Construction of Integration Vectors for Secreted Expression of Bacterial Phytase in Yarrowia Lipolytica

Daria S. Troshagina1, Anna E. Smirnova2, Catherine Madzak3, Aliya D. Suleimanova4

1, 2 Kazan Federal University, Institute of Fundamental Medicine and Biology
3 UMR GMPA, INRA, AgroParisTech, Université Paris-Saclay (Thiverval-Grignon, France)
4 Kazan Federal University, Institute of Fundamental Medicine and Biology

Email: aliya.kzn@gmail.com, Contact: +79274342839

Received: 21st October 2017 Accepted: 16th November 2017, Published: 31st December 2017

Abstract
Yeast Yarrowia lipolytica is a handy tool for efficient production of heterologous proteins. In this work we report the construction of integration vectors for secretion of bacterial phytase by these yeasts. Pantoea sp. 3.5.1 histidine acid phytase encoding gene sequence was codon-optimized and chemically synthetized. Optimized and native phytase gene sequences were cloned into integrative vector pRNA1296, containing signal sequence of XPR2 gene. Vectors were multiplied in E. coli DH5α, isolated and linearized for successful integration into Y.lipolytica genome by homologous recombination in pBR-region. Y.lipolytica strains with integrated bacterial native and optimized phytase genes were obtained.

Keywords: Phytase, Pantoea, Heterologous Expression, Yarrowia

Introduction
Heterologous proteins are of considerable interest for basic research, as well as for various biomedical and industrial applications. At present, various expression systems are used for the production of recombinant proteins, which can be divided into two groups - prokaryotic (bacterial) and eukaryotic, based on the usage of yeast cells, insects and mammals. Each of these systems has its advantages and disadvantages, which are based on the differences in the regulation of organism metabolism, transcription and translation mechanisms, as well as in the post-translational modifications and protein folding [1]. Therefore, the choice of an expression system depends on the properties of the target protein, as well as on the ability of the host cell to produce the protein of the desired quality with the observation of the minimum set of requirements.

The system based on E. coli is one of the prokaryotic expression systems, which has been used widely for several decades [2]. However, even with a high level of recombinant protein expression, these microorganisms can not be used as the producers for food, medical or agricultural industry because they are opportunistic pathogens and may produce various toxins along with proteins [3]. Besides, the usage of such expression systems is limited by the absence of post-translational modifications typical for eukaryotic proteins and by incorrect folding [4]. However, despite of all shortcomings, bacterial expression systems make an integral part of the initial stages of work with recombinant proteins, which is primarily related to the ease of working with bacteria, the development of genetic engineering and biotechnology methods for them, as well as the relative cost-effectiveness of the performed works.

The expression systems based on insect and mammalian cells carry out post-translational modifications of proteins, but unlike bacterial ones, are complex and high-cost systems, since they require the presence of expensive nutrient media and equipment [5].

The most optimal and compromise solution for the synthesis of recombinant proteins are yeast expression systems. First of all, yeasts are single-celled eukaryotic organisms that are well studied and variety of effective methods of genetic engineering has been developed for them. Secondly, they have the ability to support most of the posttranslational modifications of eukaryotic cells and to efficiently secrete recombinant proteins into the environment. Thirdly, yeast expression systems make it possible to obtain stable recombinant strains with a target gene integrated into the producer’s genome. High expression level is achieved in the presence of inexpensive inducers and on cheap nutrient media [4, 5].

The dimorphic yeast Yarrowia lipolytica has been used for several years as effective expression system [6]. They are characterized by a high level of heterologous protein production, a lesser degree of glycosylation of products and are unpretentious in cultivation. The purpose of this work was the construction of yeast integral vectors and creation of recombinant strains of Yarrowia lipolytica for the secretion of bacterial phytase from Pantoea sp. 3.5.1.

