is read as Ser in *P. farinosa*.

In a recent study, we found that the expression of SMK1 has a lethal effect in *S.cerevisiae* (Suzuki et al., 2000). It was shown that the side chain of Leu207 is required for stabilizing the hydrophobic environment that contributes to the overall structure of the SMK1 product and therefore the deletion or substitution of the residue resulted in the loss of toxicity in *S.cerevisiae*. The expression of SMK1 in *S.cerevisiae* should cause substitutions of Leu at three CUG codons.

Recharge of hydrophobic Leu residues by polar Ser residues could in some situations perturb the structure of the SMK1 product. In order to produce authentic SMK1 product, it is desirable to alter any CUG codons to a universal Ser codon. When a CUG codon corresponding to Ser206 was modified to TCG, the resulting gene, *SMK1L206S* TATATGATCAAGTTCTCGTT-, was synthesized by TATATGATCAAGTTCTCGTT-

as described previously (Suzuki, 1999). Protein A purified rabbit antiserum against the α subunit, that against the β subunit and horseradish peroxidase-linked anti-rabbit immunoglobulin (Bio-Rad Technologies, Richmond, CA) were used at dilutions of 1:5000, 1:1000 and 1:3000, respectively. For detection, the ECL Plus Western Blotting Detection Reagents (Amersham Pharmacia Biotech, Piscataway, NJ) were used.

To identify the amino acid specified by the CUG codon in *P. farinosa*, the secreted protoxin gp26 were digested by lysylendopeptidase and the purified fragments were examined by Edman sequence analysis. Two peptides that contained the amino acid residue specified by the CUG codon were sequenced as SSDSFSIVHNQGE (residues 83–96) and FSLAGGSNPDGSPCSDD (residues 205–222). Residue 87 in the γ peptide region and residue 206 in the β subunit were identified as Ser. Residue 137 in the γ peptide region could not be identified owing to the low signal.

The native SMKT was also analyzed by ESI-MS. Under the conditions described in Materials and methods, SMKT was completely dissociated into two subunits and the mass values of the α and β subunits were estimated as 6341.9 and 7847.0 by ESI-MS. The calculated molecular masses of the α subunits, β subunit with Leu206 and β subunit with Ser206 were 6342.17, 7872.54 and 7846.46, respectively. The observed molecular mass of the β subunit indicates that the 206th residue should be Ser.

For further analysis of residue 206, the β subunit was reduced, carboxymethylated and digested with lysylendopeptidase. The resultant peptide fragments were analyzed by LC–ESI IT-MS–MS (Figure 1). Each peak was clearly identified as the expected peptide fragments. Peak 7 was identified as the C-terminal peptide including carboxymethylated Cys199 and Ser206. The product ion spectrum obtained by LC–ESI IT-MS–MS measurement of peak 7 is shown in Figure 1C. Fragmentation of the peptide residues 205–222 indicated that residue 206 is unmodified Ser. Based on these data, we concluded that the CUG codon is read as Ser in *P. farinosa*.

In a recent study, we found that the expression of SMK1 has a lethal effect in *S.cerevisiae* (Suzuki et al., 2000). It was shown that the side chain of Leu207 is required for stabilizing the hydrophobic environment that contributes to the overall structure of the SMK1 product and therefore the deletion or substitution of the residue resulted in the loss of toxicity in *S.cerevisiae*. The expression of SMK1 in *S.cerevisiae* should cause substitutions of Leu at three CUG codons. Replacement of hydrophobic Leu residues by polar Ser residues could in some situations perturb the structure of the SMK1 product. In order to produce authentic SMK1 product, it is desirable to alter any CUG codons to a universal Ser codon. When a CUG codon corresponding to Ser206 was modified to TCG, the resulting gene, *SMK1L206S*, was lethal and secretion of protoxin was observed, similar to that in the expression of unmodified SMK1. The side chain of Ser206 faces solvent but is covered by the C-terminal loop of the β subunit. Therefore, the substitution of Leu for Ser206 may not affect the structure critical for the lethal effect of the SMK1 product.

