Cloning, Recombinant Expression, and Characterization of the Uricase Gene from *Pichia guilliermondii* (Postprint)

Authors: Rao Jingjing, Jing Yixian, Zou Mingyue, Hu Xiaolei, Liao Fei, Yang Xiaolan

Date: 2017-11-13T00:00:00+00:00

Abstract

Objective: To clone the protein-coding gene of uricase from a Meyerozyma guilliermondii strain and confirm it through recombinant expression and characterization. Methods: The rRNA sequence of yeast strain C.G.M.C.C 2.1008 was determined to identify its species, tandem mass spectrometry was employed to analyze the peptide sequences of this native uricase for searching homologous proteins, transcriptome analysis was conducted to verify its inducible expression property, primers were designed based on the coding sequence of uricase (Uniprot id: A5DFP1) inferred from the genome information of Meyerozyma guilliermondii ATCC 6260, the uricase gene was amplified by PCR from cDNA of strain C.G.M.C.C 2.1008, and after sequencing, it was cloned into directional expression vectors pDE1 and pDE2 to construct recombinant expression plasmids pDE1-MGU and pDE2-MGU carrying a 6His tag, while simultaneously constructing an expression plasmid R-MGU without the 6His tag. This fungal uricase was induced for expression in Escherichia coli BL21(DE3), the peptide molecular weight was determined by SDS-PAGE and MALDI-TOF-MS, kinetic parameters and inhibitor sensitivity properties were measured using uric acid as substrate, and compared with the native uricase. Results: The rRNA sequence indicated that this strain was Meyerozyma guilliermondii, tandem mass spectrometry analysis of tryptic peptides showed that this native fungal uricase was highly similar to uricase A5DFP1, and transcriptome sequencing supported the efficient inducible expression of this uricase gene. The coding sequence was rapidly obtained from cDNA by PCR using the designed primers, and sequencing with T-vector revealed that it differed from the A5DFP1 coding sequence only at the 435th base, but the amino acid remained identical. SDS-PAGE revealed that the recombinant R-MGU peptide was approximately 35 kDa; MALDI-TOF-MS showed its peptide to be approximately 17.43 kDa and consistent with the native enzyme, but the molecular weight calculated from the

amino acid sequence was only 33.98 kDa, suggesting that it might have chemical modifications. The purified recombinant uricase with 6His tag exhibited a specific activity close to 6.0 U/mg; the Michaelis constant, inhibition constant representing inhibitors, and molecular weight of R-MGU showed no difference from the native uricase, but R-MGU was difficult to purify, resulting in slightly poorer thermal stability under the same conditions compared to the purified native enzyme. Conclusion: A uricase from a Meyerozyma guilliermondii strain was successfully cloned, and its active form was recombinantly expressed.

Full Text

Clone, Expression and Characterization of the Uricase from *Meyerozyma guilliermondii*

RAO Jing-jing[1], JING Yi-xian[1], ZOU Ming-yue[1], HU Xiao-lei[1], LIAO Fei[1], YANG Xiao-lan[1],*

[1]Key Laboratory of Clinical Laboratory Diagnostics of the Education Ministry, College of Laboratory Medicine, Chongqing Medical University, Chongqing 400016, China

Abstract

Objective: To clone the protein-coding gene of a uricase from *Meyerozyma guilliermondii* and confirm its identity through recombinant expression and characterization.

Methods: The yeast strain C.G.M.C.C 2.1008 was identified by rRNA sequencing. Tandem mass spectrometry was used to analyze peptide sequences of the native uricase to search for homologous proteins, and transcriptome sequencing verified its inducible expression. Based on the genomic information of *Meyerozyma guilliermondii* ATCC 6260, primers were designed to amplify the coding sequence of the putative uricase (Uniprot id: A5DFP1) from the cDNA of strain C.G.M.C.C 2.1008. After sequencing, the gene was cloned into directional expression vectors pDE1 and pDE2 to construct recombinant plasmids pDE1-MGU and pDE2-MGU encoding uricase with a 6His tag. Simultaneously, a tag-free expression plasmid R-MGU was constructed. The fungal uricase was induced for expression in *Escherichia coli* BL21(DE3). Peptide molecular weight was determined by SDS-PAGE and MALDI-TOF-MS. Kinetic parameters and inhibitor sensitivity were measured using uric acid as substrate and compared with those of the native enzyme.