Materials and Methods
The cultivation of E. coli strains was performed on LB medium [Sambrook et al., 1989] (%): tryptone-1.0; yeast extract - 0.5; NaCl - 0.5; pH 8.5. LB (LA) agar medium contains an additional 2% of agar. During the cultivation of the recombinant E. coli DH5α strains with the plasmids pET28a/agP, kanamycin was added to the medium, with pUC57/agP-opt plasmids and pRNA1296-derived vectors, ampicillin was added to the medium.
The cultivation of *Y. lipolytica* yeasts was carried out on YPD medium (g/l): yeast extract - 10; peptone - 10; glucose - 10. YPD agar medium comprised 1.5% of agar.

The cultivation of yeast transformants was carried out on YNB-N₅₀₀₀ (g/l) medium: nitrogen base without amino acids - 6.7; glucose - 10; agar - 16.

The cloning of the optimized (agp-opt) and native (agpP) gene into the integrative yeast vector pINA1296 was performed. For this, the restriction of pINA1296 vector was performed by the restriction enzymes Kpn I and Sfi I (New England Biolabs). After restriction, the vector was treated with alkaline phosphatase (CIAP, 20 U/μl) (Thermo Fischer Scientific) to prevent self-ligation. After this, the restricted were purified by isolation from the gel using Gel extraction kit (Thermo Fisher Scientific).

The optimized agpP-opt phytase gene was amplified from the plasmid pUC57/agpP-opt using the primer pair Yar003-F (5’gtcggaacggccgttctggccgctgatggcgatatg-3’) and M13R (-48) (5’-agccgataaattcattacacagga-3’). The native agpP phytase gene was amplified from the plasmid pET28a/agpP using the primer pair Yar003-F and Yar001-R (5’-atatggtaccgcagttctgggtttgtttttg-3’).

The amplification reaction was carried out using “MJ mini” thermocycler (BioRad). 50 μl reaction mixture contained 50-200 ng of DNA, 10 μl of 5X Phusion HF Buffer, 1 μl of 10 mM deoxyribonucleoside triphosphate mixture (dNTP), 2.5 μl of each primer, 1.0 units of Phusion HF DNA polymerase (2 U/μl) (New England Biolabs). PCR program: primary denaturation of DNA at 98 °C for 30 seconds; then 10 seconds at 98 °C, the annealing of primers for 25 seconds at the temperature corresponding to each pair of primers, the elongation of 50 seconds at 72 °C, 35 cycles in total; Final synthesis at 72 °C for 10 min, the storage at 4 °C. PCR products were restricted at the Sfi I and Kpn I sites and purified by the PCR-product purification kit (Thermo Fisher Scientific).

The vector pINA1296 and the optimized agpP-opt gene or the native agpP gene were ligated using T4 DNA ligase (5 U/μl) (Thermo Fisher Scientific) followed by the transformation of the ligation mixture into *E. coli* DH5α cells. The correct embedding of the target gene was checked by sequencing using the primers to the yeast vector sequence: 6560F (5’-gatccgcatgcaacgcc-3’) and 6904R (5’-aaccggtaattctggagc-3’).

The verification of transformants was carried out by colony PCR. For this, one colony was taken and placed into sterile eppendorf tube containing 15 μl of distilled water, and was mixed. 5 μl of the resulting solution was used for the PCR reaction.

The yeast transformants were tested as follows: one colony was placed in a sterile eppendorf tube with 75 μl of sterile TE buffer (10 mM of tris (hydroxymethyl)amino methane (Tris), 1 mM of EDTA acid, pH 8.0). The resulting solution in the eppendorf tubes was irradiated in a microwave oven at maximum power for 3.5, 2, 1.5, 1 and 0.5 min with the breaks for stirring. Then the eppendorf tubes with the mixture of cells were frozen in the refrigerator at -80 °C for 10 minutes and then were centrifuged for 15 minutes at 2500 rpm. 5 μl of the resulting solution was taken for PCR reaction.

