Gene 496 (2012) 118–127

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Analysis of the 5′ untranslated region (5′UTR) of the alcohol oxidase 1 (*AOX1*) gene in recombinant protein expression in *Pichia pastoris*

Chris A. Staley [a], Amy Huang [a], Maria Nattestad [a], Kristin T. Oshiro [a], Laura E. Ray [a], Tejas Mulye [a], Zhiguo Harry Li [c], Thu Le [a], Justin J. Stephens [a], Seth R. Gomez [a], Allison D. Moy [a], Jackson C. Nguyen [a], Andreas H. Franz [b], Joan Lin-Cereghino [a], Geoff P. Lin-Cereghino [a,*]

[a] *Department of Biological Sciences, University of the Pacific, Stockton, CA 95211, USA*
[b] *Department of Chemistry, University of the Pacific, Stockton, CA 95211, USA*
[c] *Novo-Nordisk Pharmaceuticals Science and Technology, Beijing, China*

ARTICLE INFO

Article history:
Accepted 5 January 2012
Available online 25 January 2011

Keywords:
Pichia pastoris
Recombinant protein expression
Translational regulation
Untranslated regions
mRNA

ABSTRACT

Pichia pastoris is a methylotrophic yeast that has been genetically engineered to express over one thousand heterologous proteins valued for industrial, pharmaceutical and basic research purposes. In most cases, the 5′ untranslated region (UTR) of the alcohol oxidase 1 (*AOX1*) gene is fused to the coding sequence of the recombinant gene for protein expression in this yeast. Because the effect of the *AOX1* 5′UTR on protein expression is not known, site-directed mutagenesis was performed in order to decrease or increase the length of this region. Both of these types of changes were shown to affect translational efficiency, not transcript stability. While increasing the length of the 5′UTR clearly decreased expression of a β-galactosidase reporter in a proportional manner, a deletion analysis demonstrated that the *AOX1* 5′UTR contains a complex mixture of both positive and negative *cis*-acting elements, suggesting that the construction of a synthetic 5′UTR optimized for a higher level of expression may be challenging.

1. Introduction

Recombinant proteins serve many uses in our society ranging from therapeutic agents to industrial enzymes. Since the early 1990s, the methylotrophic yeast, *Pichia pastoris*, has been an excellent system for heterologous protein production, with over one thousand recombinant proteins having been produced in this yeast (Cereghino and Cregg, 2000; Lin-Cereghino et al., 2007). One key to its success is the alcohol oxidase 1 promoter (*AOX1*), which is used to drive the expression of the recombinant protein (Ellis et al., 1985; Koutz et al., 1989). The *AOX1* promoter is tightly repressed in cells grown in glucose medium but induced over 1000-fold in methanol medium (Lin-Cereghino et al., 2006). For expression of a specific recombinant protein, usually the coding sequence of the foreign gene is first inserted into a polylinker located between the *AOX1* promoter with its 5′ untranslated region (5′UTR) and the *AOX1* 3′UTR of the expression cassette of a vector (Lin Cereghino et al., 2001). The construct is then transformed into competent *P. pastoris* cells, where it integrates usually by homologous recombination, and transformants are induced on methanol medium to trigger transcription of the hybrid gene. The resulting mRNA, containing the *AOX1* 5′UTR, the foreign coding region and the *AOX1* 3′UTR, is then bound at the cap by an initiation complex, which consists of the 40S ribosomal subunit and eukaryotic initiation factors (eIFs) (Merrick, 2010; Pesole et al., 2001; Sonenberg and Hinnebusch, 2009). The 5′UTR is consequently scanned by the translation initiation complex in a 5′ to 3′ direction until the first AUG is encountered (Merrick, 2010). At this point, the 60S ribosomal subunit normally joins the complex, an 80S complex is formed, and translation commences.

In order to increase the yield of the recombinant protein, most investigators have focused on altering the *AOX1* promoter region or alternatively using a different promoter, such as glyceraldehyde-3-phosphate dehydrogenase (*GAP*) (Hartner et al., 2008; Qin et al., 2011). Little is known about the contribution of the *AOX1* 5′UTR to the expression of recombinant proteins. The 5′UTR is the region of mRNA encoded by DNA between the transcription start site and the first ATG in the gene (Lawless et al., 2009). In many eukaryotic genes, the 5′UTR has a regulatory function by affecting either the stability or the translational efficiency of the mRNA (Mokdad-Gargouri et al., 2001; Park et al., 2007; Sonenberg and Hinnebusch, 2009; Warringer et al., 2010; Wicksteed et al., 2007). The 5′UTR can exert an effect on translation efficiency either by adopting a secondary

Abbreviations: AOX1, alcohol oxidase 1; UTR, untranslated region; nt, nucleotides; eIF, eukaryotic initiation factor.