Results: rRNA sequencing identified the strain as *Meyerozyma guilliermondii*. Tandem mass spectrometry analysis of tryptic peptides revealed high similarity between the native fungal uricase and uricase A5DFP1. Transcriptome sequencing supported efficient inducible expression of this uricase gene. Using designed primers, the coding sequence was rapidly obtained from cDNA by PCR. T-vector

sequencing revealed only one nucleotide difference from the A5DFP1 coding sequence at position 435 (C instead of G), though the amino acid remained unchanged. SDS-PAGE showed the recombinant R-MGU peptide at approximately 35 kDa, while MALDI-TOF-MS revealed a peptide mass of about 17.43 kDa, consistent with the native enzyme. However, the calculated molecular weight based on the amino acid sequence was only 33.98 kDa, suggesting possible chemical modification. The purified 6His-tagged recombinant uricase exhibited specific activity close to 6.0 U/mg. The Michaelis constant (Km), inhibition constants (Ki) for representative inhibitors, and molecular weight of R-MGU showed no significant differences from the native enzyme. However, R-MGU was difficult to purify, resulting in slightly lower thermostability compared to the purified native enzyme under identical conditions.

Conclusion: A uricase gene from *Meyerozyma guilliermondii* was successfully cloned and expressed in its active form through recombinant expression.

Keywords: *Meyerozyma guilliermondii* uricase; gene cloning; recombinant expression; characterization

---

Elevated serum uric acid causes hyperuricemia, which leads to gout, cardiovascular diseases, and other conditions [1,2]. Uricase (uricase, EC 1.7.3.3) specifically catalyzes the oxidation of uric acid to the more soluble allantoin, which is primarily excreted through the kidneys. Uricase serves as a key tool enzyme for enzymatic determination of serum uric acid in diagnosing hyperuricemia-related diseases [3-4], represents a candidate protein drug for treating hyperuricemia, and is currently the only effective medication for refractory gout [5-6]. All applications of uricase require high specific activity and good stability at physiological pH. However, existing natural uricases exhibit low specific activity even under optimal conditions, and their stability is limited. Most notably, the optimal pH of available uricases is alkaline, with less than 60% activity retained at physiological pH [1]. Therefore, uricases with good thermal stability and higher specific activity at physiological pH hold greater application potential.

Microorganisms represent the most important source of uricase for convenient recombinant expression. The yeast strain C.G.M.C.C 2.1008 from the China General Microbiological Culture Collection Center was previously found to express uricase under uric acid induction, demonstrating excellent thermal stability and relatively high specific activity with significant application potential [1]. Most attractively, this uricase has an optimal pH near neutrality, making it an ideal starting enzyme for molecular modification to improve catalytic efficiency (unpublished data). After purification by repeated negative adsorption with DEAE-cellulose to meet purity requirements, the native enzyme was subjected to Edman degradation in Shanghai, which revealed that the N-terminus was blocked. Further tandem mass spectrometry analysis showed that its amino acid sequence differed significantly from known uricases [7]. Previously, based on the N-terminal amino acid sequence obtained from Edman degradation of *Bacillus*

fastidious uricase, the majority of its coding sequence was successfully cloned using genomic DNA as template and degenerate primer PCR [8]. Unfortunately, based on the amino acid sequences of peptides obtained from tandem mass spectrometry analysis in Shanghai, repeated attempts using degenerate primer PCR with the fungal genomic DNA as template failed to clone any uricase-like gene sequences or peptides matching the discovered sequences. To rapidly clone the coding sequence of this fungal uricase, we first determined the rRNA sequence to confirm the strain as *Meyerozyma guilliermondii*. Preparative SDS-PAGE was used to purify the uric acid-induced native enzyme for renewed tandem mass spectrometry analysis, which revealed high similarity between its amino acid sequence and the putative uricase (Uniprot id: A5DFP1) deduced from the *Meyerozyma guilliermondii* genome. Transcriptome sequencing confirmed efficient inducible expression of this sequence. Furthermore, precise primers were designed based on the A5DFP1 coding sequence to clone the target gene from cDNA by PCR. Recombinant expression and comparison of properties with the native enzyme confirmed that the coding sequence of the uricase from this *Meyerozyma guilliermondii* strain had been obtained.