The transformation of *E. coli* DH5α cells with plasmid DNA was performed using the CaCl₂ method [7]. The transformation of *Yarrowia lipolytica* yeast cells with linear DNA was performed according to the modified method by Le Dall [8].

**Results and Discussion**

**Optimization of the Phytase Gene from *Pantoea sp.* 3.5.1**

Phytases (EC 3.1.3.8) are the class of phosphatases that catalyze the sequential hydrolysis of myo-inositol-(1,2,3,4,5,6)-hexakisphosphate or phytic acid to less phosphorylated myo-inositol and inorganic phosphate derivatives (Fig. 1). Phytases are widely used as feed additives for monogastric animals, they can be used in agriculture to improve the bioavailability of organic soil phosphorus and to obtain specific isomers of myo-inositol phosphates, which are therapeutic compounds. In this regard, the question of scaling up the processes of these enzymes production is an acute one, which can be solved by the usage of effective expression systems for their production.

---

Fig. 1. Schematic Mode of Action of Microbial Phytase on Phytic Acid Salts (Phytates).
In this paper, the sequence of the gene for histidine acid phytase *Pantoea* sp. 3.5.1 - *agpP* (AN KJ783401.1) was used. It was previously established that *Pantoea* sp. 3.5.1 phytase hydrolyses phytate with the production of D/L-myo-inositol-1,2,4,5,6-pentakisphosphate. The knowledge about the absolute stereocchemical specificity of *Pantoea* sp. 3.5.1 phytase allows using this enzyme to obtain desired isomers of myo-inositol phosphates [9].

The sequence of the *Pantoea* sp. 3.5.1 phytase gene was optimized by eliminating its own signal peptide sequence and adding a C-terminal His-tag for the detection of the product by immunoblotting and subsequent purification of the enzyme from the yeast culture liquid by affinity chromatography.

Codon-optimization of the target gene sequence is widely used for the correct expression of the heterologous protein in yeast. During the codon-optimization process of the nucleotide sequence of *Pantoea* sp. 3.5.1 phytase gene, the following parameters were taken into account: the frequency of different codon occurrence in yeast, GC-pair content, CpG-island content, secondary mRNA structure, hidden splicing sites, premature PolyA sites, internal chi-sites and ribosome binding sites, ARE, direct and reverse repeats in DNA, the repeats with bivalent symmetry (dyad repeats), etc. (http://www.genscript.com/codon-opt.html). The codon-optimized genes were synthesized and cloned into pUC57 vector by Genscript (USA).

In order to estimate the coding potential of considered DNA sequence, the frequency of optimal codons (FOP) is used. FOP is the fraction of the optimal codons among all the codons of the gene. Figure 7 shows the frequency of the optimal codons of the *Pantoea* sp. 3.5.1 phytase gene before and after the optimization for expression in *Yarrowia lipolytica* (Fig. 2). The increase in optimal codon frequency after the optimization of the gene is shown.

In order to clone the optimized and native phytase genes from *Pantoea* sp. 3.5.1 we used the yeast vector pINA1296 (Fig.3) [11]. pINA1296 is the expression vector (7205 bp) that contains a strong *hp4d* hybrid promoter with four tandem activating repeats of UAS1B (upstream activating sequences), representing the fragment of XPR2 gene promoter, and a minimal fragment of *LEU2* gene promoter (Figure 5). Vector contains the signal peptide sequence of XPR2 gene, which allows the secretion of the heterologous protein and active leucine biosynthesis gene *LEU2* as a selective marker, which is absent in the yeast genome and allows cells to grow on a medium without leucine. Vector is an integrative one and is inserted into the *Y.lipolytica* genome by homologous recombination at pBR site.
Fig. 3. The Map of pINA1296 Secretion Vector (7205 bp), which contains the Hybrid Promoter (hp4d) and a Secretion Signal (XPR2 pre region), the Multiple Cloning Site, the pXPR2 Transcription Terminator, Leucine Selection Marker Gene (LEU2).

