Development of efficient protein expression system in filamentous fungi, *Aspergillus* species

Akio Koda

2005

Contents

Introduction 1

Chapter I Codon optimization for high-level expression of heterologous proteins in *Aspergillus niger*

Introduction 6

Materials and Methods 7

Results 11

Discussion 19

Chapter II Improvement of heterologous protein expression by enhancing translation efficiency mediated by the 5′ untranslated region in *Aspergillus oryzae*

Introduction 23

Materials and Methods 25

Results 31

Discussion 36

Chapter III Effect of the 5′ untranslated region of the *Hsp12* gene on translation efficiency in *Aspergillus oryzae*

Introduction 41

Materials and Methods 41

Results 44

Discussion 48

Conclusion 52

References 55

Acknowledgments 65

Publications 67

Introduction

Recombinant protein expression system is useful for obtaining sufficient amount of proteins of interest. Many kinds of expression systems using bacteria, yeast, plants, mammalian cells and other organisms as hosts have been developed to date. In general, the expression of higher eukaryotic proteins using prokaryotes as a host sometimes results in an inactive product without correct folding or posttranslational modifications, although the manipulation is easy. Proteins produced in yeasts such as *Saccharomyces* species and *Pichia* species invariably show high mannose-type glycosylation, although their manipulation is also easy. In contrast, in the case of expression using mammalian cells as hosts, the problems mentioned above can be solved, but the manipulation is not easy, and the production cost is high. The advantage of utility of plants is lower production costs compared with mammalian cells. However, there are a number of technical limitations. For example, a technique for gene targeting has not yet been established in plant cells, therefore, the transgene can only be inserted randomly into the chromosome of transgenic plants. Thus, each expression system has its own characteristics, so it is necessary to make a choice based on the gene (protein) of interest, the amount to be produced, the purpose and final aim. The demand for superior expression systems is increasing rapidly in biotechnological industry.

Filamentous fungi, especially several *Aspergillus* species, have been used for the traditional fermented beverage and foods, such as *sake*, *shochu*, *shoyu*, and *miso*, in Japan for over a thousand years. For about two decades, molecular genetic tools have enabled us to use these organisms to express extra copies of both homologous and heterologous genes. Recently, *Aspergillus* species have received attention as the

attractive hosts, and are widely used for the expression of homologous and heterologous genes, because they provide advantages for protein expression, such as a large capacity for protein secretion, correct post-translational processing of eukaryotic proteins, and less over-glycosylation than in yeast. (Archer and Peberdy 1997; Verdoes et al. 1995). Previous efforts at enhancing heterologous gene expression in filamentous fungi have mostly been directed at boosting transcription. To date, several strong promoters have been isolated and used for the production of heterologous proteins in *Aspergillus* species (Christensen et al. 1988; Tada et al. 1991; Hata et al. 1992; Toda et al. 2001; Ishida et al. 2001).

In author's laboratory, improved promoters such as P*agdA*142 (Minetoki et al. 1998), P-No8142 (Ozeki et al. 1996; Minetoki 2000), and P-*enoA*142 (Tsuboi et al. 2005) have already been developed using a functional *cis*-element, Region III (Region IIIa and IIIb). Region IIIa was found as one of the highly conserved sequences in the promoter regions of the *A. oryzae* amylolytic genes, Taka-amylase A (*amyB*) (Tada et al. 1989), glucoamylase (*glaA*) (Hata et al. 1990, 1991), and α-glucosidase (*agdA*) (Minetoki et al. 1995) (Fig. 1). Deletion analyses of the promoter regions indicated that Region IIIa is a functional element essential for starch or malto-oligosaccharide induction (Minetoki et al. 1996, 1998). Region IIIa overlaps SRE (starch-responsive element) which binds to the transcriptional factor AmyRp (Gomi et al. 2000; Tsukagoshi et al. 2001) (Fig. 2). Region IIIb has a CCAAT box (Hap complex binding sequence), which is considered to enhance overall promoter activity and to increase expression in conjunction with other transcriptional factors (Tsukagoshi et al. 2001) (Fig. 2). Introduction of the multiple copies of the fragment comprising Region III (including both IIIa and IIIb) into various promoters resulted in a significant increase

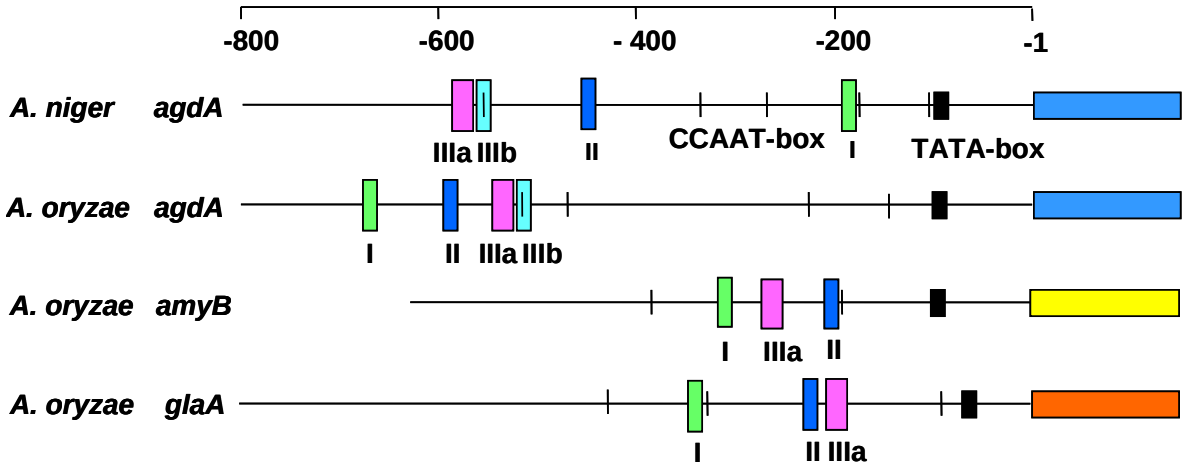

FIG. 1. Diagram of the homologous sequence elements identified in the promoter region of the *Aspergillus* amylase genes. The putative CCAAT sequence and the TATA-box are indicated by the *vertical line* and *solid box*, respectively. Each consensus sequence element is indicated by different box pattern. The numbers indicate the nucleotide distance to the translation start point.

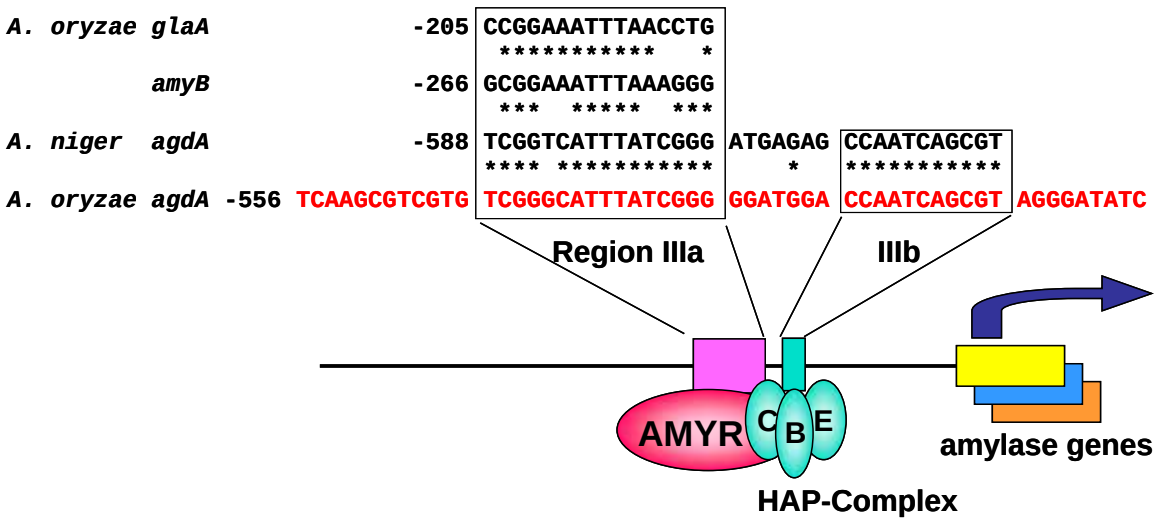

FIG. 2. The conserved nucleotide sequences, including Region III, in the upstream region of *Aspergillus* amylase genes. The numbers indicate the nucleotide distance to the translation start point.

in their promoter activities at the transcriptional level (Minetoki et al. 1998; Tsuboi et al. 2005) (Fig. 3). Indeed, using these improved promoters, a high yield of α-glucosidase, nuclease S1, and 1,2-α-mannosidase, the production levels of which are very low by nature because of their own gene expression level in wild-type strains, could be gained.

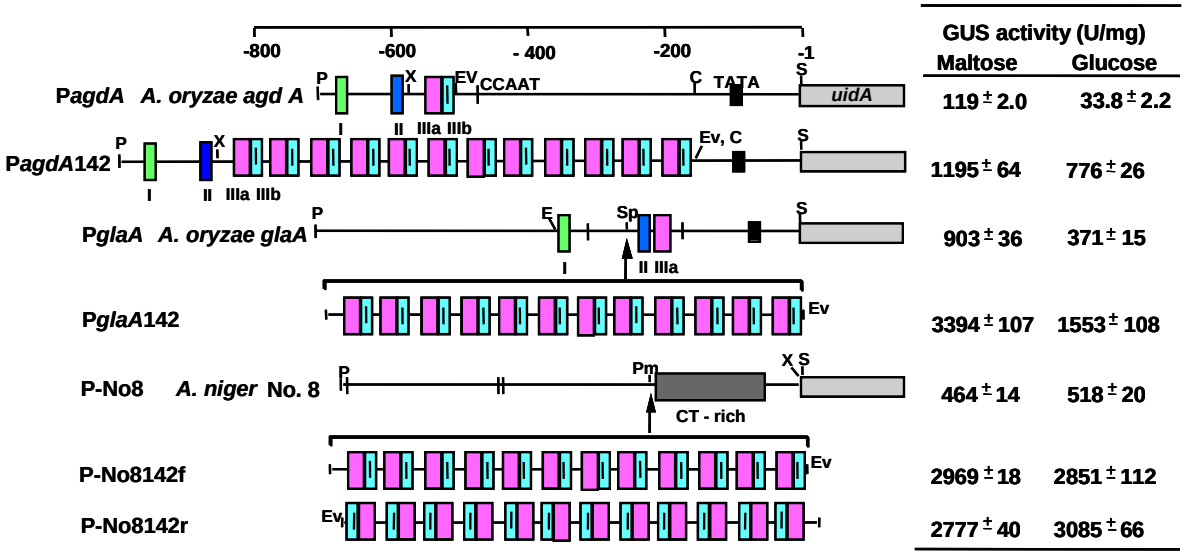

FIG. 3. Improvement of promoter activity by the introduction of the Region III. The location of each conserved element (Region I, II, IIIa and IIIb) within the promoter region is indicated by a different box pattern. The putative CCAAT-box and the TATA-box are indicated by the *vertical line* and *solid box*, respectively. The numbers marked at the top indicate the nucleotide distance to the translation start point. Two independently isolated single copy transformants with each plasmid integrated at the *niaD* locus were grown in 15 ml of CD-P medium containing 2% glucose or maltose at 30°C for 48 h. The GUS activities of cell extracts obtained from two independent experiments are presented as the average with standard errors. C, *Cla*I; E, *Eco*RI; Ev, *Eco*RV; P, *Pst*I; Pm, *Pma*CI; S, *Sal*I; Sp, *Spe*I; X, *Xho*I.

However, in addition to the transcriptional effect, the level of heterologous protein expression can also depend on other steps, including posttranscription initiation events such as transcription elongation, mRNA processing, mRNA stability and/or efficiency of translation. Little is known about the effects of these events on protein expression in filamentous fungi.

In the present study, I addressed posttranscriptional enhancement in heterologous gene expression in filamentous fungus *Aspergillus* species. In Chapter I, I focused on the coding region of a gene to be expressed, and investigated the effect of codon optimization on the expression of a heterologous protein in *Aspergillus niger*, using potato α-glucan phosphorylase (GP) as a model protein. In Chapter II, it was investigated whether alteration in the 5′ untranslated region (5′UTR) of an expression vector has a large influence on translation efficiency in *Aspergillus oryzae*, using a model β-glucuronidase (GUS) expression system. Moreover, this paper describes the 5′UTRs of the fungal heat-shock protein (Hsp) genes that dramatically increase the expression level of an exogenous gene in *A. oryzae* in Chapter III.

Chapter I

Codon optimization for high-level expression of heterologous proteins in *Aspergillus niger*

Introduction

Filamentous fungi, especially several *Aspergillus* species, have a high capacity for protein secretion and are therefore exploited for the industrial production of homologous and heterologous proteins (Archer and Peberdy 1997; Verdoes et al. 1995). However, for most nonfungal proteins the secreted yields are very low, often 2 to 3 orders of magnitude lower than homologous proteins (Gouka et al. 1997). Although significant production levels have been obtained for several nonfungal proteins by using strong promoters and the secretion carrier (gene fusion) approach (Ward et al. 1990; Jeenes et al. 1993; Contreras et al. 1991), the yields in most cases are still orders of magnitude lower than those obtained for fungal proteins. The bottlenecks that limit heterologous protein production in *Aspergillus* species are a subject of great interest (Conesa et al. 2001). In comparison with other foreign proteins, such as mammalian proteins, there are only a few reports of the production of plant proteins in *Aspergillus* species (Gouka et al. 1996; Juge et al. 1998; Moralejo et al. 2000). Moreover, it is not known whether plant cytosolic proteins can be produced in *Aspergillus* species. Despite this, heterologous expression of plant genes in *Aspergillus* species may serve as an important alternative for producing plant proteins.

In this study, potato (*Solanum tuberosum*) α-glucan phosphorylase (GP, EC

2.4.1.1), an enzyme that catalyzes the reversible phosphorolysis of α-1,4-glucan, was used as a model plant cytosolic protein. GP has been shown to be useful for the synthesis of glucose 1-phosphate, which has potential as an antibiotic or immunosuppressive drug (Weinhäusel et al. 1994, 1995), as well as for generating amylose molecules that have a narrow molecular weight distribution (Kitamura 1996; Niemann et al. 1991; Fujii et al. 2003). Thus, this enzyme is of interest for potential industrial application. The cDNA encoding potato GP has been cloned, and its nucleotide sequence was determined (Mori et al. 1991). Here, it was attempted to express the potato GP in *Aspergillus niger* using the strong promoter P-No8142 (Ozeki et al. 1996; Minetoki 2000). Special attention was paid to the influence of different codon usage and A+T content in the coding region on GP protein expression, and a significant increase in GP gene expression by using an optimized synthetic gene was observed.

Materials and Methods

Strains and media *A. niger* ND48prt42A122A, a derivative strain of *A. niger* IFO4343, was used as a recipient strain for transformation and GP expression experiments. The strain has deleted *niaD* and *pepA* loci, and it was obtained by mutagenesis and screening for reduced extracellular protease activity. *Escherichia coli* DH5α was used for construction and propagation of plasmids. Modified dextrin-peptone (DP) medium containing 2% (w/v) dextrin, 1% (w/v) polypeptone (Polypepton; Nihon Pharmaceutical), 0.5% (w/v) KH_2PO_4, 0.05% (w/v) $MgSO_4 \cdot 7H_2O$, disodium hydrogen citrate buffer (50 mM, pH 6.5), and 1 mM pyridoxine hydrochloride

was used as the medium for production of GP.

Construction of a potato GP expression vector using the wild-type GP gene

The wild-type potato GP cDNA (GP-wt, Fig. I.3) inserted into a unique *Eco*RI site in pBluescript II SK was kindly provided by Prof. K. Tanizawa (Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka) (Mori et al. 1991). This was digested with *Xho*I and *Xba*I to release the GP-wt coding region, which was then cloned into *Xho*I/*Xba*I-digested pNAN8142 (Minetoki 2000), an *Aspergillus* expression vector controlled by the P-No8142 promoter (Fig. I.1), to produce the GP-wt expression vector, pNAN-GP-wt.

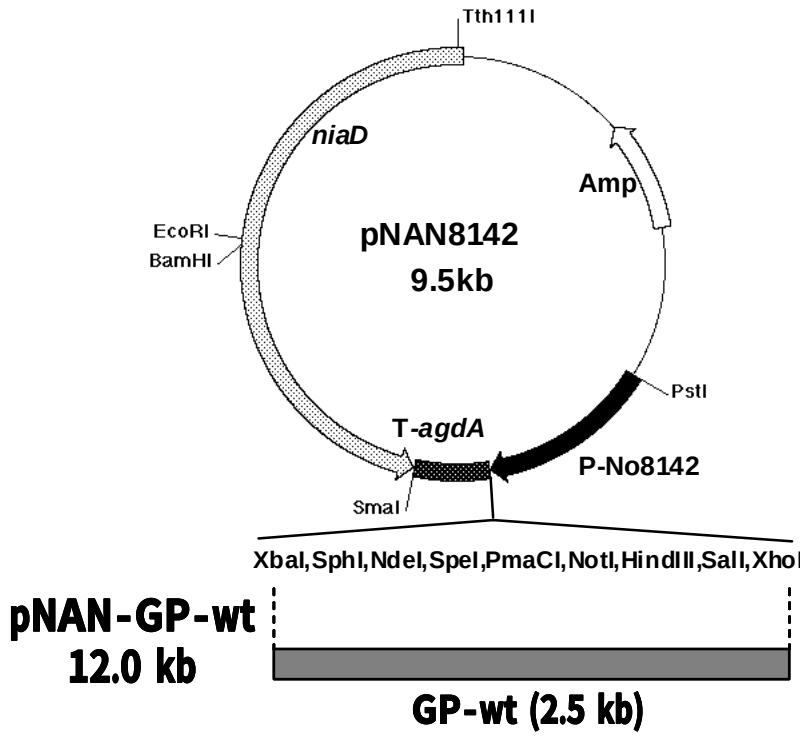

FIG. I.1. Structure of the expression vector pNAN8142 and pNAN-GP-wt using the improved promoter. P-No8142 is the improved promoter with Region III. T-agdA is the transcriptional terminator from the *A. oryzae agdA* gene. The *niaD* gene is used as a selectable marker gene.