1.1.1 Strains and Plasmid Vectors

The yeast strain was purchased from the China General Microbiological Culture Collection Center (C.G.M.C.C 2.1008). The pDE1 and pDE2 plasmid vectors were derived from a directional cloning kit. Competent *E. coli* DH5 and BL21(DE3) cells were purchased from Beijing Dingguo Changsheng Biotechnology Co., Ltd.

1.1.2 Reagents

DNA ladder (DL2000) and rTaq polymerase were purchased from TaKaRa. Protein markers were from Thermo Fisher Scientific. T-vector kits and directional cloning kits were from Beijing Tsingke New Industry Biotechnology Co., Ltd. Plasmid extraction kits were from OMEGA (USA). Fungal RNA extraction kits and gel recovery kits were from MAGE. Reverse transcription kits were from Roche. DEAE-cellulose 52 chromatography columns and Ni-NTA columns were from Solarbio. All other reagents were analytical grade from domestic sources.

1.2.1 Strain Identification

An appropriate amount of bacterial suspension was spread on malt extract plates and incubated at 37°C for 24 h. The plates were transported on dry ice to Beijing Zhongmei Taihe Biotechnology Co., Ltd. and Chongqing Pandian Biotechnology Co., Ltd. for strain identification.

1.2.2 Induced Expression and Purification of Native Uricase

The yeast was cultured and induced with uric acid to obtain native uricase (MGU). The enzyme was purified three times by negative adsorption with

DEAE-cellulose 52 [1], concentrated, and dialyzed overnight against 20 mM Tris-HCl (pH 8.0). A 1.5 ml aliquot of enzyme solution was concentrated to 300 l using an ultrafiltration tube, and 25 l was taken for SDS-PAGE analysis.

1.2.3 Tandem Mass Spectrometry Identification

After SDS-PAGE, the gel was stained with Coomassie brilliant blue. After complete destaining, the visible bands were excised and numbered 1, 2, 3, and 4, then stored in sterilized EP tubes. The gel pieces were washed three times with sterilized double-distilled water and sent to the Biomedical Analysis and Testing Center of the Third Military Medical University for tandem mass spectrometry identification.

1.2.4 Transcriptome Sequencing Analysis

Uric acid-induced culture broth was collected and transported on dry ice to Shanghai Majorbio Bio-pharm Technology Co., Ltd. for transcriptome sequencing.

1.2.5 Amplification of the Uricase Gene

Fungal culture broth after uric acid induction was used to extract RNA with a fungal RNA extraction kit, followed by reverse transcription to obtain cDNA template, which was aliquoted and stored at -20°C. Based on the coding sequence of Uniprot id A5DFP1, a first pair of precise PCR primers was designed targeting the coding sequence for the first 8 amino acid residues at the N-terminus and the genomic sequence containing the stop codon plus 21 additional bases at the C-terminus (synthesized by Beijing Tsingke New Industry Biotechnology Co., Ltd.). The forward primer PF1 sequence was CACCATGTCTTTGGTCGCATCTTCATACG, and the reverse primer PR1 sequence was TTATAACTTGGCGTTGGCGTCTCTG. A second pair of precise primers was designed to facilitate connection of the PCR product to the expression vector: forward primer PF2 was CACCATGTCTTTGGTCGCATCTTCATACG, and reverse primer PR2 was TAACTTGGCGTTGGCGTCTCTG. The PCR reaction system (100 l) contained: 1 l rTaq polymerase, 10 l 10× buffer, 6 l MgCl , 8 l dNTP mix, 4 l PF1/PF2, 4 l PR1/PR2, with sterile water added to 100 l. PCR parameters were: initial denaturation at 95°C for 3 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, extension at 72°C for 1 min, and a final extension at 72°C for 5 min. PCR products were analyzed by 1.7% agarose gel electrophoresis and gel-purified.