To examine the effects of two other Ser residues, all three CUG codons in SMK1 corresponding to Ser87, Ser137 and Ser206 were replaced by universal Ser codons. The modified SMK1 gene, SMK1L/S, was inserted under the control of a galactose-inducible promoter, pGal1. The resultant plasmid,
Alternative CUG codon usage in Pichia farinosa

Fig. 1. (A) Base peak intensity trace of LC–ESI IT-MS–MS of the lysylendopeptidase digest of reduced and carboxymethylated (RCM) β subunit of SMKT. The base peak intensity is 3.3×10^6. (B) Calculated and observed masses of lysylendopeptidase fragments of RCM-β subunit. Residues represent amino acid numbers corresponding to the preprotoxin. * indicates the number of amino acid residues corresponding to the preprotoxin. The N-terminal amino acid of the β subunit is Gly146. ** indicates the mass containing carboxymethyl-Cys. (C) Product ion spectrum obtained by LC–ESI IT-MS–MS measurement of peak 7 in (A). The product ion spectrum was obtained from the doubly charged ion (m/z 871.2) of the peptide corresponding to peak 7. The base peak intensity is 6.7×10^5. C* in the amino acid sequence represents the carboxymethylated cysteine.

Fig. 2. Toxicity of endogenously expressed SMK1. The vector (pYES2), a plasmid carrying unmodified SMK1 (pCS224) or that carrying SMK1L/S (pCS262) was transformed into wild-type cells (W303–1A) and spf1 disruptants (CS701C). The transformants were streaked on a methylene blue plate containing 2% galactose. Dead cells stained blue.

pCS262, was transformed into the wild-type strain (W303–1A) of S. cerevisiae. When SMK1L/S was induced by galactose, the colonies of transformants showed a blue color in the presence of methylene blue employed to stain dead cells. As shown in Figure 2, the colonies of cells expressing SMK1L/S product turned blue, as was observed when unmodified SMK1 and SMK1L206S were expressed, suggesting that the side chains of residues Ser87, Ser137 and Ser206 do not affect the structure stability and/or the toxicity of the SMK1 product. When pCS224 or pCS262 was transformed into the spf1 disruptant cells (CS701C), which are highly resistant to exogenous SMKT, the transformant colonies did not show a blue color on YPGal/MB plates (Figure 2). We have recently demonstrated that the spf1 disruptants acquire resistance to SMKT by trapping the toxin at the modified cell surface (Suzuki and Shimma, 1999). Glycosylation is defective in spf1 disruptants, which may generate molecules with affinity to SMKT. However, the spf1 disruptants also acquired resistance to endogenous SMK1 products, suggesting that the intracellular target site of SMKT is also modified in the absence of Spf1p.

The expression of SMK1L/S products was examined by immunoblotting. Transformants were grown in YPGal overnight. Secreted proteins in the culture medium were precipitated with trichloroacetic acid. Collected cells were broken with glass beads and cell lysates were separated into precipitated fraction (P100) and supernatant fraction (S100) by ultracentrifugation. Immunoblotting was performed using antibodies against the α subunit and against the β subunit (Figure 3). Immunoblotting showed that the mature α subunit was secreted from pCS224 and pCS262 both in wild-type cells and in spf1 disruptant cells. No mature α subunit was detected in P100 and S100. The anti-α subunit antibody specifically recognizes the C-terminal region of the mature α subunit, because the mature α subunit, but not other SMK1 products, was detected.

Processing of precursors should result in the release of mature α subunit, β subunit and/or γ–β peptide. When anti-β subunit antibody was used, two SMK1L/S products were detected in the secreted fraction from pCS262. The larger SMK1L/S product was the same as the protoxin, gp26. The smaller product may be a glycosylated γ–β peptide. The β subunit was not detected either in secreted fractions or in P100.

The amount of the secreted SMK1L/S product from pCS262 was greater than that of the SMK1 product from pCS224.
Especially the secretion of protoxin with authentic amino acid sequence was significantly increased. To determine whether the loss of the rare CUG codon increases the expression of Ser residues in the γ region are involved in the secretion processes, the transcription level was examined by RT-PCR. As shown in Figure 3C, the presence or absence of CUG codon did not have much effect on the transcription level of $SMK1$ either in the wild-type or in $spf1$ disruptant cells. Therefore, the increased expression of protoxin with the residues of Ser87, Ser137 and Ser206 is not due to the elevated level of transcription but is probably due to the stable conformation by the authentic amino acid residues.