Design and synthesis of a GP gene for expression in *A. niger* For optimal expression in *A. niger*, the nucleotide sequence of the GP gene was modified to have *A. niger* codon usage preference. To facilitate cloning into the *Aspergillus* expression vector pNAN8142, unique restriction endonuclease sites for *Hin*dIII and *Xba*I were added to the 5′- and 3′-ends of the coding region, respectively. The actual codons used for each amino acid are summarized in Table I.1.

The synthetic GP cDNA (GP-syn) thus designed was constructed by assembly PCR (Fig. I.2) using a set of overlapping 40-mer oligonucleotides with Pyrobest DNA polymerase (Takara Bio) essentially according to the procedure of Stemmer et al. (Stemmer et al. 1995). The amplified GP-syn was digested with *Hin*dIII and *Xba*I and then cloned into the multiple cloning site in pNAN8142, yielding pNAN-GP-syn, which was verified by sequencing and used for protein expression.

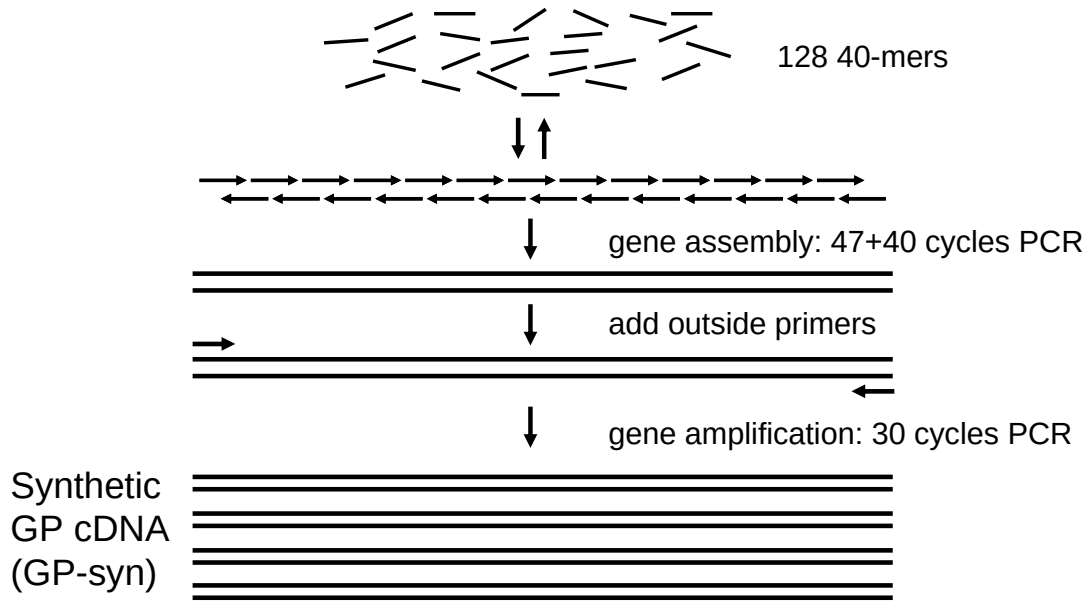

FIG. I.2. Schematic illustration of the assembly PCR-based gene synthesis. The synthetic gene is assembled by DNA polymerization from a set of overlapping complementary oligonucleotides in a first PCR (gene assembly). The assembled product is then amplified using the two outermost primers (gene amplification).

Fungal transformation Transformation of *A. niger* was carried out according to Gomi et al. (Gomi et al. 1987). Protoplasts for this procedure were prepared using 20 mg/ml of Lysing Enzymes (Sigma).

GP activity assay *A. niger* transformants were cultured in 15 ml of modified DP medium at 30°C with shaking. After 60 h of cultivation, the mycelia were harvested and washed with distilled water. Preparation of cell-free extracts of the transformants was carried out as described by Tada et al. (1991) using an extraction buffer of 10 mM Tris-HCl (pH 7.0) containing 0.1% (w/v) Triton X-100.

GP activity was assayed by determining the amount of inorganic phosphate released from glucan and glucose 1-phosphate according to Saheki et al. (1985) with minor modifications. The reaction mixture (200 μl) containing 5 mM Tris-HCl (pH 7.0), 0.025% (w/v) Triton X-100, 1% (w/v) soluble starch, and 12.5 mM glucose 1-phosphate was incubated at 37°C for 15 min. The reaction was stopped by the addition of 800 μl of molybdate reagent (15 mM ammonium molybdate and 100 mM zinc acetate, pH 5.0). Two hundred μl of 10% (w/v) ascorbic acid (pH 5.0) was then added to the mixture. This mixture was incubated at 30°C for 15 min, and the absorbance was measured at 850 nm. A reaction mixture without soluble starch was used as a blank for each sample. One unit was defined as the amount of enzyme that produced 1 μmol of phosphate in 1 min. Total protein was determined with a protein assay kit (Bio-Rad) using bovine serum albumin as a standard.

Northern blot analysis *A. niger* transformants were cultured in 15 ml of the modified DP medium for 40 h at 30°C with shaking, after which the mycelia were transferred into fresh modified DP medium and cultured for 8 h. Next, total RNA was isolated with an ISOGEN RNA isolation kit (Nippon Gene). Total RNA (10 μg) was

separated on a formaldehyde gel and transferred onto a Hybond-N^+ membrane (Amersham Biosciences) using the capillary transfer method. The entire GP-wt or GP-syn cDNA fragment was used as a probe for the transformants containing pNAN-GP-wt or pNAN-GP-syn, respectively. Probes were labeled and detected using the AlkPhos direct labeling and detection system (Amersham Biosciences).

SDS-PAGE analysis Electrophoresis was performed according to the method of Laemmli (1970) using 7.5% polyacrylamide gels. Proteins were stained with Coomassie brilliant blue R-250.

Immunoblot analysis The samples were loaded on a 7.5% SDS-polyacrylamide gel. Separated proteins were then electrophoretically transferred onto an Immobilon-P membrane (Millipore). Immunoblotting was performed using an Immun-Blot assay kit (goat anti-rabbit IgG [H+L] AP, Bio-Rad) with a rabbit polyclonal anti-GP antibody (kindly provided by Ezaki Glico Co. Ltd., Osaka). Prestained broad-range protein markers (Bio-Rad) were utilized for estimation of protein molecular weights.

Results

Expression of a wild-type GP gene in *A. niger* In this chapter, I used GP, a cytosolic protein derived from potato, which I expected to be expressed in the cytosol of *A. niger* as a properly folded and active protein. At first, a wild-type GP gene was tried to be expressed in *A. niger*. The wild-type GP cDNA (GP-wt, Fig. I.3) was cloned in the *Aspergillus* expression vector pNAN8142 under the control of a strong promoter, P-No8142 (Minetoki 2000), which is derived from the No. 8AN promoter

isolated from *A. niger* (Ozeki et al. 1996) (Fig. I.1). This construct was introduced into the *A. niger* ND48prt42A122A strain using *niaD* as the selectable marker, and several independent transformants were obtained. PCR analysis confirmed that the

TABLE I.1. Comparison of the codon usage for wild-type (wt) and synthetic (syn) GP genes

Amino acid	Codon	*A. niger* fraction [a]	GP		Amino acid	Codon	*A. niger* fraction	GP		Amino acid	Codon	*A. niger* fraction	GP	
			wt	syn				wt	syn				wt	syn
Tyr	TAT	0.27	21	9	Ser	TCA	0.06	8	0	Cys	TGT	0.34	4	4
	TAC	0.73	12	24		AGT	0.10	15	5		TGC	0.66	3	3
						TCG	0.16	5	5					
Phe	TTT	0.25	28	7		AGC	0.21	7	17	Arg	AGG	0.07	13	0
	TTC	0.75	7	28		TCC	0.30	2	10		AGA	0.08	8	0
						TCT	0.17	16	16		CGG	0.15	5	5
Gly	GGA	0.18	21	9							CGA	0.08	5	0
	GGG	0.09	9	0	Ile	ATA	0.04	8	0		CGC	0.36	5	18
	GGC	0.39	6	27		ATT	0.30	23	12		CGT	0.26	7	20
	GGT	0.34	15	15		ATC	0.66	15	34					
										Asn	AAT	0.24	30	12
Lys	AAG	0.84	24	46	His	CAT	0.34	17	7		AAC	0.76	14	32
	AAA	0.16	27	5		CAC	0.66	4	14					
										Gln	CAA	0.27	15	7
Leu	CTA	0.05	9	0	Val	GTC	0.45	14	18		CAG	0.73	19	27
	TTA	0.02	8	0		GTG	0.28	10	10					
	TTG	0.16	11	11		GTA	0.05	5	1	Thr	ACA	0.11	20	0
	CTT	0.15	27	6		GTT	0.23	21	21		ACG	0.17	5	5
	CTC	0.29	11	31							ACT	0.22	13	13
	CTG	0.33	14	32	Asp	GAT	0.42	33	13		ACC	0.50	4	24
						GAC	0.58	17	37					
Glu	GAG	0.70	23	44						Pro	CCC	0.41	5	15
	GAA	0.30	34	13	Ala	GCC	0.42	11	26		CCT	0.26	22	12
						GCA	0.12	28	4		CCA	0.12	9	0
Met	ATG	1.00	19	19		GCT	0.29	26	25		CCG	0.21	2	11
						GCG	0.16	1	11					
Stop	TAA	0.39	0	0						Total			839	839
	TAG	0.32	0	0	Trp	TGG	1.00	18	18	%A+T			57	44
	TGA	0.29	1	1						%C+G			43	56

[a] The codon usage for *A. niger* was determined from the codon usage database (http://www.kazusa.or.jp/codon/) (Nakamura et al. 2000).

TABLE I.2. GP production in different transformants of *A. niger*

Strain	Plasmid	GP activity (U/mg protein)	μg GP/mg protein[a]
Recipient	−	0.00489	<0.1
TFwt#1	pNAN-GP-wt	0.00883	<0.1
TFwt#2		0.00532	<0.1
TFwt#3		0.00511	<0.1
TFsyn#1	pNAN-GP-syn	3.79	43.6
TFsyn#2		3.45	39.6
TFsyn#3		3.51	40.3
TFsyn#5		8.23	94.6

[a] The amount of GP protein was calculated on the basis of the specific activity of purified recombinant GP (87 U/mg protein).

construct was integrated in the genome of these transformants. A GP assay (Table I.2) revealed that the transformants containing GP-wt (pNAN-GP-wt) failed to produce GP activity. In addition, Northern blot analysis showed that GP transcripts were hardly detectable despite the strong promoter (Fig. I.4C, lane 2), indicating that the lack of successful expression was apparent even at the mRNA level.

Design and synthesis of a GP gene for expression in *A. niger* Heterologous gene expression can be inhibited by unsuitable codons and/or differences in base contents between the expression host and the DNA donor. In my case, expression of GP-wt was inhibited at the mRNA level despite using the strong promoter, P-No8142, suggesting that characteristic features of the DNA sequence of GP-wt contributed to the lack of successful expression in *A. niger*. Therefore, the sequence of GP-wt (2517 bases) was analyzed. It was found that there are large differences between the most frequently used codons in GP-wt and *A. niger* genes and that there are 73 codons in the entire sequence that are extremely rare (fraction < 10%) in *A. niger* (Table I.1). Previous research in both prokaryotes and eukaryotes demonstrated that rare codons can

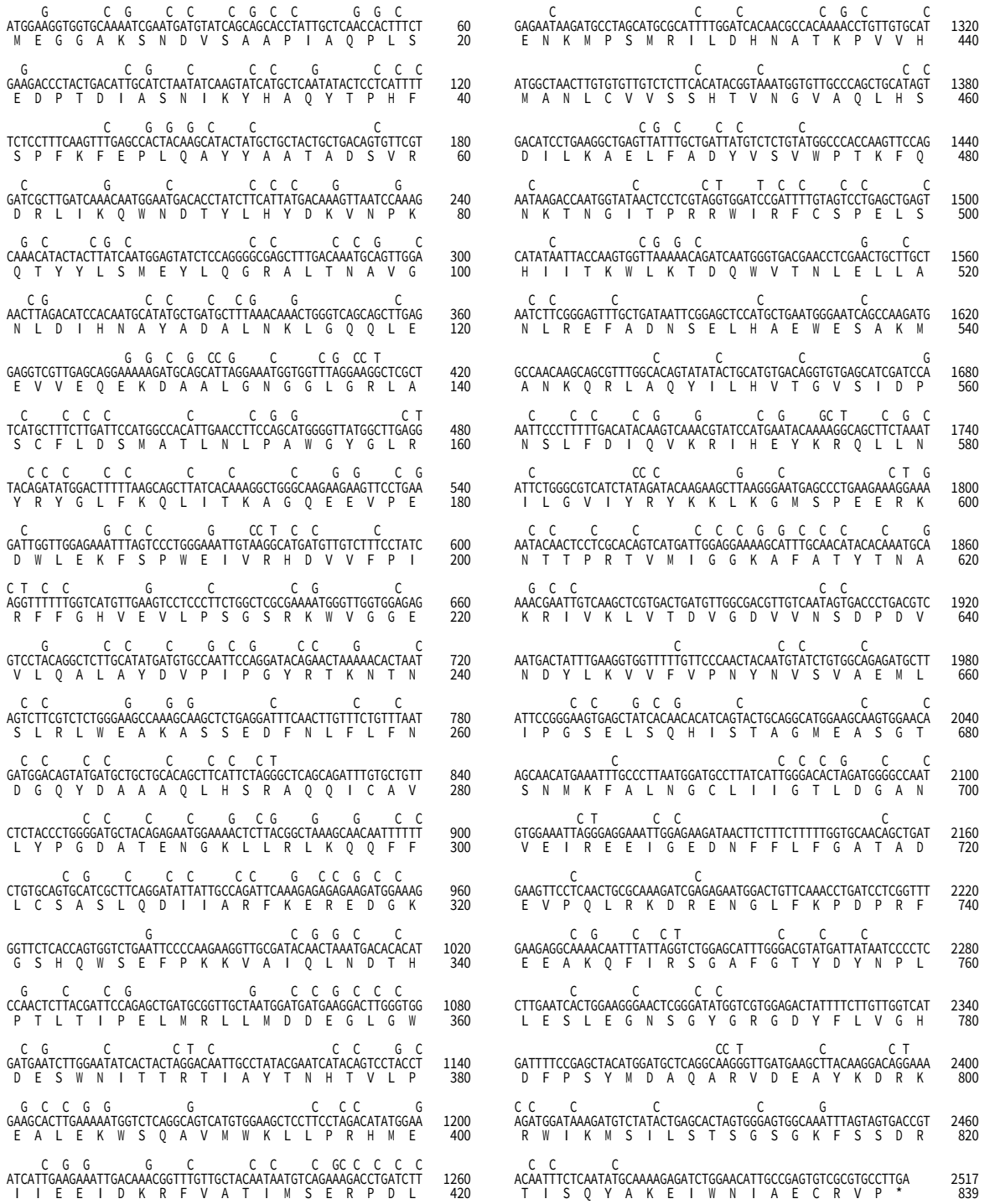

FIG. I.3. Nucleotide and deduced amino acid sequence of the wild-type GP gene (GP-wt). Differences between the GP-syn and GP-wt sequences are indicated above the nucleotide sequence of the latter.