1.2.6 T-Vector Ligation and Sequencing

The gel-purified PCR product was ligated into the T-vector according to the kit instructions and transformed into DH5 competent cells. After recovery in LB medium without antibiotics, an appropriate amount of bacterial suspension was spread on LB selection plates containing 100 mg/L ampicillin to screen

for transformants, which were incubated inverted at 37°C overnight. The next day, single colonies were picked and cultured in medium containing 100 mg/L ampicillin, then sent for sequencing.

1.2.7 Construction of Recombinant Uricase Expression Plasmids

Primers PF1/PR1 and PF2/PR2 enabled the PCR product to carry a CACC adapter for direct ligation with directional cloning expression vectors pDE1 and pDE2. The constructed plasmid pDE1-MGU expressed recombinant uricase with a 6His tag at the N-terminus, while pDE2-MGU expressed uricase with a 6His tag at the C-terminus. The pDE1-MGU plasmid was sent to Beijing Taihe Gene Co., Ltd. for mutagenesis to remove the N-terminal 6His tag and linker sequence, constructing the tag-free recombinant uricase plasmid R-MGU. These plasmids were transformed into *E. coli* DH5 competent cells by heat shock at 42°C, and transformants were selected on LB plates containing kanamycin and incubated inverted at 37°C overnight. The next day, positive clones were identified by colony PCR, cultured, and plasmids were extracted using a plasmid extraction kit and stored at -20°C.

1.2.8 Induced Expression of Recombinant Uricase

Recombinant uricase plasmids were transformed into *E. coli* BL21(DE3) competent cells by heat shock at 42°C, and transformants were selected on LB plates containing 100 mg/L kanamycin and incubated inverted at 37°C overnight. The next day, single colonies were inoculated into 3 ml of medium containing kanamycin and cultured overnight at 37°C with shaking at 180 rpm. The overnight culture was then inoculated into fresh 250 ml medium at a 1:1000 ratio and cultured at 37°C with shaking at 180 rpm until the OD reached 0.6–1.0. Expression was induced by adding isopropyl- -D-thiogalactoside (IPTG) to a final concentration of 0.2 mmol/L, followed by culture at 16°C with shaking at 110 rpm for 20–22 h. Cells were harvested by centrifugation at 4°C, 8000 rpm for 15 min. The cell pellet was resuspended in 20 mmol/L Tris-HCl buffer (pH 8.0), and cells were disrupted by ultrasonication on ice (33% amplitude, 5 s work, 5 s interval, 30 min total). The lysate was centrifuged at 4°C, 12,000 rpm for 20 min, and the supernatant was collected as crude enzyme solution.

1.2.9 Purification of Recombinant Uricase

Recombinant uricase proteins with N-terminal and C-terminal 6His tags were purified using Ni-NTA columns. Eluates collected with 300 mmol/L imidazole were considered once-purified enzyme solution. The Ni-NTA purification was repeated once, and the target protein eluates were collected and stored at 4°C.

The recombinant uricase expressed without a 6His tag had an isoelectric point of approximately 8.97. Based on the principle of negative adsorption, purification was performed using a DEAE-cellulose 52 chromatography column. The crude enzyme solution was loaded onto a DEAE-cellulose 52 column pre-equilibrated

with 20 mmol/L Tris-HCl (pH 8.0) buffer at a flow rate of 5–7 s/drop. After adsorption, the column was washed/eluted with equilibration buffer, and fractions of 1–2 ml were collected. Fractions with highest uricase activity were pooled, and the DEAE-cellulose 52 purification was repeated once. The purified target protein was stored at 4°C.