* Corresponding author at: Department of Biological Sciences, University of the Pacific, 3601 Pacific Avenue, Stockton, California 95211, USA. Tel.: +1 209 946 2988; fax: +1 209 946 3022.
E-mail address: glincere@pacific.edu (G.P. Lin-Cereghino).
URL: http://web.pacific.edu/x12191.xml (G.P. Lin-Cereghino).

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2012.01.006

structure which can influence the rate of ribosome movement or by binding *trans*-acting factors that affect the function of the translation machinery (Chatterjee and Pal, 2009; Mignone et al., 2002; Mittelmeier and Dieckmann, 1995). Some eukaryotic 5′UTRs even contain sequences called internal ribosome entry sites (IRES), which allow the translation initiation complex with the 40S to skip over the cap and bind nearer to the first AUG in a cap-independent manner (Baird et al., 2006; Mokrejs et al., 2009). However, at the present time, IRES are defined by functional criteria only. Although some contain a Y-shaped stem-loop just upstream of the AUG initiation codon, there is no identified consensus sequence (Lopez-Lastra et al., 2005; Mignone et al., 2002). 5′UTR structure, thus, has the potential to influence protein expression, but much of this potential has yet to be understood.

Even though the *AOX1* 5′UTR is found in most *P. pastoris* expression vectors, to date no systematic studies have been done to determine if the *AOX1* 5′UTR plays a significant role in controlling protein expression. Because the 114 nucleotide-long 5′UTR is unusually long for yeast 5′UTRs (most are shorter than 50 nucleotides) (Lawless et al., 2009), this sequence drew our interest. Our first goal was to determine if the *AOX1* 5′UTR contains sequence elements that influence protein production by acting on transcript stability or translatability. We envisioned using this data to engineer an artificial 5′UTR to provide enhanced protein production. For instance, mutagenesis of a 5′ UTR in *Aspergillus oryzae* increased expression of a GUS reporter up to eight-fold (Koda et al., 2004). Our second goal was to comprehend how increases in the length of 5′UTR affect expression of heterologous proteins. Often when investigators are inserting their coding sequences into expression vectors, they may add a large amount of foreign sequence upstream of the ATG or use a restriction site located in the 3′ region of a polylinker (Bagga, 2008; Grillo et al., 2010). Both of these actions lengthen the 5′UTR. Although it was reported that an increase of 36 nucleotides in the *AOX1* 5′UTR reduced the expression of human serum albumin by about 50-fold (Sreekrishna, 1993), no systematic analysis has been performed to correlate the effect of incremental 5′UTR extension with protein expression levels.

To meet these two goals, site-directed mutagenesis was carried out on the *AOX1* 5′UTR region fused to a β-galactosidase reporter. β-galactosidase assays were performed to measure the effect of these mutations on protein expression while northern analysis was utilized to explore the possibility that these mutations affected transcript stability. Although increasing the length of the 5′UTR had a clear effect on protein expression, our results also suggest that the *AOX1* 5′UTR contains an intricate network of both positive and negative *cis*-acting elements that affect the translational efficiency of the messenger mRNA.

2. Materials and methods

2.1. Strains, media and reagents

P. pastoris strain yJC100 (wild type) is a derivative of the original wild-type *P. pastoris* strain NRRL Y11430 (North Regional Research Laboratories, US Department of Agriculture, Peoria, IL) and has been described previously (Lin-Cereghino et al., 2006). yJC100 cells were cultured in either YPD medium (1% yeast extract, 2% peptone, and 2% glucose), YND (minimal medium with 1% dextrose) or YNM (minimal medium with 0.5% methanol) (Cregg and Madden, 1988; Cregg et al., 1985).