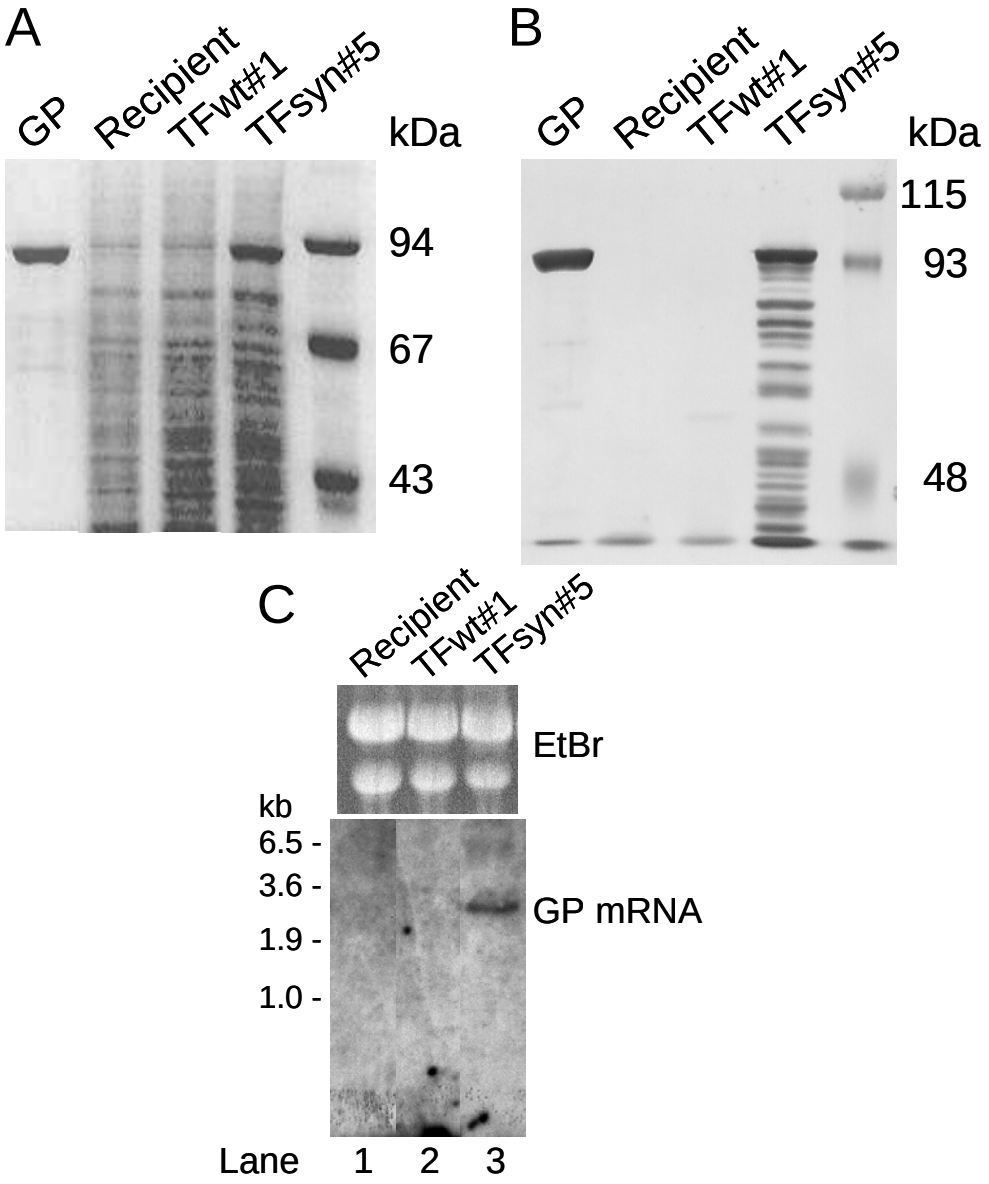

FIG. I.4. Expression analysis of GP-wt and GP-syn in *A. niger*. (A) SDS-PAGE analysis of cell-free extracts of transformants containing different constructs. The cell-free extracts were the same as those used for the GP assay in Table I.2. Each lane was loaded with 10 μg of protein. The first lane contains a control of purified *E. coli* cell-derived recombinant GP (40 mU; kindly provided by Ezaki Glico Co. Ltd.). The positions of the molecular mass markers are shown on the right. (B) Immunoblot analysis of intracellular GP protein in different transformants. The cell-free extracts (10 μg of protein) were resolved by SDS-PAGE and subjected to immunoblot analysis. The first lane contains a control of purified *E. coli* cell-derived recombinant GP (40 mU). The positions of prestained broad-range protein markers (Bio-Rad) and their sizes are indicated. (C) Northern blot analysis of GP mRNA in different transformants. The entire GP-wt fragment (for transformant TFwt#1) or the entire GP-syn fragment (for transformant TFsyn#5 and the recipient strain) was used as a probe. Ethidium bromide-stained ribosomal RNA is shown as a loading control.

affect mRNA stability (abundance) as well as the rate of translation. Therefore, it was suspected that the lack of detectable mRNA is due, in part, to the high content of rare codons. Additionally, based on Kazusa's Codon Usage Database (http://www.kazusa.or.jp/codon/), the A+T content is significantly higher in GP-wt (57%) than in *A. niger* genes (44%). Consequently, GP-wt contains many A+T-rich stretches that resemble a 'terminator' consensus (AATAAA) and can cause fortuitous polyadenylation (Fig. I.3). Thus, premature termination of the transcripts may also be responsible for the lack of detectable GP mRNA.

To overcome these possible problems, an optimized synthetic gene (GP-syn) was designed for expression in *A. niger* wherein the codon usage and A+T content was modified without altering the amino acid sequences (Fig. I.3). First, the A+T content was decreased from 57% to 44%, and special attention was paid to eliminating A+T-rich stretches. Second, the entire coding region was designed to have a suitable codon bias for *A. niger* (Table I.1). Most of the rare codons were converted to ones that are more frequently used in *A. niger*. As a result, 39% of the codons in the coding sequence were replaced by more suitable codons. These modifications resulted in an overall change of approximately 15% in the nucleotide sequence. GP-syn was then synthesized by assembly PCR and was cloned into the multiple cloning site in the expression vector pNAN8142, yielding pNAN-GP-syn.

Enhanced expression with the optimized synthetic gene The construct pNAN-GP-syn containing the optimized synthetic gene, GP-syn, was introduced into the *A. niger* recipient strain, and several independent transformants were obtained. By using GP-syn (pNAN-GP-syn), the expression level was dramatically improved (Table I.2). The highest GP activity was observed in the transformant TFsyn#5. The

estimated copy number of the expression cassettes in the transformant TFsyn#5 by Southern blot analysis is 2 or 3, and the other transformants, TFsyn#1, TFsyn#2, and TFsyn#3, were estimated to be a single copy transformants (data not shown).

Intracellular GP production was also examined by SDS-PAGE (Fig. I.4A) and immunoblot analysis (Fig. I.4B). TFsyn#5 showed a strong band of approximately 94 kDa, the expected size for GP protein, whereas this band was hardly detectable in the extract of transformant TFwt#1, which contained GP-wt (pNAN-GP-wt). In addition, several smaller immunoreactive proteins were detected in transformant TFsyn#5 (Fig. I.4B). An assay of GP activity and quantitative scanning of the gel indicated that the GP protein accounts for approximately 10% of the total soluble protein in strain TFsyn#5. This expression level is much higher than that previously reported using an *E. coli* expression system (Mori et al. 1991).

To investigate whether the differences in GP production among the constructs could be explained by differences in mRNA levels, Northern blot analysis was performed (Fig. I.4C). A GP mRNA band of approximately 2.5 kb, corresponding to the size of the full-length GP gene, was detected only in the transformant containing GP-syn (TFsyn#5). Thus, a drastic increase in both the mRNA level and GP protein expression was obtained using the synthetic GP gene.

Product stabilization Proteolytic degradation can often be an important problem in protein production, especially for heterologous expression. Immunoblot analysis of the expression of intracellular GP protein in transformant TFsyn#5 revealed many immunoreactive bands smaller than the intact protein (Fig. I.4B), suggesting that the protein is degraded intracellularly or in the cell extracts. I therefore investigated whether the addition of protease inhibitors to the extraction buffer can prevent the

degradation of GP protein (Fig. I.5). It was found that a serine protease inhibitor, phenylmethylsulfonyl fluoride (PMSF), effectively prevented the degradation of GP protein (Fig. I.5, lanes 3, 6, and 7), whereas the other protease inhibitors had no effect (Fig. I.5, lanes 4, 5). These results indicate that most of the proteolytic degradation occurs not intracellularly but in the cell extracts and that it can be prevented by the addition of PMSF. Indeed, addition of PMSF to the extract of TFsyn#5 increased the amount of GP protein by approximately 30% to approximately 100 mg GP protein/liter.

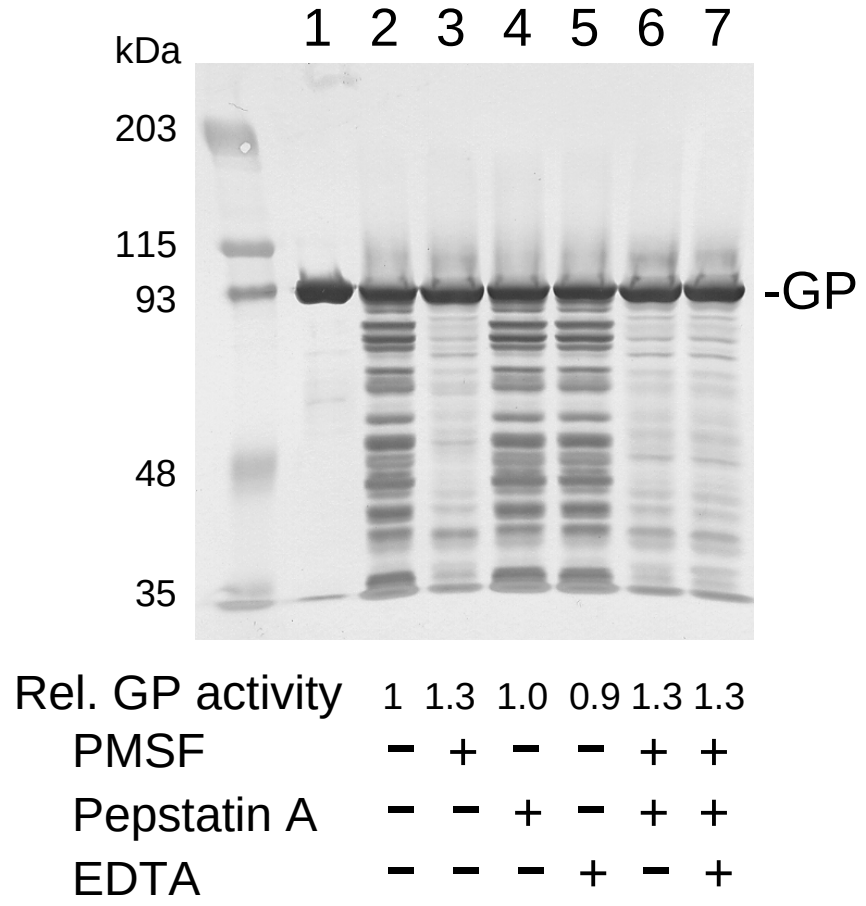

FIG. I.5. Immunoblot analysis of intracellular GP protein. Equal amounts of cell-free extract (10 μg of protein) of transformant TFsyn#5 with or without protease inhibitors were resolved by SDS-PAGE and subjected to immunoblot analysis. Lane 1, a control of purified *E. coli* cell-derived recombinant GP (40 mU); lane 2, no inhibitor; lane 3, phenylmethylsulfonyl fluoride (PMSF; 1 mM); lane 4, pepstatin A (1 μg/ml); lane 5, EDTA (5 mM); lane 6, PMSF (1 mM) and pepstatin A (1 μg/ml); lane 7, PMSF (1 mM), pepstatin A (1 μg/ml), and EDTA (5 mM). The positions of prestained broad-range protein markers (Bio-Rad) and their sizes are indicated. The relative GP activities (U/mg protein) are shown at the bottom of the blot.

Discussion

Matching the codon usage of heterologous genes to that of the expression host is a common strategy for increasing the expression of heterologous proteins in many species, including bacteria, plants, yeasts, mammalian cells, and transgenic animals. On the other hand, this strategy has not yet been investigated extensively in filamentous fungi, and there have been only a few studies that investigated the effect of codon optimization in *Aspergillus* species (Gouka et al. 1996; Faus et al. 1996, 1998; Krasevec et al. 2000; Cardoza et al. 2003; Nelson et al. 2004; and Tokuoka *et al*., Abstr. Annu. Meet. Jpn. Soc. Biosci. Biotechnol. Agrochem., p. 24, 2004). In this chapter, the successful expression of potato GP protein in *A*. *niger* was demonstrated. The breakthrough that allowed development of a system for high expression of the GP protein in *A*. *niger* was the redesigning of the primary DNA sequence of the gene.

In my early attempts, I failed to express the GP protein in *A*. *niger* transformants containing the wild-type GP gene (GP-wt). Northern blot analysis showed that the reason was very low abundance of the mRNA. Given the fact that the GP-wt was under control of the strong promoter in pNAN8142, the undetectable mRNA level observed was almost certainly not due to effects on transcription initiation but instead probably the result of post-transcription initiation effects, such as incorrect termination of transcription and/or mRNA stability, which are related to the characteristic features of the coding region. Previously, Miura et al. reported that the A+T-rich native α-amylase gene from *Sulfolobus solfataricus* was expressed at very low levels in *Candida utilis*, and they suggested that premature termination of the transcripts was responsible for the low production (Miura et al. 1999). Similar results have been reported for the

expression of A+T-rich genes in *Saccharomyces cerevisiae* and *Pichia pastoris*, wherein protein expression could not be observed and Northern blotting revealed truncated mRNAs (Romanos et al. 1991; Scorer et al. 1993). In my case, the A+T content of GP-wt is significantly higher (57%) than of *A. niger* genes (44%), and there are many A+T-rich stretches in GP-wt. A+T-rich sequences resembling a 'terminator' consensus (AATAAA) might have functioned as fortuitous, aberrant polyadenylation sites in *A. niger*, although truncated mRNAs were hardly detectable by Northern blot analysis. Such truncated mRNAs with wild-type gene sequences may be unstable in *A. niger* cells, preventing their accumulation.

One other possible reason for the lack of detectable mRNA for GP-wt is a high content of rare codons. It is generally accepted that rare codons can cause translational problems, that is, ribosomal pausing created by rare codons may cause inhibition of proper translation. The negative effects of rare codons on protein expression in *E. coli*, such as a reduced amount of the synthesized protein, frameshifting, and in-frame deletion of amino acids, have been reviewed by Kane (1995). Also, in filamentous fungi, introduction of three rare codons into the glutamate dehydrogenase gene of *Neurospora crassa* causes a 70% loss in protein production without affecting the mRNA level (Kinnaird et al. 1991), suggesting the importance of proper codon usage for effective translation. In addition, it is now well-established that the degradation of certain mRNAs is tightly coupled to their translation (Hentze 1991). Translation elongation can be interrupted when ribosomes reach a 'translation pause site' (Wolin and Walter 1988). In some cases, it has been shown that rare codons can slow the elongation rates along an mRNA, causing such a ribosomal pause, which can stimulate mRNA degradation (Caponigro et al. 1993; Hennigan and Jacobson 1996; Lemm and

Ross 2002). In these cases, it appears that a translational pause generates a ribosome-deficient region downstream of the pause site, and this unprotected region is exposed to endonuclease attack and/or other factors, such as mRNA-binding proteins, that can influence mRNA stability. Little is known about the influence of rare codons on mRNA stability in filamentous fungi. In the present case, the GP-wt contains many rare codons (Table I.1), so its mRNA may be rapidly degraded through similar mechanisms. Degradation of unusual mRNAs derived from heterologous species can prevent the synthesis of unnecessary, sometimes harmful proteins, resulting in a significant benefit to cell survival. Therefore, the regulation in mRNA level may be a general phenomenon among eukaryotes.

It was suspected that redesigning the primary DNA sequence of GP gene would solve these potential problems. The GP gene was reconstructed to reduce the AT content and to change the codon preference to that favored in *A. niger*. The synthetic GP gene (GP-syn) resulted in a dramatic increase in the level of both mRNA and protein for GP, although it is possible that there still are limitations in both. The expression of GP in *A. niger* using GP-syn accounted for as much as approximately 10% of the total soluble protein. Future research should focus on the extracellular production of GP protein, which should improve the efficiency of production in *A. niger* and facilitate purification. Furthermore, in such cases, the protease-deficient strain, also used in this study as a host, should alleviate the problem of extracellular degradation of the secreted GP protein.

An interesting question raised by the present data and some previous reports (Romanos et al. 1991; Scorer et al. 1993) is what the differences are in mRNA 3′-end formation among different eukaryotic species. In potato, *S. tuberosum*, GP-wt mRNA

should be processed properly. If inappropriate, premature termination of the GP-wt transcripts occurs in *A*. *niger* cells, it could be due to different mechanisms of mRNA 3′-end formation. Further research is required to ascertain whether and how the mechanisms of mRNA 3′-end formation can affect mRNA abundance.

In this chapter, it was demonstrated that redesigning of the potato GP gene can lead to substantial improvement in its expression in *A*. *niger*, although further research is required to understand the precise mechanisms involved. It is expected that redesigning the primary sequence of heterologous genes can be an effective strategy for the expression of heterologous proteins in filamentous fungus *Aspergillus* species.

Chapter II

Improvement of heterologous protein expression by enhancing translation efficiency mediated by the 5′ untranslated region in *Aspergillus oryzae*

Introduction

In Chapter I, the coding region of a gene to be expressed was focused, and the effect of codon optimization on the protein expression was investigated. In this chapter, the efficiency of translation initiation mediated by the 5′ untranslated region (5′UTR) was addressed (Fig. II.1).