In order to obtain pINA1296-agpP-opt and pINA1296-agpP constructs the restriction of pINA1296 vector was performed by KpnI and SfiI (Fig. 4a). Optimized and native phytase gene of Pantoea sp. 3.5.1 were amplified from plasmids pUC57/agpP-opt and pET28a/agpP using the primers Yar003-F and M13R (for the optimized phytase gene agpP-opt) (Fig. 4b) and Yar003-F and Yar001-R (for the native phytase gene agpP) (Fig. 4c), respectively. PCR products about 1690 bp were isolated from the gel and restricted by Kpn I and SfiI restriction enzymes.

Further, the ligation of pINA1296 vector was performed with the optimized and native Pantoea sp. 3.5.1 phytase genes. The ligase mixture was transformed into E. coli DH5α cells. The transformants were selected on the medium with ampicillin (100 μg/ml). In order to test the clones, PCR was performed from six colonies of each variant using the combination of primers to the yeast vector - 6560F and 6904R (Fig. 5). The PCR assay confirmed the presence of the phytase gene in all selected transformants.
The results of genotyping were confirmed by restriction analysis of the plasmids isolated from transformants and by sequencing. The following transformants were selected for further work: pINA1296-aggP-opt3 with the optimized phytase gene and pINA1296-aggP5 - with the native phytase gene.

Recombinant Strains of Yarrowia Lipolytica carrying the Phytase Gene of Pantoea sp. 3.5.1

The dimorphic yeasts Yarrowia lipolytica are recognized as an efficient system for the expression of heterologous proteins, which is characterized by a high yield of heterologous protein followed by the application for large-scale production. These yeasts are currently used for the production of petroleum products, the bioremediation of contaminated soils and waters, the bioconversion of whole cells and the modernization of industrial wastes [12].

The vectors pINA1296-aggP-opt3 and pINA1296-aggP5 obtained by us from E. coli were isolated and the restriction of these vectors was carried out within Spe I genomic integration site, which is located in the pBR region, to obtain a linearized structure.

The Y. lipolytica strain Po1g was used for transformation. In the genome of this strain the leucine biosynthesis gene (leu2), the alkaline extracellular protease gene (spr2) and the acid extracellular protease (asp) gene were deleted. In order to integrate the expression vectors based on pBR, the pBR322 docking platform was integrated into the yeast genome to the ura3 locus [11].

The transformation of the Yarrowia lipolytica strain was carried out with linearized genetic constructs pINA1296-aggP-opt3 and pINA1296-aggP5. The transformants were selected on the medium without leucine. On the second day of the incubation the results were evaluated.

In order to analyze the transformants, 3 colonies were selected from each transformation and genotyping was performed. The electrophoresis of PCR products showed that all transformants contained the genes of bacterial phytase integrated into the yeast genome (Fig.6).

Conclusions

Optimization of agpP phytase gene sequence was performed - 59.79% of the nucleotide residues were replaced, the codon adaptation index (CAI) and the frequency of the optimal codons were increased in the gene sequence.

Genetic constructs containing optimized (aggP-opt) and native phytase gene of Pantoea sp. 3.5.1 (aggP) under the control of the strong hybrid promoter hp4d and the signal peptide of the alkaline extracellular protease XPR2 gene were obtained.

The transformation of Yarrowia lipolytica was performed and 8 recombinant strains of Y. lipolytica with the bacterial phytase gene integrated in the genome were obtained.
Summary
Thus, recombinant strains of *Yarrowia lipolytica* containing the gene of the bacterial histidine acid phytase from *Pantoea* sp. 3.5.1 in their genome were obtained. Further evaluation of bacterial phytase expression by yeast, the isolation and the purification of the enzyme, the study of its properties may be an important stage for the creation of a competitive biotechnological production of phytases effective for the preparation of individual isomers of myo-inositol phosphates for pharmacology.

Acknowledgements
This work was supported by the Russian Foundation for Basic Research (project no. 16-34-60191). The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

References