Recombinant DNA manipulations were carried out in the *Escherichia coli* strain TOP10 (Invitrogen Corp., Carlsbad, CA). TOP10 cells were cultured in LB medium (0.5% yeast extract, 1% glucose, and 0.5% NaCl) at 37 °C. Zeocin was added to LB medium at the final concentration of 25 μg/mL Zeocin for plasmid selection. Recombinant DNA methods, including bacterial transformation, were performed essentially as described (Sambrook et al., 1989). Plasmid DNA was purified from *E. coli* cultures using a QIAprep Spin Miniprep Kit (Qiagen, Chatsworth, CA). PCR products were purified with the QIAprep PCR Cleanup Kit (Qiagen, Chatsworth, CA) prior to restriction digestion. Restriction enzymes were purchased from MBI Fermentas (Hanover, MD). DNA digested with restriction enzymes was resolved in TBE agarose gels. DNA fragments were purified from agarose gels by using the Geneclean II kit (Qbiogene, Carlsbad, CA). Chromosomal DNA from *P. pastoris* transformants was prepared using the OmniPrep™ kit from GenoTechnology, Inc. (St. Louis, MO). Oligonucleotides were synthesized by Sigma Genosys (Plano, TX). All mutated sites and ligation junctions in newly synthesized vectors were confirmed by DNA sequencing (Geneway Research, Hayward, CA).

2.2. Construction of plasmids

The coding sequence the *lacZ* gene was amplified from pGC181 (Lin-Cereghino et al., 2006) with the primers chriss5bgal2 CAAG AATTCATGCCAGGGGATCCCGTCGTT and 0103R CAAGCGGCCGCTTT TTGACACCAGACCAA. The resulting PCR fragment was restricted with the enzymes *Eco*RI and *Not*I and then ligated into the corresponding sites of pPICZB (Invitrogen Corp, Carlsbad, CA) to yield pCS1, the parent deletion plasmid. All deletion plasmids were created by site-directed mutagenesis using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). The parent addition plasmid, pLR1, was created from pCS1. The plasmid pLR1 and pCS1 are identical except pLR1 has a *Spe*I restriction site inserted in the 5′UTR for use in vector construction. Plasmid pLR1 was created using the QuikChange II XL Site-Directed Mutagenesis Kit with primers Laura 1a GAGAAGATCAAAAAACAACTAGTTATTCGAAACGAGG and Laura 1b CCTCGTTTTCGAATAACTAGTTGTTTTTTGATCTTCTC. All addition plasmids were produced from pLR1 using the double oligonucleotide reannealing technique (DORT) (see below). pAH201, in which the *AOX1* 5′UTR was replaced by the *GAP* 5′UTR, was constructed by removing the 140 bp *HindIII-EcoRI* fragment of pCS1 and inserting the following two complementary oligonucleotides Amyswapgaptop CCTTTTTTTTTATCATCACTAGTAGCTTACTTTCATAATGCG and Amyswapgapbottom CGCAATTATGAAAGTAAGCTACTAGTGATGATAAAA AAAAAGG with the DORT.

2.3. Site-directed mutagenesis

Deletions of the *AOX1* 5′UTR were generated using QuikChange II XL Site-Directed Mutagenesis Kit from Stratagene (La Jolla, CA). Primers were designed to anneal to the parent plasmid, pCS1, and "loop" out an area of the 5′UTR creating a deletion. The primers used in this work, which were HPLC purified, are listed in Table 1. Mutagenesis was performed according to manufacturer's directions. Briefly, the primers were diluted to a concentration of 25 ng/μL and the following reaction mixture was set up: 5 μL of 10× reaction buffer, 1 μL of pCS1, 5 μL of each primer, 1 μL of 10 mM dNTP, 3 μL of QuikSolution, 1 μL of *Pfu Turbo* DNA polymerase, and sterile water to a final volume of 50 μL. The cycling parameters were a) one cycle of 95 °C for 5 min, b) eighteen cycles of 95 °C for 30 s, 55 °C for 60 s, 68 °C for 7 min and c) one cycle of 68 °C for 10 min. After temperature cycling, 1 μL *Dpn*I was added directly to the reaction mix and incubated at 37 °C for 1 h to destroy the methylated, parental DNA. The final step of the mutagenesis was the transformation of the mutated plasmid reaction into One Shot TOP 10 Ultra competent *E. coli* cells (Invitrogen, Carlsbad, CA) and selection on LB plates supplemented with Zeocin® (25 μg/mL).

2.4. Double oligo reannealing technique (DORT)

All addition plasmids were produced using the DORT to insert polymers of adenine (A) nucleotides into the *AOX1* 5′UTR. The pairs of complementary oligonucleotides (Table 2) were designed in a manner that would create *Spe*I and *Eco*RI "sticky ends" when