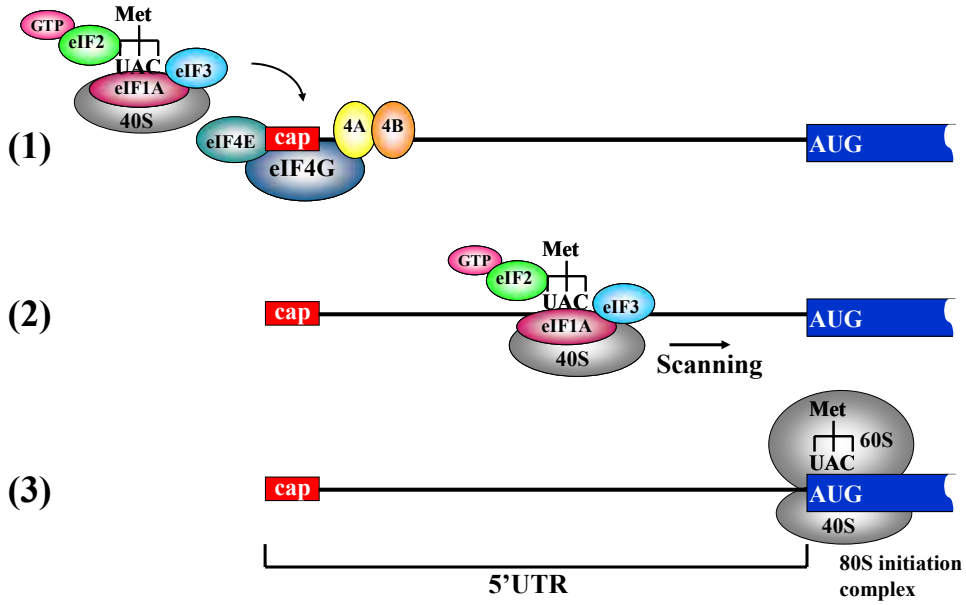

FIG. II.1. The scanning model for initiation of translation in eukaryotes. (1) The 40S ribosomal subunit is primed for initiating translation by binding of the ternary complex comprising eIF2, Met-tRNAi, and GTP. The resulting 43S preinitiation complex is recruited to the mRNA via interactions with the eIF4 factors (eIF4G/eIF4A/eIF4E) bound at or near the cap structure of the mRNA. (2) The 43S complex then scans the 5′UTR to locate the initiator codon. (3) Finally, initiation factors are released and the large (60S) ribosomal subunit joins to form an 80S ribosome.

The main strategy for increasing the production of heterologous proteins has been introducing the heterologous gene under the control of a strong promoter and/or *cis* elements in filamentous fungi. In addition to increasing transcription, it is important to maximize the amount of protein translated per unit of mRNA. In filamentous fungi, the exact mechanism of translation initiation remains unknown, so the translation process has often been either neglected or its importance for heterologous protein expression has been underestimated. Therefore, little attention has been paid to the 5′UTR in expression vectors and there are few data on the role of the 5′UTR in translation efficiency. However, as a result of recent studies of higher eukaryotes and yeast, it is generally accepted that the context of the translation initiation codon (AUG), length of the 5′UTR, secondary structure, GC content, and upstream AUGs (uAUGs) or upstream open reading frames (uORFs) play important roles in determining the translation rate (Kozak 1991a, b; Morris and Geballe 2000) (Fig. II.2). Therefore, in filamentous fungi, it is greatly desirable not only to elucidate the effect of the 5′UTR on translation efficiency, but also to develop a novel strategy to improve recombinant protein yields with a 5′UTR that contributes to efficient translation.

In this chapter, the translation (initiation) component of gene expression was focused in *A. oryzae*. It was investigated whether alteration of the 5′UTR of an expression vector has a large influence on translation efficiency in *A. oryzae*, using a model β-glucuronidase (GUS) expression system. In such alterations, we used the 5′UTR of the enolase gene (*enoA*) derived from *A. oryzae*, expecting positive effects on translation efficiency, because this gene is known to be one of the most strongly expressed genes in *A. oryzae* (Machida et al. 1996).

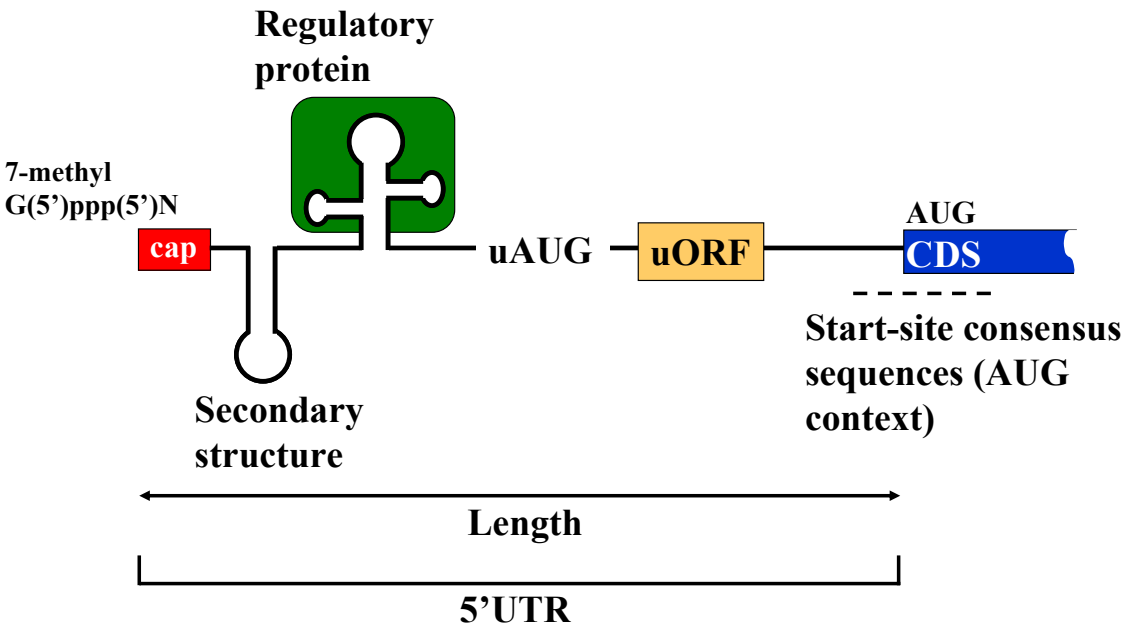

FIG. II.2. General elements within 5′UTR that regulate translation in eukaryotes.

Materials and Methods

Strains and media *A. oryzae* niaD300, a *niaD* mutant derived from the wild-type strain RIB40, was used as a recipient strain for transformation and was kindly provided by Dr. J. R. Kinghorn (University of St Andrews, UK). *E. coli* DH5α was used for construction and propagation of plasmids. The medium for fungal cultivation in flasks was dextrin-peptone (DP), consisting of 2% (w/v) dextrin, 1% (w/v) polypeptone (Polypepton; Nihon Pharmaceutical), 0.5% (w/v) KH_2PO_4, and 0.05% (w/v) $MgSO_4 \cdot 7H_2O$.

Plasmid construction Starting vectors for construction of the expression

vectors containing different 5′UTRs were pNAN8142 (Ozeki et al. 1996; Minetoki 2000) and pBI221 (Jefferson et al. 1987). pNAN8142 is a fungal high expression vector containing the P-No8142 promoter for protein expression, the terminator region of the *A. oryzae agdA* gene, and the *A. oryzae niaD* gene for selection in *A. oryzae*. The multiple cloning site is between the P-No8142 and the *agdA* terminator region (see Fig. I.1 of Chapter I). pBI221 was used as a donor for the reporter gene, *E. coli uidA*, encoding β-glucuronidase (GUS).

pNANG8142, used as a control expression vector, was constructed as follows: the *uidA* gene was isolated as a 1.9-kb *Sal*I-*Xba*I fragment from a derivative of pBI221 in which the *Sma*I site present upstream of the *uidA* initiation codon and the *Sac*I site present downstream of the *uidA* stop codon were converted to a *Sal*I site and a *Xba*I site, respectively, with the corresponding linkers. This fragment was cloned into *Sal*I/*Xba*I-digested pNAN8142 (*Sal*I and *Xba*I sites are located in the multiple cloning site of the vector), resulting in pNANG8142 (Fig. II.3). The chimeric construct and the 5′UTR sequence of pNANG8142 are shown in Fig. II.5.

pNANG-8142/enoUTR, in which the distal 14 bp of the 5′UTR derived from the P-No8142 promoter was replaced by the *enoA* 5′UTR, was constructed as follows: for modification of the 5′UTR, a *Stu*I site was created 8 bp downstream of the *uidA* translation initiation site with a two-nucleotide substitution by PCR. PCR was performed with pNANG8142 as a template using the primers 5′-AGA*AGGCCT*GTAGAAACCCCAACCCGTGA-3′ (bases mutated are the first and third bases (A, G) in the *Stu*I site, in italics) and 5′-TG*TCTAGA*CACACAAACGGTGATACGTACAC-3′ (the synthesized *Xba*I site is in italics). The amplified 0.4-kb fragment, a portion of the *uidA* gene, was cloned into

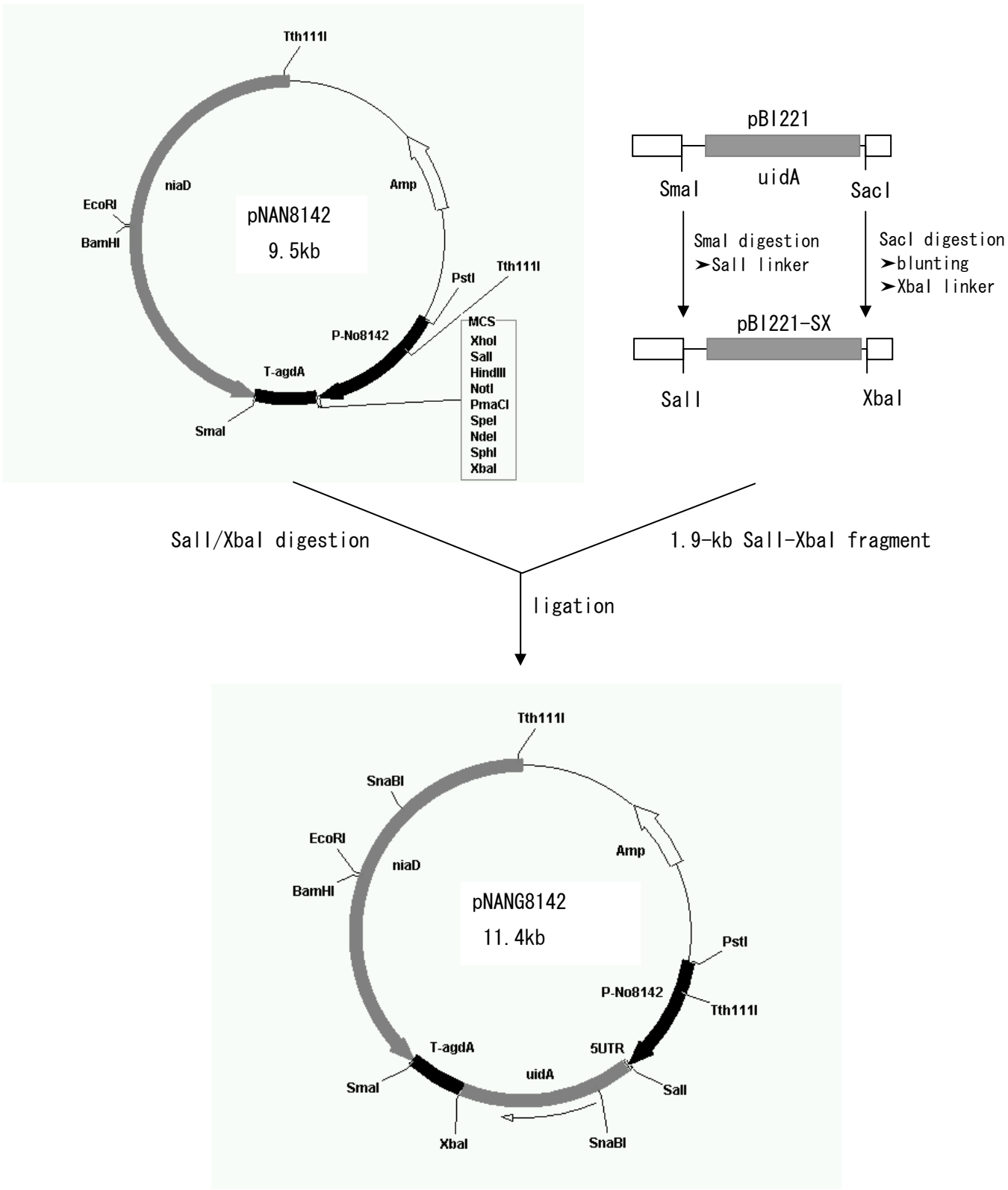

FIG. II.3. Construction of the plasmid pNANG8142 (see text for details).

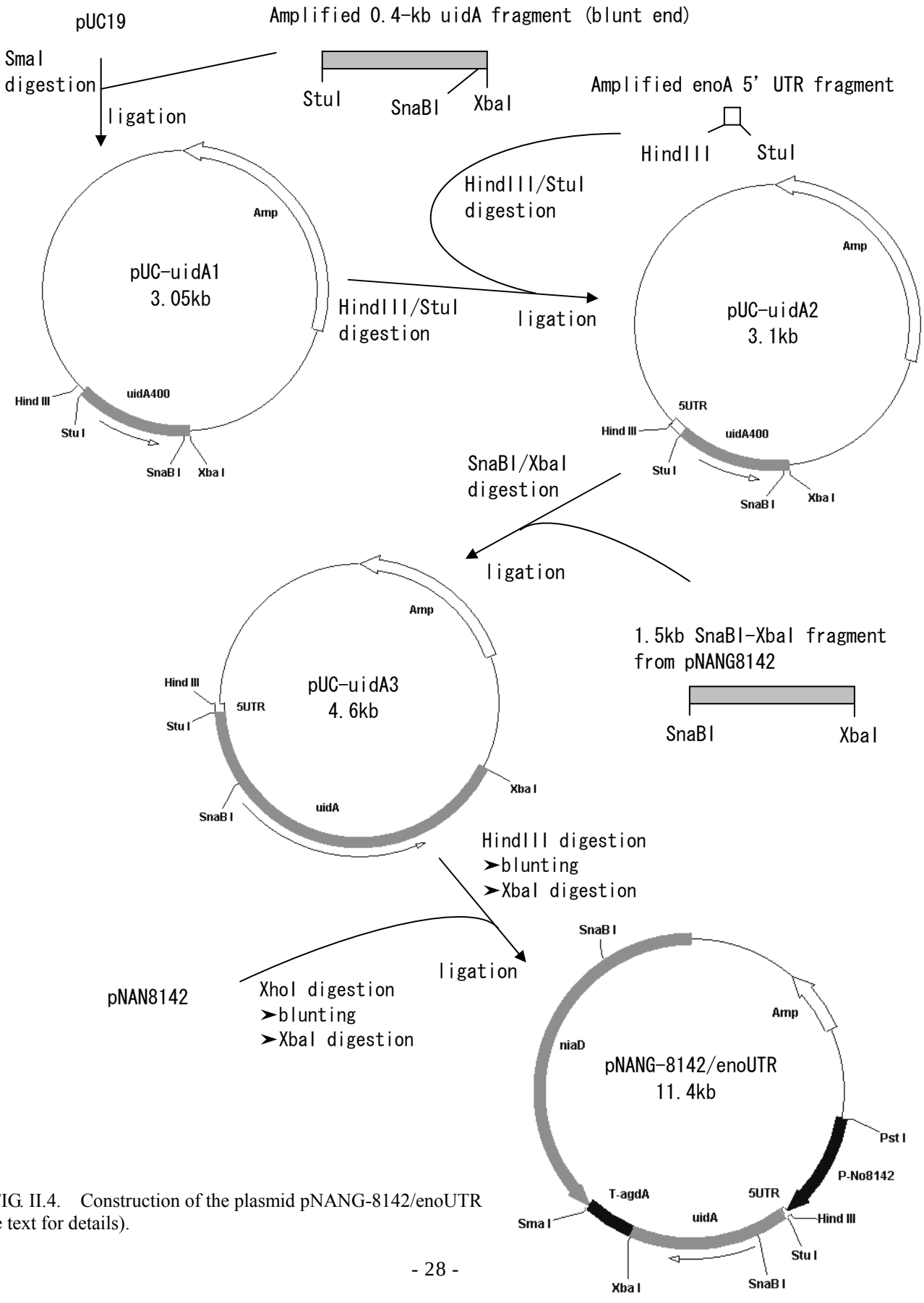

FIG. II.4. Construction of the plasmid pNANG-8142/enoUTR (see text for details).

the *Sma*I site of pUC19, resulting in pUC-uidA1 (Fig. II.4). The *enoA* 5′UTR was amplified by PCR on *A. oryzae* niaD300 genomic DNA using the primers 5′-ACT*AAGCTT*CTTCCGTCCTCCAAGTTAGTC-3′ (the synthesized *Hin*dIII site is in italics) and 5′-CTT*AGGCCT*TAACATTTTGACGAGCTGCGGAATTG-3′ (the synthesized *Stu*I site is in italics). The amplified 72-bp 5′UTR fragment of the *enoA* gene contained its transcription start sites (Machida et al. 1996). This fragment was digested with *Hin*dIII/*Stu*I and cloned into *Hin*dIII/*Stu*I-digested pUC-uidA1, resulting in pUC-uidA2 (Fig. II.4). The 1.5-kb *Sna*BI-*Xba*I restriction fragment (a portion of the *uidA* gene) from pNANG8142 was cloned into *Sna*BI/*Xba*I-digested pUC-uidA2 to obtain pUC-uidA3 (Fig. II.4). Finally, the 1.9-kb *Hin*dIII (blunted) /*Xba*I restriction fragment from pUC-uidA3, containing the *enoA* 5′UTR::GUS (*uidA*) chimeric gene construct, was cloned into *Xho*I (blunted) /*Xba*I-digested pNAN8142 to obtain pNANG-8142/enoUTR (the ends were blunted by filling-in) (Fig. II.4). The *Hin*dIII site (AAGCTT) was regenerated by this blunt-end ligation. The chimeric construct and the 5′UTR sequence of pNANG-8142/enoUTR are shown in Fig. II.5.

pNANG-8142UTR, containing the complete 5′UTR derived from the original P-No8142 promoter, was constructed as follows: using pNAN8142 as a template, a 1.3-kb fragment, including the P-No8142 promoter and its complete 5′UTR, was amplified with the primers 5′-TAG*CTGCAG*CGGGATTGCCATTG-3′ (*Pst*I site is in italics) and 5′-ACT*AGGCCT*TAACATGCTGACAGGGGAGTGTGCGCGATTGACGCGCAAGA TAGAAAAGAAAGGTAC-3′ (the synthesized *Stu*I site is in italics). The amplified fragment was digested with *Pst*I/*Stu*I and cloned into *Pst*I/*Stu*I-digested pNANG-8142/enoUTR (replacement of the *Pst*I-*Stu*I fragment) to obtain

pNANG-8142UTR. The 5′UTR sequence of pNANG-8142UTR are shown in Fig. II.5.

pNANG-enoUTR, in which most of the 5′UTR derived from the P-No8142 promoter was replaced by the *enoA* 5′UTR, was constructed as follows: using pNAN8142 as a template, a 1.2-kb fragment including the P-No8142 promoter and only the first 15-bp of its 5′UTR was amplified with the primers 5′-TAG*CTGCAG*CGGGATTGCCATTG-3′ (*Pst*I site is in italics) and 5′-ACT*AAGCTT*GAAGGGCAAGGAAGTGGGAGGA-3′ (the synthesized *Hin*dIII site is in italics). The amplified fragment was digested with *Pst*I/*Hin*dIII and cloned into *Pst*I/*Hin*dIII-digested pNANG-8142/enoUTR (replacement of the *Pst*I-*Hin*dIII fragment) to obtain pNANG-enoUTR. The 5′UTR sequence of pNANG-enoUTR are shown in Fig. II.5. All fragments obtained by PCR were sequenced completely to exclude mutations within the sequences.