The $SMK1$ product accumulated in a significant amount in P100 of $spf1$ disruptants but was hardly detected in those of the wild-type cells. Recently, we have shown by subcellular fractionation experiments that the fractionation pattern of Spf1p was similar to that of an early Golgi protein, Och1p (Suzuki, 2001). Glycosylation is defective in $spf1$ disruptants. Loss of Spf1p may perturb the ion homeostasis in the secretory pathway, which may result in glycosylation defects and disturbance of the secretion process. Accumulation of the $SMK1$ product may be due to the partially perturbed secretory pathway.

Alternative codon usage in $P.sorbitophila$ and $C.cacaoi$ was predicted from the sequence information from the genome and from in vitro experiments, respectively. However, there was no discussion about each Ser residue encoded by CUG codon. So far, two genes in $P.sorbitophila$ have been sequenced. In the case of Pichia TIM10, Ser83 is encoded by CUG codon and is conserved in $S.cerevisiae$ and Candida albicans (Kayingo et al., 2000). In the case of $PsLYS2$, the gene encoding the α-aminoadipate reductase from $P.sorbitophila$, three CUG codons corresponding to Ser41, Ser1196 and Ser1322 were found in the ORF, although these residues were specified as Leu (Bleykasten-Grosshans et al., 2001). The corresponding residues to Ser1196 and Ser1322 in Calbicans are also Ser. Because $PsLYS2$ is functional in $S.cerevisiae$, these residues are not critical for its function. There are 112 Ser residues in the $PsLYS2$ product including those encoded by CUG codon, suggesting that the CUG codon is a rare codon in $P.sorbitophila$. On the other hand, CUG is relatively abundant in the $SMK1$ gene, in which three out of 20 Ser residues (15%) are encoded by CUG codons. Since no sequence similar to $SMK1$ was found, the origin of the chromosomal killer gene is unknown. The abundance of CUG codon might be peculiar to the killer gene, which is irrelevant to the evolution of genome.

$P.farinosas/P.sorbitophila$ is highly resistant to osmotic stress in general and particularly to salt stress up to 4 M NaCl. Previous studies showed that glyceral, the major solute in this yeast, is actively co-transported with protons to maintain osmotic balance (Lages and Lucas, 1995; Oliveira et al., 1996). In addition to the high halotolerance, $P.farinosas$ secretes proteins with relatively small sugar chains. gp26, the secreted protoxin of SMKT, has a mannos-containing sugar chain about 2 kDa in size (Suzuki, 1999). Highly glycosylated sugar chains observed in $S.cerevisiae$ were not added to the protoxin. These properties are potentially useful for heterologous expression of proteins, especially those derived from mammals. Because the CUG codon encodes Ser in most Candida species, the character of $P.farinosas$ is promising for recombinant protein production derived from such species. The expression system of $P.farinosas$ is currently being sought based on the need for further elucidation of the killer mechanism and

**Fig. 3.** Expression of $SMK1$. Immunoblotting of P100, S100 and TCA precipitates of culture filtrates of wild-type cells (A) and $spf1$ disruptants (B) and RT-PCR (C). After overnight culture in the selective medium, each transformant carrying the vector (pYES2), unmodified $SMK1$ (pCS224) or $SMK1L$ (pCS262), was cultured overnight in YPGal (pH 4.5). Cell lysates were subjected to ultracentrifugation to obtain P100 (P) and S100 (S). Secreted proteins in culture filtrates were precipitated with TCA (sec). Samples were analyzed by using anti-α subunit and anti-β subunit antibodies. The positions of bands of gp26, β subunit and α subunit of SMKT are indicated by arrows. Total RNA was also prepared from cells cultured in YPGal overnight. RT-PCR was performed to detect the $SMK1$ transcript. A 2 µg amount of total RNA (RNA), PCR product without reverse transcriptase (–RT), PCR products amplified for 18 cycles (18) and for 22 cycles (22) from each transformant are shown.
potential use of this yeast as a host in functional expression of heterologous genes.

Acknowledgements

We thank Dr J. Shima for the helpful discussions and critical reading of the manuscript. This work was supported in part by a grant-in-aid (Pioneer program) from the Ministry of Agriculture, Forestry and Fisheries, Japan.

References


Received 16 November 2001; revised 6 December 2001; accepted 7 December 2001