Fungal transformation Transformation of *A. oryzae* was done by the method of Gomi et al. (1987).

Southern blot analysis *A. oryzae* transformants were grown in DP medium with shaking. Preparation of genomic DNA from the transformants was carried out as described by Tsuchiya et al. (1992). Probe labeling and detection were done by use of AlkPhos direct labelling and detection system (Amersham Biosciences).

Northern blot analysis *A. oryzae* transformants were cultured in 15 ml of DP medium for 40 h at 30°C with shaking, after which the mycelia were transferred into fresh DP medium and cultured for 8 h. Then, total RNA was isolated with an ISOGEN RNA isolation kit (Nippon Gene). Total RNA (10 μg) was separated on a formaldehyde gel and transferred onto a Hybond-N^+ membrane (Amersham

Biosciences), using the capillary transfer method. The entire *uidA* fragment from pNANG8142 was used as a probe. Probe labeling and detection were done by use of AlkPhos direct labelling and detection system (Amersham Biosciences).

GUS assay *A. oryzae* transformants were cultured in 15 ml of DP medium at 30°C with shaking. After 40 h of cultivation, the mycelia were harvested and washed with distilled water. Preparation of cell-free extracts of the transformants was done as described by Tada et al. (1991). GUS activity of the cell-free extract was measured according to Jefferson et al. (1986), using *p*-nitrophenyl glucuronide as the substrate. One unit was defined as the amount of enzyme producing 1 nmol/min of *p*-nitrophenol at 37°C. Total protein was determined with a protein assay kit (Bio Rad), using bovine serum albumin as a standard.

SDS-PAGE analysis Electrophoresis was performed according to the method of Laemmli (1970) in 10% polyacrylamide gels. Proteins were stained with Coomassie brilliant blue R-250.

Results

Effect of the 5′UTRs on GUS expression in *A. oryzae* To investigate whether GUS protein expression in *A. oryzae* is influenced by the mRNA 5′UTR sequence, four expression vectors with different 5′UTRs fused upstream of the *E. coli* gene encoding GUS (*uidA*) were constructed (Fig. II.5). All four constructs used the strong promoter P-No8142, which is derived from the No. 8AN promoter isolated from *A. niger* (Ozeki et al. 1996). First, the control expression vector pNANG8142, which contains most of the 5′UTR of the P-No8142 promoter as the cap-proximal side of the 5′UTR; and 17 bp

of non-fungal sequence from pBI221 served as the distal portion of the 5′UTR, was constructed. Second, pNANG-8142UTR contained the complete 5′UTR derived from the original P-No8142 promoter. The remaining two vectors contained sequences from the 5′UTR of the *A*. *oryzae* enolase gene (*enoA*), which is one of the most strongly expressed genes in *A*. *oryzae* (Machida et al. 1996). In pNANG-8142/enoUTR and pNANG-enoUTR, part of the P-No8142 5′UTR or nearly all of this 5′UTR, respectively, was replaced by the *enoA* 5′UTR. It was confirmed that no uAUGs or uORFs were introduced within the 5′UTRs of these constructs, because such sequences can hamper translation initiation of the main ORF in eukaryotes (Kozak 1991b; Morris and Geballe 2000), including filamentous fungi (Vilela and McCarthy 2003).

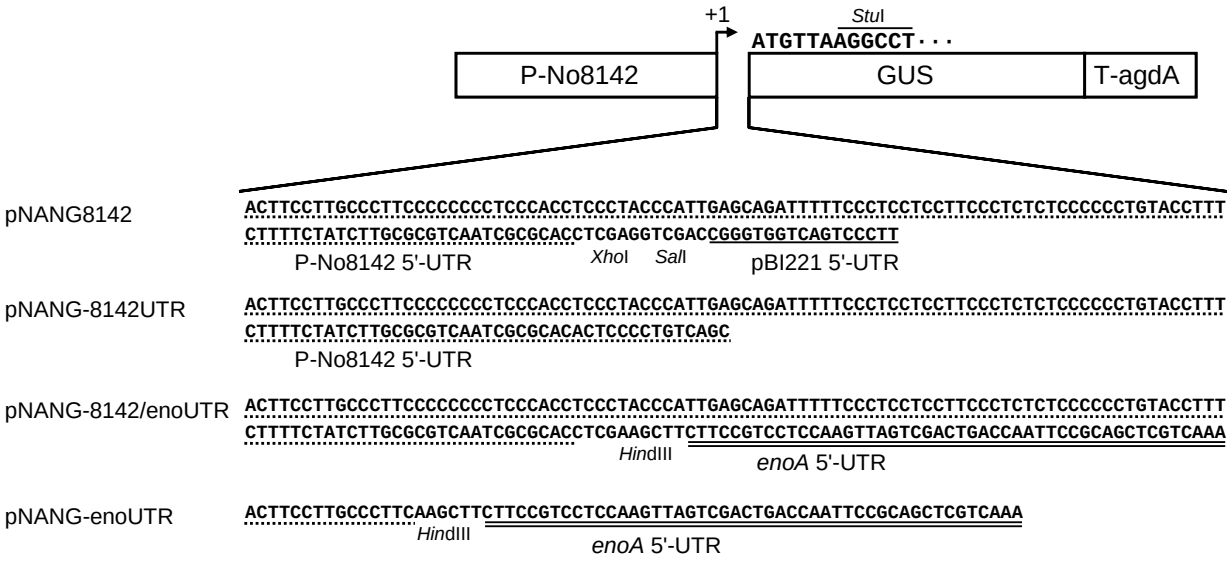

FIG. II.5. Structure of the promoter (P-No8142)::5′UTR::GUS chimeric gene constructs and their entire 5′UTR sequences. T-agdA is the transcriptional terminator from the *A. oryzae agdA* gene. An *arrow* indicates the major transcription start site determined by RNA ligase-mediated rapid amplification of 5′ cDNA ends (data not shown). Translation of all constructs starts from the AUG initiation codon of the GUS gene. Sequences derived from the P-No8142 5′UTR are indicated by *broken lines*. Relevant restriction sites in the plasmids are indicated (see text for details).

These plasmids were introduced into the *A. oryzae* niaD300 chromosome by means of *niaD*-based homologous recombination. Transformants having one copy of the plasmid at the *niaD* locus were selected by Southern blot analysis (data not shown). These transformants were subjected to GUS assay. Figure II.6A shows the specific GUS activity of the transformants containing the different 5′UTR constructs. Alterations in the 5′UTR, including partial or nearly complete replacement of the 5′UTR, significantly affected GUS protein expression with no change in the cellular growth rate. GUS activity was strongly increased in transformants containing pNANG-8142UTR, pNANG-8142/enoUTR, and pNANG-enoUTR, compared with the control expression vector pNANG8142. Approximately 4-fold higher GUS activity was observed with pNANG-8142UTR and pNANG-8142/enoUTR and 8-fold higher with pNANG-enoUTR. The GUS activity resulting from this construct (40,000 units/mg protein) was approximately 40-fold higher than the activities reported previously for widely used strong promoters, such as the *amyB* promoter (Tada et al. 1991) and the *glaA* promoter (Hata et al. 1992). Figure II.6B shows a SDS-PAGE analysis of cell-free extracts. As expected, there was good correlation between the amount of GUS protein seen in the gel and the enzyme activity.

The 5′UTR-specific enhancement is post-transcriptional To investigate whether the enhancements in GUS activity mediated by these 5′UTR constructs resulted from transcriptional or post-transcriptional effects, we determined the levels of GUS transcripts in the transformants harboring these 5′UTR constructs (Fig. II.6C). The levels of GUS transcripts were the same for all transformants. This indicates that the enhancements in GUS activity were not due to increased transcription or increased mRNA stability. In addition, alterations in the 5′UTR in these plasmids did not cause

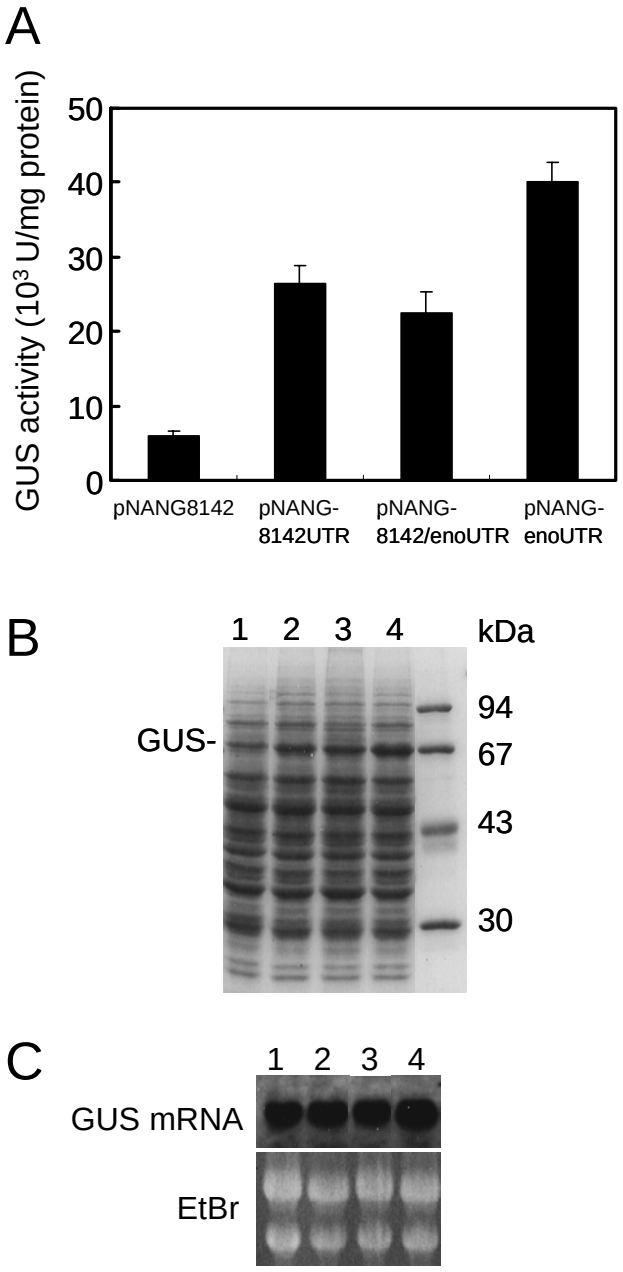

FIG. II.6. Effect of the 5′UTRs on translation efficiency in *A. oryzae*. (A) GUS activity of transformants containing the different 5′UTR constructs. Values are the means of three independent experiments; and error bars represent the standard deviation. (B) SDS-PAGE analysis of cell-free extracts. The cell-free extracts were the same as those used for the GUS assays in (A). Each lane was loaded with 10 μg of protein. The positions of the molecular mass markers are shown on the *right*. Lanes 1–4 correspond to pNANG8142, pNANG-8142UTR, pNANG-8142/enoUTR, and pNANG-enoUTR, respectively. (C) Northern blot analysis of GUS transcripts. Ethidium bromide-stained ribosomal RNA is shown as a loading control. Lanes 1–4 correspond to pNANG8142, pNANG-8142UTR, pNANG-8142/enoUTR, and pNANG-enoUTR, respectively.

any change in the amino acid sequence of the GUS protein (Fig. II.5). Thus, it is very likely that the 5′UTRs in pNANG-8142UTR, pNANG-8142/enoUTR, and pNANG-enoUTR enhanced GUS expression by enhancing translation.

GUS overproduction in a transformant carrying multiple copies of pNANG-enoUTR As shown in Fig. II.6A, the *enoA* 5′UTR construct (pNANG-enoUTR) dramatically enhanced translation in *A*. *oryzae*. To further examine the ability of this 5′UTR construct to direct GUS overproduction in *A*. *oryzae*, a transformant, strain AO9-pNANG-enoUTR, carrying 4 or 5 copies of pNANG-enoUTR, was isolated according to Southern blot analysis. The growth rate of the strain AO9-pNANG-enoUTR was similar to that of the recipient strain. When grown on DP medium for 40 h, the strain AO9-pNANG-enoUTR produced remarkably high GUS activity (Fig. II.7A). GUS production was also examined by SDS-PAGE analysis (Fig. II.7B). Quantification of GUS protein in the gel using the public domain ImageJ program (http://rsb.info.nih.gov/ij/) indicated that GUS made up more than 50% of the total soluble protein in this strain. This is the highest level of GUS production ever reported in filamentous fungi.

A

Strain	Plasmid copy number	GUS activity (U/mg protein)
Recipient	–	<0.1
AO9-pNANG-enoUTR	4–5	161,000 ± 15,200

B

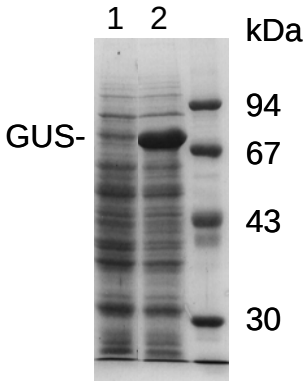

FIG. II.7. GUS overproduction in a transformant carrying multiple copies of pNANG-enoUTR. (A) GUS activity of the transformant (AO9-pNANG-enoUTR). GUS activity is presented as the mean value and the standard deviation from three independent measurements. Plasmid copy number was determined by Southern blot analysis. (B) SDS-PAGE analysis of cell-free extracts. The cell-free extracts were the same as those used for the GUS assays in (A). Each lane was loaded with 10 μg of protein. Lane 1, recipient strain; lane 2, AO9-pNANG-enoUTR.

Discussion

In this chapter, it was investigated whether the 5′UTR of mRNA has a large influence on translation efficiency in *A. oryzae*. Using the GUS reporter gene assay, various constructs containing the same P-No8142 promoter with different 5′UTR sequences were analyzed. Alterations in the 5′UTR significantly affected GUS protein expression. Although there is a possibility either that transcription might be up-regulated through the binding of *trans*-factors to *cis*-elements within the 5′UTR or that mRNA stability may be enhanced, no significant difference was detected in the

level of GUS transcripts among all constructs. These results clearly show that the enhancements in expression were due to an increase in translation efficiency.

How did the alterations in the 5′UTR enhance the translation efficiency? Characteristics of the 5′UTR reported to affect translation efficiency include length, start-site consensus sequences, GC content, and secondary structure (Kozak 1991b). Statistically, 5′UTRs that enable efficient translation in eukaryotes are short (Kochetov et al. 1998), although recognition of the first AUG codon may be impaired if it is positioned too close to the cap because scanning ribosomes are likely to bypass it (Kozak 1991b). There are, however, exceptions to this generalization. For example, extending the 5′UTR from 41 bases to 69 bases resulted in a 10-fold increase in reporter activity in tobacco protoplasts (Dansako et al. 2003). In the present study, the highest GUS activity was conferred by the shortest 5′UTR tested: pNANG-enoUTR (5′UTR 69 bases long), which conferred about 8-fold higher GUS activity compared with pNANG8142 (145 bases). However, the GUS activity resulting from the longest 5′UTR (174 bases, in pNANG-8142/enoUTR) was about 4-fold higher than that from pNANG8142. Thus, the translational enhancements cannot be explained simply by differences in 5′UTR length.

The most likely explanation is context effects on AUG codon recognition. The efficiency with which the 43S preinitiation complex recognizes the AUG codon depends on the nucleotides surrounding that codon (Kozak 1986). A consensus sequence is GCCA/GCCaugG for vertebrates (Kozak 1987), and A/UAA/CAA/CAaugUCU/C for yeast (Hamilton et al. 1987). The most highly conserved position within these consensus sequences is the purine, usually A, in position −3 (3 nt upstream of the AUG codon, which covers positions +1 to +3). The importance of the −3 position was tested

in yeast by substituting C, G, and U for the A at the −3 position in the *CYC7* gene (Yun et al. 1996). With AAC (sequence in front of the AUG codon) designated as 100%, the relative levels of translation for A:G:C:U were 100:93:70:54. In contrast, the context in higher eukaryotes can modulate translation at least 10-fold (Kozak 1986). In filamentous fungi, UCAA/CA/CaugG/U has been reported as a consensus (Ballance 1986). Recent database survey of more than 100 *A. oryzae* gene and mRNA sequences (including hypothetical genes) revealed that −3 A and −4 C were highly conserved and the other positions were less conserved (data not shown), suggesting that −3 A and −4 C are important for efficient recognition of the AUG codon in *A. oryzae*. In this study, pNANG-8142UTR, pNANG-8142/enoUTR, and pNANG-enoUTR have the consensus sequence at both −3 and −4, whereas pNANG8142, which conferred much lower GUS activity, has a C at −3. This observation suggests that the presence of *A. oryzae* consensus sequences surrounding the AUG codon contributed to efficient translation.

One other possible explanation is RNA secondary structure. Secondary structure between the cap and AUG codon can inhibit translation initiation. It is suggested that such structure impedes binding or scanning of the 40S ribosomal subunit (Kozak 1991b). The secondary structure of each 5′UTR used in this study was predicted by computer analysis with the mfold program (Zuker 2003), and the free energy value (ΔG) was calculated (Table II.1). The 5′UTR which had the lowest stability in its secondary structure showed the highest GUS activity (Fig. II.6A). This observation is in agreement with the generally accepted view that the removal of secondary structures within the 5′UTR often results in higher translation.

TABLE II.1. Structural features of the 5′UTRs used in this study

Plasmid	5′UTR		
	Length (bases)	GC content (%)	ΔG[a] (kcal/mol)
pNANG8142	145	59	−27.2
pNANG-8142UTR	130	58	−12.0
pNANG-8142/enoUTR	174	55	−21.0
pNANG-enoUTR	69	51	−5.0

[a] Free energy of the secondary structure. mRNA secondary structure prediction was performed with the mfold program ver. 3.1 (Zuker 2003), using default settings.

In addition to above-mentioned explanations, there are some other possibilities to account for the enhancement obtained. Compared with pNANG-8142/enoUTR, pNANG-enoUTR had a deletion of 105 bp of P-No8142 5′UTR; and it conferred GUS activity about 2-fold higher. The enhancement of translation obtained by deletion of part of the P-No8142 5′UTR suggests that the original P-No8142 5′UTR may contain specific sequences or structures that function as binding sites for regulatory proteins that decrease translation efficiency. The P-No8142 promoter is a derivative of the No. 8AN promoter isolated as a strong promoter from *A. niger*, although it is unknown what gene is expressed under the control of this promoter (Ozeki et al. 1996). The P-No8142 promoter has a long C+T-rich motif just downstream of the transcription start site (Fig. II.5). Such C+T-rich motifs are often found in highly expressed genes in *Aspergillus* species and they may contribute to efficient transcription. In addition, the present results raise the possibility that the C+T-rich motif within the 5′UTR may decrease translation efficiency.

Taken together, these results suggest that length, AUG context, secondary structure, and putative regulatory sequences may contribute cooperatively to overall

translation efficiency.

In eukaryotes, it is generally accepted that translation initiation is an important step in both global and mRNA-specific gene regulation. However, in filamentous fungi, not even the basic mechanism of translation initiation is known; and, so, little attention has been paid to the efficiency of translation initiation in heterologous gene expression. In this study, using a model GUS expression system, it was demonstrated that the 5′UTR of mRNA plays an important role in determining the translation efficiency in *A*. *oryzae*, although additional work is required to elucidate the exact mechanisms underlying this phenomenon. Furthermore, from a comparison of pNANG-8142UTR and pNANG-enoUTR (Fig. II.6A), it is speculated that the *enoA* 5′UTR can promote translation initiation more effectively than the P-No8142 5′UTR and that the 5′UTRs of specific genes may serve as important regulators of gene expression in *A*. *oryzae*.

To my knowledge, this is the first report of alterations in a 5′UTR increasing translation efficiency of a heterologous gene in filamentous fungi. The findings provide a novel approach for improving recombinant protein yield in filamentous fungi.

Chapter III

Effect of the 5′ untranslated region of the *Hsp12* gene on translation efficiency in *Aspergillus oryzae*

Introduction

In Chapter II, it was demonstrated that the 5′UTR plays an important role in determining translation efficiency in the filamentous fungus *A. oryzae* and that it is feasible to improve recombinant protein yield by promoting translation initiation. Finding 5′UTRs that strongly increase translation efficiency would allow the development of a stronger expression system. In plants, Dansako et al. (2003) showed that the 5′UTRs of several *Arabidopsis* heat-shock protein (Hsp) genes contribute to efficient expression of a foreign gene by enhancing translation. In this chapter, it was investigated whether the 5′UTRs of the fungal Hsp genes possess the ability to enhance gene expression at the translational level in *A. oryzae*.

Materials and Methods

Strains and media *A. oryzae* niaD300 was used as a recipient strain for transformation (see Materials and Methods of Chapter II). *E. coli* DH5α was used for construction and propagation of plasmids. The medium for fungal cultivation in flasks was dextrin-peptone (DP) (see Materials and Methods of Chapter II).

Hsp 5′UTR sequences employed in this study The *A. oryzae* EST database

(http://www.aist.go.jp/RIODB/ffdb/index.html) (Machida 2002) was searched using the BLAST algorithm for cDNA sequences showing high homology to the *Saccharomyces cerevisiae HSP12* gene (GenBank accession number X55785), the *Aspergillus nidulans Hsp30* gene (GenBank accession number D32070), and the *Neurospora crassa Hsp70* gene (GenBank accession number U10443). This search yielded three candidate sequences: a putative *Hsp12* gene, a putative *Hsp30* gene (the complete mRNA sequence was recently submitted to GenBank by Matsushita et al. with accession number AB126868), and a putative *Hsp70* gene. Next, from the EST (cDNA) sequence data, the 5′UTRs of each gene were predicted. These putative 5′UTR sequences were amplified by PCR from *A*. *oryzae* niaD300 genomic DNA using the following primer pairs:
5′-TAG*AAGCTT*ATCAATTTCGACAAGACAAACAAACAAATC-3′ (an added *Hin*dIII site is in italics) and 5′-ACT*AGGCCT*TAACATGTTTGCGAGTGGTTTG-3′ (added *Stu*I site in italics) for the *Hsp12* 5′UTR;
5′-TAG*AAGCTT*ATCAGAACAACAACAATCAAACAC-3′ (added *Hin*dIII site in italics) and 5′-ACT*AGGCCT*TAACATTTTGGCTGTGTGTTGA-3′ (added *Stu*I site in italics) for the *Hsp30* 5′UTR; and
5′-TAG*AAGCTT*ACTCACTTGGTCTCATCCTCTATAG-3′ (added *Hin*dIII site in italics) and 5′-ACT*AGGCCT*TAACATTGTGAAAAGAGTAGGAG-3′ (added *Stu*I site in italics) for the *Hsp70* 5′UTR. All fragments obtained by PCR were sequenced completely to exclude mutations within the sequences.

Construction and transformation of GUS expression vectors containing different 5′UTRs The starting vector for construction of the expression vectors was pNANG-enoUTR (described in Chapter II) (Fig. III.1A), which contains the strong

promoter P-No8142, the *A. oryzae enoA* 5′UTR, and the *E. coli uidA* gene, encoding β-glucuronidase (GUS), as reporter.

For construction of pNANG-hsp12UTR containing the *Hsp12* 5′UTR, the PCR-amplified fragment of the *Hsp12* 5′UTR described above was digested with *Hin*dIII/*Stu*I and cloned into *Hin*dIII/*Stu*I-digested pNANG-enoUTR for replacement of the 5′UTR sequence (Fig. III.1B). pNANG-hsp30UTR, containing the *Hsp30* 5′UTR, and pNANG-hsp70UTR, containing the *Hsp70* 5′UTR, were made similarly (Fig. III.1B).

Transformation of *A. oryzae* was done as described in Materials and Methods of Chapter II.

Southern blot analysis Southern blot analysis was done as described in Materials and Methods of Chapter II.

Northern blot analysis *A. oryzae* transformants were cultured in 15 ml of DP medium for 40 h at either 30°C or 37°C with shaking. Then, total RNA was isolated with an ISOGEN RNA isolation kit (Nippon Gene). Total RNA (10 μg) was separated on a formaldehyde gel and transferred onto a Hybond-N^+ membrane (Amersham Biosciences), using the capillary transfer method. The entire *uidA* fragment from pNANG-enoUTR was used as a probe. Probe labeling and detection were done as described in Materials and Methods of Chapter II.

GUS assay *A. oryzae* transformants were cultured in 15 ml of DP medium at either 30°C or 37°C with shaking for 40 h. Preparation of cell-free extracts and GUS assay of the transformants were done as described in Materials and Methods of Chapter II.

Results

The *Hsp12* 5′UTR increases GUS protein expression in *A. oryzae* To investigate whether the 5′UTRs of the *A. oryzae* Hsp genes can increase the efficiency of exogenous protein expression, a series of 5′UTR::GUS (*uidA*) fusions having different 5′UTRs derived from Hsp genes was constructed (Fig. III.1B). As control (nonheat-shock) 5′UTR constructs, three expression vectors, pNANG8142,

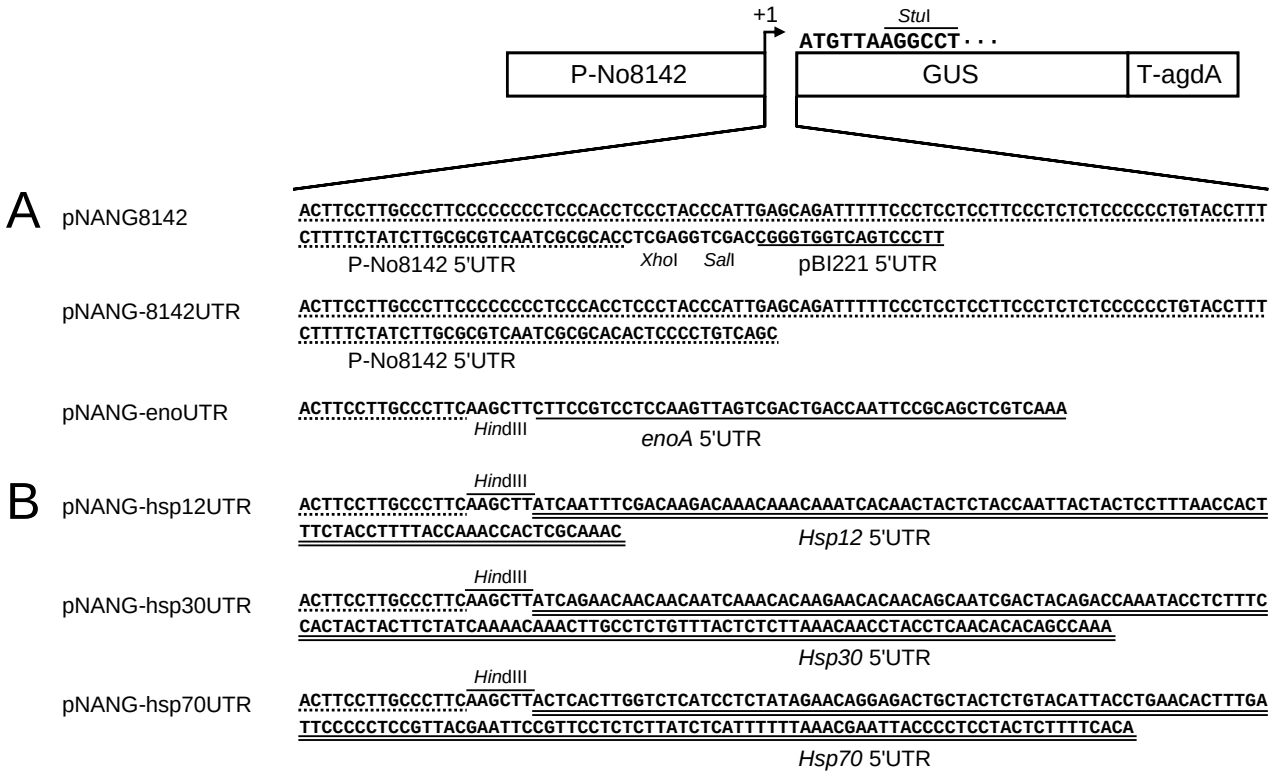

FIG. III.1. Structure of the chimeric gene constructs and their 5′UTR sequences. The promoter is P-No8142. T-agdA is the transcriptional terminator from the *A. oryzae agdA* gene. An *arrow* indicates the major transcription start site determined by RNA ligase-mediated rapid amplification of 5′ cDNA ends (data not shown). Translation of all constructs starts from the AUG initiation codon of the GUS gene. Sequences derived from the P-No8142 5′UTR are indicated by *broken underlining*. Relevant restriction sites in the plasmids are indicated. (A) Sequences of the nonheat-shock 5′UTR constructs. (B) Sequences of the Hsp 5′UTR constructs.

pNANG-8142UTR, and pNANG-enoUTR were used (Fig. III.1A, also see Chapter II). pNANG8142 contained most of the 5′UTR of the P-No8142 promoter as the cap-proximal side of the 5′UTR; and 17 bp of non-fungal sequence from pBI221 (Jefferson et al. 1987) served as the distal portion of the 5′UTR. pNANG-8142UTR contained the complete 5′UTR derived from the original P-No8142 promoter. pNANG-enoUTR contained sequences from the 5′UTR of the *A. oryzae* enolase gene (*enoA*), which is one of the most strongly expressed genes in *A. oryzae* (Machida et al. 1996). All these constructs contained the strong promoter P-No8142.

These plasmids were introduced into the *A. oryzae* niaD300 chromosome by means of *niaD*-based homologous recombination. Transformants having one copy of the plasmid at the *niaD* locus were selected by Southern blot analysis (data not shown).

Hsp proteins are known to be transiently induced under a wide variety of stress conditions to protect cells against the stress, so the 5′UTRs of the Hsp genes may function more effectively under stressful conditions such as heat shock. Therefore, the effect of these Hsp 5′UTRs on GUS expression was examined at different temperatures. Figure III.2A shows the specific GUS activity of the transformants containing the different 5′UTR constructs at 30°C and 37°C. In all transformants, the amounts of total soluble protein per cell were similar between both temperature conditions, although the growth rates were a little slower at 37°C than at 30°C. At both normal (30°C) and high (37°C) temperatures, GUS activity strongly depended on the type of 5′UTR. The Hsp 5′UTRs, especially *Hsp12* and *Hsp30* 5′UTRs, strongly enhanced GUS expression at both temperature conditions. Interestingly, the highest GUS activity was observed with the *Hsp12* 5′UTR construct (pNANG-hsp12UTR) at 37°C; expression was increased about 20- and 1.7-fold compared to the nonheat-shock control

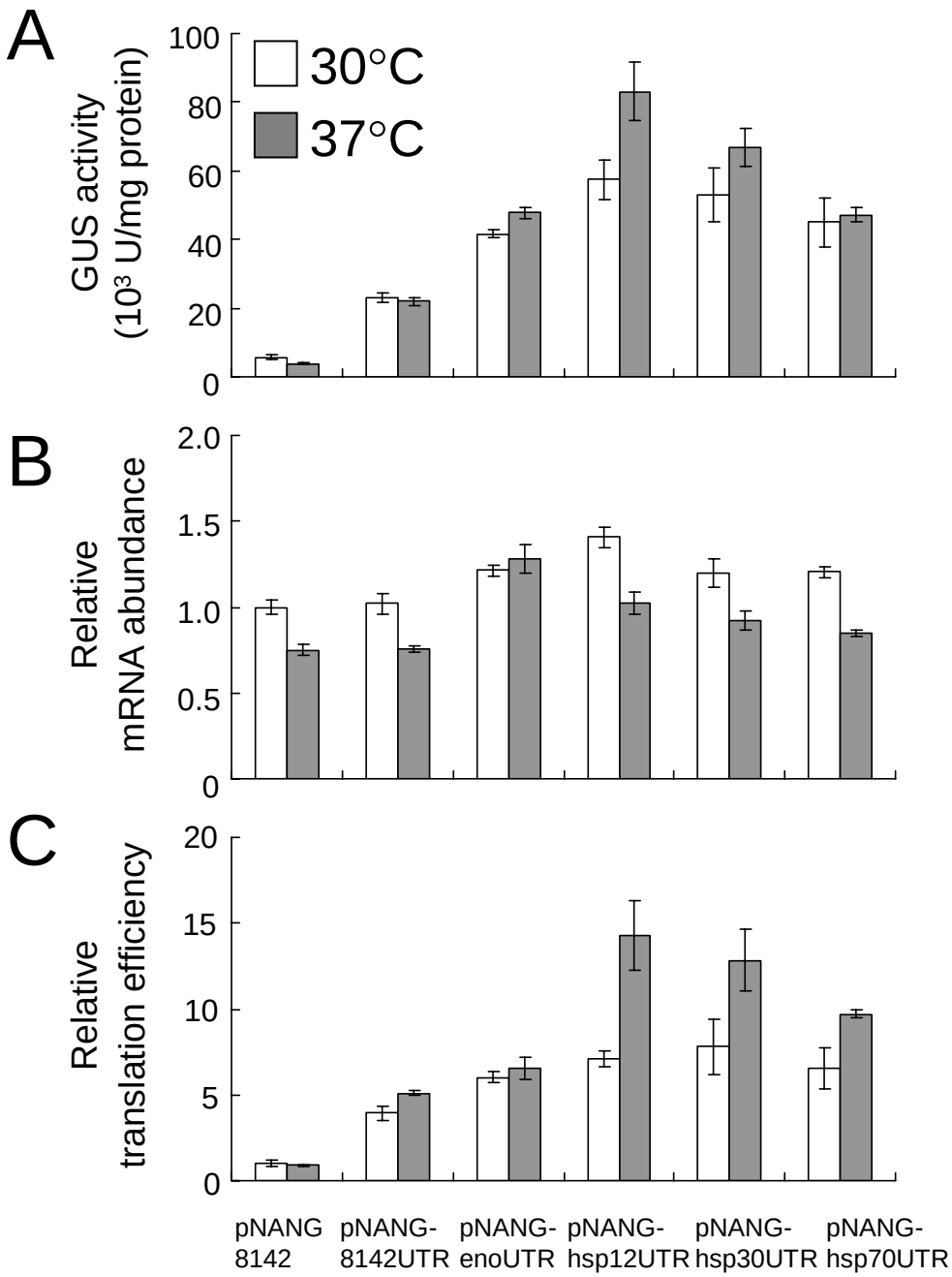

FIG. III.2. Effect of the 5′UTRs on translation efficiency in *A. oryzae* at different temperatures. (A) GUS activity of transformants containing the different 5′UTR constructs at 30°C (*white*) and 37°C (*gray*). Values are the means of three independent experiments, and error bars represent the standard deviation. (B) Relative GUS mRNA levels of transformants containing the different 5′UTR constructs at 30°C (*white*) and 37°C (*gray*). RNA was isolated from the same mycelia samples used for GUS assay in (A) and assayed by Northern blot. GUS mRNA levels for each construct were quantified using the public domain ImageJ program (http://rsb.info.nih.gov/ij/) and normalized to the intensity of the corresponding 18S rRNA band. GUS mRNA levels were calculated relative to the average value of pNANG8142 at 30°C. Values are the means of three independent experiments, and error bars represent the standard deviation. (C) Relative translation efficiency of the different mRNAs (5′UTR constructs) at 30°C (*white*) and 37°C (*gray*). The relative translation efficiency (GUS activity/mRNA) was calculated as the ratio of relative GUS activity to relative GUS mRNA level (shown in (A) and (B), respectively), with the efficiency of pNANG8142 at 30°C defined as 1.0. Values represent data from three determinations, and error bars represent standard deviations.

constructs pNANG8142 and pNANG-enoUTR, respectively. The level of GUS productivity resulting from this construct (83,000 units/mg protein) was similar to that reported quite recently with the *sodM* promoter (Ishida et al. 2004) and is the highest level ever reported in filamentous fungi for a transformant with a single copy of the GUS expression cassette.

High expression at 37°C mediated by the *Hsp12* 5′UTR is due to efficient translation The *Hsp12* and *Hsp30* 5′UTRs may have increased GUS expression by transcriptional activation, stabilization of the transcripts, and/or elevation of translation efficiency. To assess the translation efficiency of the mRNAs containing different 5′UTRs, the levels of GUS transcripts in the transformants harboring these 5′UTR constructs were determined (Fig. III.2B), and relative translation efficiencies were calculated (see the legend in Fig. III.2C). At 30°C, as seen in Fig. III.2B, GUS transcript abundance was slightly affected by the 5′UTR (up to 1.4-fold). Translation efficiency, however, was much more strongly affected by the 5′UTR (Fig. III.2C). At 30°C, the *Hsp12* 5′UTR construct (pNANG-hsp12UTR) and the *Hsp30* 5′UTR construct (pNANG-hsp30UTR) had approximately eightfold higher translation efficiency compared to that for pNANG8142 and were slightly higher than that for pNANG-enoUTR. Thus, both increased mRNA abundance, and higher translation efficiencies contributed to efficient expression of GUS protein at 30°C mediated by the *Hsp12* and *Hsp30* 5′UTRs. Furthermore, pNANG-hsp12UTR and pNANG-hsp30UTR displayed a significant increase (approximately twofold) in the translation efficiency at 37°C compared to 30°C, whereas for the other constructs, the translation efficiencies at 37°C were very similar to or only slightly higher than those at 30°C (Fig. III.2C). These results apparently indicate that increased translational

efficiency at elevated temperature (37°C) accounts for the much more efficient expression of GUS protein mediated by the *Hsp12* and *Hsp30* 5′UTRs seen in Fig. III.2A.

Discussion

What is the mechanism for more efficient translation at the elevated temperature observed with the *Hsp12* and *Hsp30* 5′UTRs? The simplest explanation is that increased temperature alters the RNA structure. It is generally accepted that secondary structure in the 5′UTR largely affects translation efficiency. The *Hsp12* and the *Hsp30* 5′UTRs have comparatively high AT content, and the stability of secondary structure at 37°C predicted with the mfold program (Zuker 2003) is lower than that of the other 5′UTRs (data not shown). As a result, it is possible that such structural change might make the 5′UTRs more accessible to ribosomes and/or facilitate ribosome 'scanning' compared to the other 5′UTRs at 37°C. Recent studies in bacteria revealed that mRNA secondary structure can act as a direct temperature sensor, with conformational change in the mRNA regulating translation (Morita et al. 1999; Johansson et al. 2002). Further work is required to elucidate whether and how temperature-dependent structural change in the 5′UTR affects translation in filamentous fungi. In addition, the predicted stability of secondary structure of 5′UTRs does not always correlate well with the translation efficiency, suggesting that the translation efficiency cannot be explained by the 5′UTR secondary structure alone. Another possible explanation for the higher translation efficiency at 37°C is that temperature-specific factors may interact with regulatory sequences or structures within the *Hsp12* 5′UTR and the *Hsp30* 5′UTR.

Interestingly, the *Hsp12* 5′UTR and the *Hsp30* 5′UTR have several A+C-rich stretches (Fig. III.3). Such sequences may function as *cis*-elements that contribute to efficient translation at 37°C. Additional work, such as point mutation analysis and deletion analysis, will give us information on the molecular mechanisms involved.

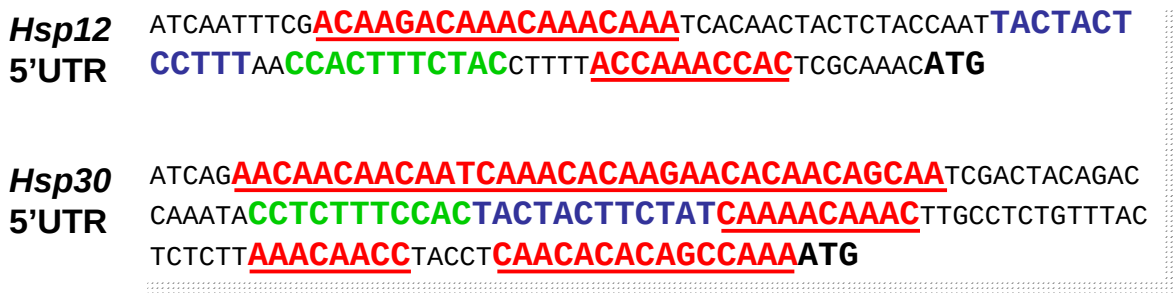

FIG. III.3. Sequences and putative motifs in the *Hsp12* and *Hsp30* 5′UTRs (see text).

In addition to above-mentioned explanations, there are some other possibilities to account for the translational enhancement obtained with the *Hsp12* and the *Hsp30* 5′UTRs. The majority of eukaryotic mRNAs initiate translation by a cap-dependent mechanism ‘ribosome scanning’ (Kozak 1989). However, translation of some mRNAs can occur in a cap-independent manner termed ‘internal ribosome entry site (IRES)-dependent translation’ (Hellen and Sarnow 2001). It was reported that some types of IRES are activated under heat shock stress (Kim and Jang 2002), although it is not clear how heat shock enhances internal ribosomal initiation directed by IRES elements. An interesting question raised by the present data is whether the *Hsp12* and the *Hsp30* 5′UTRs have an IRES-like element to mediate cap-independent translation

that contributes to efficient translation at 37°C. Traditionally, the 5′UTRs containing IRES usually have significant secondary structure spanning approximately over several hundred bases and punctuated by multiple uAUGs. On the other hand, the *Hsp12* and the *Hsp30* 5′UTRs are relatively short and do not form a stable secondary structure (data not shown). In addition, IRES-dependent translation has rarely been documented in filamentous fungi. Therefore, further work is required not only to confirm the hypothesis in the case of the Hsp 5′UTRs, but also to provide significant new insight into regulatory mechanisms of gene expression in filamentous fungi.

In this chapter, it was demonstrated that the *Hsp12* and the *Hsp30* 5′UTRs strongly enhanced the translation efficiency of the chimeric GUS mRNA compared to the control, nonheat-shock 5′UTRs at 30°C. A further observation is that temperature had differential effects on translation efficiency of different mRNAs, and the enhanced translation at elevated temperature (37°C) mediated by the *Hsp12* and *Hsp30* 5′UTRs contributed to much more efficient expression of GUS protein. This phenomenon, previously unrecognized in filamentous fungi, may be an important posttranscriptional mechanism in the regulation of fungal gene expression. It is noted that the data obtained in this study should be interpreted with caution. The 5′UTR sequences used in this study were predicted from *A. oryzae* EST database (Machida 2002), and may not contain entire (full-length) 5′UTR sequence of the corresponding mRNAs if there was truncation of the full-length mRNA and/or incomplete synthesis by reverse transcriptase during cDNA preparation. Thus, translational characteristics of the full-length 5′UTRs may differ from those used in this study. Future research should focus on the nature of this phenomenon to elucidate the mechanisms of posttranscriptional regulation of gene expression in filamentous fungi.

Rapid regulation of Hsp expression must be crucial for cells to adapt to heat and other environmental stresses. Efficient translation mediated by the Hsp 5′UTRs may contribute to such rapid control of gene expression. The 5′UTRs of several *Arabidopsis* Hsp genes also contribute to efficient translation (Dansako et al. 2003). In the light of these results, it is possible that enhancement of translation mediated by the Hsp 5′UTRs might be a general phenomenon that is widely retained among eukaryotes, although the other various Hsp 5′UTRs should be tested.

Lastly, it is expected that these Hsp 5′UTRs, especially the *Hsp12* 5′UTR, could be used in combination with strong promoters to improve foreign gene expression in *A. oryzae*.

Conclusion

The aim of this study was to investigate the possibility of enhancing heterologous protein expression by improvement of posttranscription initiation events in filamentous fungi.

In Chapter I, the coding region of a gene to be expressed was focused, and the effect of codon optimization on the expression of a heterologous protein was investigated in *A. niger*, using potato α-glucan phosphorylase (GP) as a model protein. In the early attempts, expression of the wild-type GP gene (GP-wt) was inhibited at the mRNA level. Given the fact that the GP-wt was under control of the strong promoter in pNAN8142, the undetectable mRNA level observed was almost certainly not due to effects on transcription initiation but instead probably the result of post-transcription initiation effects, such as incorrect termination of transcription and/or mRNA stability, which are related to the characteristic features of the coding region. It was found that there were large differences between the most frequently used codons in GP-wt and *A. niger* genes and that there were many codons in the entire sequence that are extremely rare (fraction < 10%) in *A. niger*. Additionally, it was found that the A+T content was significantly higher in GP-wt (57%) than in *A. niger* genes (44%), so GP-wt contained many A+T-rich stretches that resemble a 'terminator' consensus (AATAAA). Therefore, it was suspected that the lack of detectable mRNA was due, in part, to the high content of rare codons and/or premature termination of the transcripts. To overcome these possible problems, an optimized synthetic gene (GP-syn) was designed. Use of *A. niger*-preferred codon usage and lower A+T content in a synthetic gene (GP-syn) resulted in a significant improvement in the level of the GP mRNA and a

dramatic increase in the quantity of GP protein produced such that it accounted for approximately 10% of the total soluble protein. These results suggest that redesigning the primary DNA sequence encoding a desired protein product can be an extremely effective method for improving heterologous protein production in filamentous fungi.

In Chapter II, I focused on the translation initiation step mediated by the 5′UTR and investigated whether alteration in the 5′UTR of an expression vector has a large influence on translation efficiency in *A. oryzae*. Using the GUS reporter gene assay, various constructs containing the same P-No8142 promoter with different 5′UTR sequences were analyzed. Alterations in the 5′UTR, including partial or nearly complete replacement of the 5′UTR, resulted in an increase in GUS activity of up to eight-fold, without affecting mRNA levels. These results clearly show that the enhancements in GUS expression were due to an increase in translation efficiency and that the 5′UTR plays an important role in determining translation efficiency in *A. oryzae*. Moreover, using the most effective 5′UTR construct, remarkable intracellular overproduction of GUS protein was achieved; and the GUS level reached more than 50% of the total soluble protein. This is the first experimental evidence that indicates the feasibility of improving recombinant protein yield by promoting translation initiation in filamentous fungi.

In Chapter III, it was investigated specifically whether the 5′UTRs of *A. oryzae* Hsp genes can increase the efficiency of expression of a foreign gene by enhancing translation. Using a series of 5′UTR::GUS (*uidA*) fusion constructs, the translation efficiency of chimeric mRNAs with different 5′UTRs was analyzed at different temperatures. It was found that the 5′UTR of a heat-shock protein gene, *Hsp12*, greatly enhanced the translation efficiency of the chimeric GUS mRNA at normal

temperature (30°C). A further observation was that temperature had differential effects on translation efficiency of different mRNAs containing different 5′UTRs. At high temperature (37°C), the translation efficiency of the mRNA containing the *Hsp12* 5′UTR was far superior to that of mRNAs containing nonheat-shock 5′UTRs, resulting in much more efficient expression of GUS protein (about 20-fold higher GUS activity compared to the control construct). This 5′UTR can be used in combination with various strong promoters to enhance the expression of foreign proteins in *A. oryzae*.

The processes involved in protein expression are highly complicated, and future research on each aspect of protein expression will lead to the elimination of limitations in heterologous protein production in filamentous fungi. As demonstrated in this study, the novel strategies for expressing heterologous proteins has been established with improved characteristics at the level of posttranscription initiation in *Aspergillus* species. These strategies should be applied extensively for heterologous protein expression in filamentous fungi, and will be able to contribute to biotechnological industry, such as the production of pharmaceutical drugs and enzymes. I am now trying to produce various useful enzymes using these strategies. In Fig. 1, I summarize the mentioned strategies by presenting the construction of an efficient expression vector. Furthermore, it will be expected that these results provide important information also for the study of basic mechanisms of gene expression in filamentous fungi.

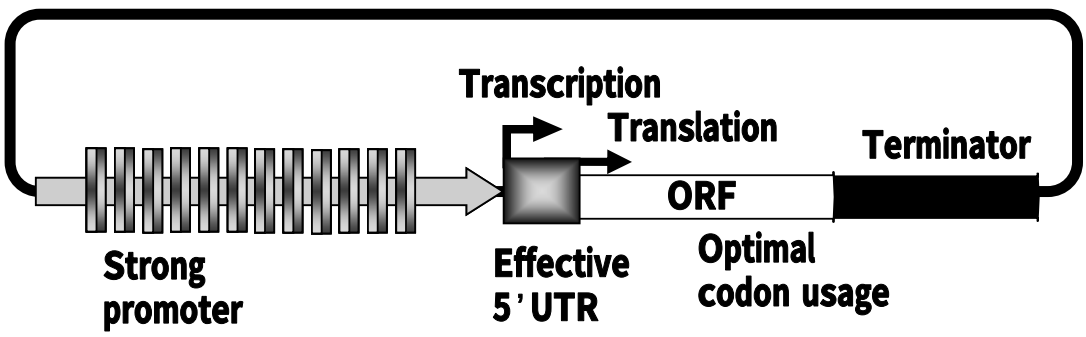

FIG. 1. Schematic representation of the construction of an efficient expression vector.

References

Archer DB, Peberdy JF (1997) The molecular biology of secreted enzyme production by fungi. Crit Rev Biotechnol 17: 273–306

Ballance DJ (1986) Sequences important for gene expression in filamentous fungi. Yeast 2: 229–236

Caponigro G, Muhlrad D, and Parker R (1993) A small segment of the *MATα1* transcript promotes mRNA decay in *Saccharomyces cerevisiae*: a stimulatory role for rare codons. Mol Cell Biol 13: 5141–5148

Cardoza RE, Gutierrez S, Ortega N, Colina A, Casqueiro J, and Martin JF (2003) Expression of a synthetic copy of the bovine chymosin gene in *Aspergillus awamori* from constitutive and pH-regulated promoters and secretion using two different pre-pro sequences. Biotechnol Bioeng 83: 249–259

Christensen T, Woeldike H, Boel E, Mortensen SB, Hjortshoej K, Thim L, and Hansen MT (1988) High-level expression of recombinant genes in *Aspergillus oryzae*. Bio/Technology 6: 1419–1422

Conesa A, Punt PJ, van Luijk N, and Hondel CAMJJ van den (2001) The secretion pathway in filamentous fungi: a biotechnological view. Fungal Genet Biol 33: 155–171

Contreras R, Carrez D, Kinghorn JR, Hondel CAMJJ van den, and Fiers W (1991) Efficient KEX2-like processing of a glucoamylase-interleukin-6 fusion protein by *Aspergillus nidulans* and secretion of mature interleukin-6. Bio/Technology 9: 378–381

Dansako T, Kato K, Satoh J, Sekine M, Yoshida K, and Shinmyo A (2003) 5′

untranslated region of the *HSP18.2* gene contributes to efficient translation in plant cells. J Biosci Bioeng 95: 52–58

Faus I, Patino C, Rio JL, del Moral C, Barroso HS, and Rubio V (1996) Expression of a synthetic gene encoding the sweet-tasting protein thaumatin in *Escherichia coli*. Biochem Biophys Res Commun 229: 121–127

Faus I, del Moral C, Adroer N, del Rio JL, Patino C, Sisniega H, Casas C, Blade J, and Rubio V (1998) Secretion of the sweet-tasting protein thaumatin by recombinant strains of *Aspergillus niger* var. *awamori*. Appl Microbiol Biotechnol 49: 393–398

Fujii K, Takata H, Yanase M, Terada Y, Ohdan K, Takaha T, Okada S, and Kuriki T (2003) Bioengineering and application of novel glucose polymers. Biocatal Biotransform 21: 167–172

Gomi K, Iimura Y, and Hara S (1987) Integrative transformation of *Aspergillus oryzae* with a plasmid containing the *Aspergillus nidulans argB* gene. Agric Biol Chem 51: 2549–2555

Gomi K, Akeno T, Minetoki T, Ozeki K, Kumagai C, Okazaki N, and Iimura Y (2000) Molecular cloning and characterization of a transcriptional activator gene, *amyR*, involved in the amylolytic gene expression in *Aspergillus oryzae*. Biosci Biotechnol Biochem 64: 816–827

Gouka RJ, Punt PJ, Hessing JG, and Hondel CAMJJ van den (1996) Analysis of heterologous protein production in defined recombinant *Aspergillus awamori* strains. Appl Environ Microbiol 62: 1951–1957

Gouka RJ, Punt PJ, and Hondel CAMJJ van den (1997) Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. Appl

Microbiol Biotechnol 47: 1–11

Hamilton R, Watanabe CK, and de Boer HA (1987) Compilation and comparison of the sequence context around the AUG start codons in *Saccharomyces cerevisiae* mRNAs. Nucleic Acids Res 15: 3581–3593

Hata Y, Tsuchiya K, Kitamoto K, Gomi K, Kumagai C, Tamura G, and Hara S (1990) Nucleotide sequence and expression of the glucoamylase-encoding gene (*glaA*) from *Aspergillus oryzae*. Gene 108: 145–150

Hata Y, Kitamoto K, Gomi K, Kumagai C, Tamura G, and Hara S (1991) The glucoamylase cDNA from *Aspergillus oryzae*: its cloning, nucleotide sequence, and expression in *Saccharomyces cerevisiae*. Agric Biol Chem 55: 941–949

Hata Y, Kitamoto K, Gomi K, Kumagai C, and Tamura G (1992) Functional elements of the promoter region of the *Aspergillus oryzae glaA* gene encoding glucoamylase. Curr Genet 22: 85–91

Hellen CU, Sarnow P (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev 15: 1593–1612

Hennigan AN, Jacobson A (1996) Functional mapping of the translation-dependent instability element of yeast *MATα1* mRNA. Mol Cell Biol 16: 3833–3843

Hentze MW (1991) Determinants and regulation of cytoplasmic mRNA stability in eukaryotic cells. Biochim Biophys Acta 1090: 281–292

Ishida H, Hata Y, Kawato A, Abe Y, and Kashiwagi Y (2004) Isolation of a novel promoter for efficient protein production in *Aspergillus oryzae*. Biosci Biotechnol Biochem 68: 1849–1857

Ishida H, Matsumura K, Hata Y, Kawato A, Suginami K, Abe Y, Imayasu S, and Ichishima E (2001) Establishment of a hyper-protein production system in

submerged *Aspergillus oryzae* culture under tyrosinase-encoding gene (*melO*) promoter control. Appl Microbiol Biotechnol 57: 131–137

Jeenes DJ, Marczinke B, MacKenzie DA, and Archer DB (1993) A truncated glucoamylase gene fusion for heterologous protein secretion from *Aspergillus niger*. FEMS Microbiol Lett 107: 267–271

Jefferson RA, Burgess SM, and Hirsh D (1986) β-Glucuronidase from Escherichia coli as a gene-fusion marker. Proc Natl Acad Sci USA 83: 8447–8451

Jefferson RA, Kavanagh TA, and Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901–3907

Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, and Cossart P (2002) An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. Cell 110: 551–561

Juge N, Svensson B, and Williamson G (1998) Secretion, purification, and characterization of barley α-amylase produced by heterologous gene expression in *Aspergillus niger*. Appl Microbiol Biotechnol 49: 385–392

Kane JF (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 6: 494–500

Kim YK, Jang SK (2002) Continuous heat shock enhances translational initiation directed by internal ribosomal entry site. Biochem Biophys Res Commun 297: 224–231

Kinnaird JH, Burns PA, and Fincham JR (1991) An apparent rare-codon effect on the rate of translation of a *Neurospora* gene. J Mol Biol 221: 733–736

Kitamura S (1996) Starch polymers, natural and synthetic, p. 7915–7922. *In* Salamone,

J. C. (ed.), Polymeric materials encyclopedia CRC Press, Boca Raton

Kochetov AV, Ischenko IV, Vorobiev DG, Kel AE, Babenko VN, Kisselev LL, and Kolchanov NA (1998) Eukaryotic mRNAs encoding abundant and scarce proteins are statistically dissimilar in many structural features. FEBS Lett 440: 351–355

Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44: 283–292

Kozak M (1987) An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 15: 8125–8148

Kozak M (1989) The scanning model for translation: an update. J Cell Biol 108: 229–241

Kozak M (1991a) Effects of long 5′ leader sequences on initiation by eukaryotic ribosomes in vitro. Gene Expr 1: 117–125

Kozak M (1991b) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem 266: 19867–19870

Krasevec N, Hondel CAMJJ van den, and Komel R (2000) Expression of human lymphotoxin α in *Aspergillus niger*. Pflug Arch Eur J Phy 440: Suppl R83–R85

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685

Lemm I, Ross J (2002) Regulation of c-*myc* mRNA decay by translational pausing in a coding region instability determinant. Mol Cell Biol 22: 3959–3969

Machida M, Chang YC, Manabe M, Yasukawa M, Kunihiro S, and Jigami Y (1996) Molecular cloning of a cDNA encoding enolase from the filamentous fungus, *Aspergillus oryzae*. Curr Genet 30: 423–431

Machida M (2002) Progress of *Aspergillus oryzae* genomics. Adv Appl Microbiol 51: 81–106

Minetoki T, Gomi K, Kitamoto K, Kumagai C, and Tamura G (1995) Nucleotide sequence and expression of α-glucosidase-encoding gene (*agdA*) from *Aspergillus oryzae*. Biosci Biotechnol Biochem 59: 1516–1521

Minetoki T, Nunokawa Y, Gomi K, Kitamoto K, Kumagai C, and Tamura G (1996) Deletion analysis of promoter elements of the *Aspergillus oryzae agdA* gene encoding α-glucosidase. Curr Genet 30: 432–438

Minetoki T, Kumagai C, Gomi K, Kitamoto K, and Takahashi K (1998) Improvement of promoter activity by the introduction of multiple copies of the conserved region III sequence, involved in the efficient expression of *Aspergillus oryzae* amylase-encoding genes. Appl Microbiol Biotechnol 50: 459–467

Minetoki T (2000) Development of the heterologous gene expression system for *Aspergillus* species. Kagaku to Seibutsu 38: 831–838 (in Japanese)

Miura Y, Kettoku M, Kato M, Kobayashi K, and Kondo K (1999) High level production of thermostable α-amylase from *Sulfolobus solfataricus* in high-cell density culture of the food yeast *Candida utilis*. J Mol Microbiol Biotechnol 1: 129–134

Moralejo FJ, Cardoza RE, Gutierrez S, Sisniega H, Faus I, and Martin JF (2000) Overexpression and lack of degradation of thaumatin in an aspergillopepsin A-defective mutant of *Aspergillus awamori* containing an insertion in the *pepA* gene. Appl Microbiol Biotechnol 54: 772–777

Mori H, Tanizawa K, and Fukui T (1991) Potato tuber type H phosphorylase isozyme. Molecular cloning, nucleotide sequence, and expression of a full-length cDNA in

Escherichia coli. J Biol Chem 266: 18446–18453

Morita MT, Tanaka Y, Kodama TS, Kyogoku Y, Yanagi H, and Yura T (1999) Translational induction of heat shock transcription factor σ^{32}: evidence for a built-in RNA thermosensor. Genes Dev 13: 655–665

Morris DR, Geballe AP (2000) Upstream open reading frames as regulators of mRNA translation. Mol Cell Biol 20: 8635–8642

Nakamura Y, Gojobori T, and Ikemura T (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res 28: 292

Nelson G, Kozlova-Zwinderman O, Collis AJ, Knight MR, Fincham JR, Stanger CP, Renwick A, Hessing JG, Punt PJ, Hondel CAMJJ van den, and Read ND (2004) Calcium measurement in living filamentous fungi expressing codon-optimized aequorin. Mol Microbiol 52: 1437–1450

Niemann C, Saenger W, Pfannemüller B, Eigner WD, and Huber A (1991) Phosphorolytic synthesis of low-molecular-weight amyloses with modified terminal proups. Comparison of potato phosphorylase and muscle phosphorylase b, p. 189–204. *In* Friedman, R. B. (ed.), Biotechnology of amylodextrin oligosaccharides. American Chemical Society, Washington, D.C.

Ozeki K, Kanda A, Hamachi M, and Nunokawa Y (1996) Construction of a promoter probe vector autonomously maintained in *Aspergillus* and characterization of promoter regions derived from *A. niger* and *A. oryzae* genomes. Biosci Biotechnol Biochem 60: 383–389

Romanos MA, Makoff AJ, Fairweather NF, Beesley KM, Slater DE, Rayment FB, Payne MM, and Clare JJ (1991) Expression of tetanus toxin fragment C in

yeast: gene synthesis is required to eliminate fortuitous polyadenylation sites in AT-rich DNA. Nucleic Acids Res., **19**, 1461–1467

Saheki S, Takeda A, and Shimazu T (1985) Assay of inorganic phosphate in the mild pH range, suitable for measurement of glycogen phosphorylase activity. Anal Biochem 148: 277–281

Scorer CA, Buckholz RG, Clare JJ, and Romanos MA (1993) The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris*. Gene 136: 111–119

Stemmer WPC, Crameri A, Ha KD, Brennan TM, and Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene 164: 49–53

Tada S, Iimura Y, Gomi K, Takahashi K, Hara S, and Yoshizawa K (1989) Cloning and nucleotide sequence of the genomic Taka-amylase A gene of *Aspergillus oryzae*. Agric Biol Chem 53: 593–599

Tada S, Gomi K, Kitamoto K, Takahashi K, Tamura G, and Hara S (1991) Construction of a fusion gene comprising the Taka-amylase A promoter and the *Escherichia coli* β-glucuronidase gene and analysis of its expression in *Aspergillus oryzae*. Mol Gen Genet 229: 301–306

Toda T, Sano M, Honda M, Rimoldi OJ, Yang Y, Yamamoto M, Takase K, Hirozumi K, Kitamoto K, Minetoki T, Gomi K, and Machida M (2001) Deletion analysis of the enolase gene (*enoA*) promoter from the filamentous fungus *Aspergillus oryzae*. Curr Genet 40: 260–267

Tsuboi H, Koda A, Toda T, Minetoki T, Hirotsune M, and Machida M (2005) Improvement of the *Aspergillus oryzae* enolase promoter (P-*enoA*) by the

introduction of *cis*-element repeats. Biosci Biotechnol Biochem 69: 206–208

Tsuchiya K, Tada S, Gomi K, Kitamoto K, Kumagai C, and Tamura G (1992) Deletion analysis of the Taka-amylase A gene promoter using a homologous transformation system in *Aspergillus oryzae*. Biosci Biotechnol Biochem 56: 1849–1853

Tsukagoshi N, Kobayashi T, and Kato M (2001) Regulation of the amylolytic and (hemi-)cellulolytic genes in aspergilli. J Gen Appl Microbiol 47: 1–19

Yun DF, Laz TM, Clements JM, and Sherman F (1996) mRNA sequences influencing translation and the selection of AUG initiator codons in the yeast *Saccharomyces cerevisiae*. Mol Microbiol 19: 1225–1239

Verdoes JC, Punt PJ, and Hondel CAMJJ van den (1995) Molecular genetic strain improvement for the overproduction of fungal proteins by Filamentous fungi. Appl Microbiol Biotechnol 43: 195–205

Vilela C, McCarthy JE (2003) Regulation of fungal gene expression via short open reading frames in the mRNA 5′ untranslated region. Mol Microbiol 49: 859–867

Ward M, Wilson LJ, Kodama KH, Rey MW, and Berka RM (1990) Improved production of chymosin in *Aspergillus* by expression as a glucoamylase-chymosin fusion. Bio/Technology 8: 435–440

Weinhäusel A, Nidetzky B, Kysela C, and Kulbe KD (1995) Application of *Escherichia coli* maltodextrin phosphorylase for the continuous production of glucose 1-phosphate. Enzyme Microb Technol 17: 140–146

Weinhäusel A, Nidetzky B, Rohrbach M, Blauensteiner B, and Kulbe KD (1994) A new maltodextrin-phosphorylase from *Corynebacterium callunae* for the production of glucose 1- phosphate. Appl Microbiol Biotechnol 41: 510–516

Wolin SL, Walter P (1988) Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J 7: 3559–3569

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415

Acknowledgments

I am very grateful to Professor Dr. Yoshiki Tani, Nara Institute of Science and Technology (NAIST), for his kind guidance and encouragement while carrying out this study.

I am deeply grateful to Dr. Ko Kato in NAIST for his helpful discussions. Without him, this thesis could not have been completed.

I also wish to express thanks to Professor Dr. Katsuya Gomi in Tohoku University for his valuable advices. His extensive discussions around my work have been very helpful for this study.

I also thank Drs. Takeshi Takaha and Hiroki Takata, Ezaki Glico Co., Ltd., for their valuable discussions and excellent technical support.

I express special thanks to Mr. Bunjiro Osabe, Honorary Chairman of Ozeki Co., Ltd., Mr. Yasuo Hashimoto, President of Ozeki Co., Ltd., and Mr. Tadayuki Yamamoto, Managing Director of Ozeki Co., Ltd., for their encouragement in carrying out this work.

I am most grateful to Dr. Masato Hirotsune, Research Manager of Ozeki Co., Ltd., for his management of this work.

My warm thanks are due to Professor Dr. Kenji Ozeki in Kanazawa Institute of Technology, Dr. Toshitaka Minetoki of Ozeki Co., Ltd., and Mr. Takayuki Bogaki of Ozeki Co., Ltd., who directed me in fungal biotechnological research, including this study. Their kind support and guidance have been of great value in this study.

I am indebted to Mrs. Akiko Nishi and Mr. Hirokazu Tsuboi for their helpful discussions, and to all members of the Gen. Res. Lab. of Ozeki Co., Ltd. for their

encouragement and collaboration during this study.

Special thanks are due to Mr. Jesse Olsen and my little sisters Aya and Miho, for their kind advice about writing English reports.

I can never thank enough my parents. Their wholehearted affection and dedicated support have made me what I am today. I would like to express my sincere gratitude to my parents-in-law for their tenderness and great understanding.

Finally, my deep appreciation should go to my wife, Yoko, and my daughters, Sana and Yuna, for their love and heartful support.

Publications

Original articles

Koda A, Minetoki T, Ozeki K, and Hirotsune M (2004) Translation efficiency mediated by the 5′ untranslated region greatly affects protein production in *Aspergillus oryzae*. Appl Microbiol Biotechnol 66: 291–296

Koda A, Bogaki T, Minetoki T, and Hirotsune M (2005) 5′ Untranslated region of the *Hsp12* gene contributes to efficient translation in *Aspergillus oryzae*. Appl Microbiol Biotechnol, in press

Koda A, Bogaki T, Minetoki T, and Hirotsune M (2005) High expression of a synthetic gene encoding potato α-glucan phosphorylase in *Aspergillus niger*. J Biosci Bioeng 100, in press

Related articles

Tsuboi H, **Koda A**, Toda T, Minetoki T, Hirotsune M, and Machida M (2005) Improvement of the *Aspergillus oryzae* enolase promoter (P-*enoA*) by the introduction of *cis*-element repeats. Biosci Biotechnol Biochem 69: 206–208

Koda A, Minetoki T (2005) Development of high expression system for foreign proteins in *Aspergillus* species. Seibutsu-kogaku Kaishi 83: 295–297 (in Japanese)

Koda A, Tokuoka M, and Gomi K (2005) Efficient recombinant protein production improved at the post-transcriptional level in filamentous fungi. BIOSCIENCE & INDUSTRY 63: 543–546 (in Japanese)