1.2.10 Characterization of Recombinant Uricase

Appropriate amounts of enzyme solution were subjected to 10% SDS-PAGE. Uricase activity was measured using uric acid as substrate [9-10], and protein concentration was determined by the Coomassie brilliant blue method to calculate specific activity. The Michaelis constant (Km) was determined by double reciprocal analysis [11]. The response of enzyme activity inhibition rate to inhibitor concentration logarithm was measured, and the half-maximal inhibitory concentration (IC) was determined by fitting the linear portion using the built-in functions of MS Excel 6.0. Inhibition constants for competitive inhibitors potassium oxonate and xanthine were calculated based on the Km values of different recombinant uricases. For thermostability assessment, purified uricase was filter-sterilized and added to filter-sterilized phosphate buffer (pH 7.4) and borate buffer (pH 9.2). The thermostability assay system contained final concentrations of 0.1 g/mL ampicillin and kanamycin, 0.1 mmol/L EDTA, and 2.0 mmol/L p-aminobenzamidine dihydrochloride. The stability assay system was placed in a 37°C incubator, and initial activity was measured after 20 min. Activity decay was recorded by measuring activity at different time intervals. The effect of potassium oxonate at a final concentration of 30 mol/L on thermostability was also examined. Native uricase MGU and recombinant uricase R-MGU purified by ion exchange negative adsorption were further subjected to acidic native electrophoresis followed by gel recovery [7]. Samples were concentrated and desalted using ultrafiltration tubes, then sent to the Biomedical Analysis and Testing Center of the Third Military Medical University for MALDI-TOF-MS analysis to compare their molecular weights.

2.1 Coding Sequence Information of MGU

The uricase from yeast strain C.G.M.C.C 2.1008 (MGU, hereafter referring to uricase from this strain) requires uric acid induction for expression [1,7], with its greatest advantage being an optimal pH near neutrality (unpublished data). Obtaining partial peptide or coding gene sequences enables rapid cloning of the complete coding sequence using degenerate primer PCR or cDNA screening [8]. After purifying MGU to a single peptide band by HPLC, Edman degradation sequencing revealed that the N-terminus was blocked. Tandem mass spectrometry analysis of tryptic peptides showed that its sequence differed significantly from known fungal uricases but was similar to a special dehydrogenase amino acid sequence (completed in Shanghai) [7]. Based on this similar protein peptide sequence, degenerate primers were designed for PCR amplification of the coding sequence using genomic DNA as template, but no meaningful sequence

information was obtained. Therefore, an alternative route was urgently needed to rapidly determine partial amino acid sequences or partial coding sequences of MGU to facilitate rapid gene cloning.

The Uniprot database contains a putative uricase A5DFP1 deduced from the genome sequence of *Meyerozyma guilliermondii* ATCC 6260, but recombinant expression has not been confirmed. To quickly confirm the amino acid sequence similarity between MGU and uricase A5DFP1, rRNA sequencing was first performed to identify the strain. The results showed that this strain belongs to *Meyerozyma guilliermondii*, though it could not be identified to the strain level (its 28S rRNA D1/D2 and ITS region sequences are available in Genbank under accession numbers KY688098 and KY688099, respectively). The initial tandem mass spectrometry analysis found no obvious similarity between MGU amino acid sequence and A5DFP1, possibly due to unknown experimental errors. After uric acid induction of the yeast cells to express native MGU, purification was performed three times using DEAE-cellulose 52 column chromatography, followed by SDS-PAGE resolution of four protein components for tandem mass spectrometry analysis (Figure 1). The most similar proteins corresponding to the four polypeptides in order of decreasing molecular weight were A5DJ39, A5DR94, A5DFP1, and A5DNM4. Among these, A5DJ39 is formate dehydrogenase; A5DR94 is an uncharacterized protein predicted to have hydrolase activity; and A5DNM4 is peptidyl-prolyl cis-trans isomerase. Unexpectedly, all high-abundance peptides successfully resolved by tandem mass spectrometry from component 3 (the most abundant) were found as identical peptides in the amino acid sequence of *Meyerozyma guilliermondii* uricase A5DFP1. These high-abundance peptides collectively covered nearly 90% of the A5DFP1 amino acid sequence (Figure 2 [Figure 2: see original paper]), and even some low-abundance peptides could be matched to A5DFP1 peptide sequences. MALDI-TOF-MS revealed a peptide molecular weight of approximately 17.4 kDa for MGU, while SDS-PAGE analysis showed approximately 35 kDa [1,7], but the molecular weight calculated from the A5DFP1 amino acid sequence was approximately 33.8 kDa [1,7]. Therefore, MGU is highly similar to A5DFP1 in sequence, but the MGU peptide may be longer or possess chemical modifications.

To confirm the sequence similarity between A5DFP1 and MGU, transcriptome sequencing was further performed. MGU expression depends on uric acid induction [1,7]. Transcriptome sequencing yielded two coding fragment reads that matched the DNA sequence of gene PGUG_02092 perfectly. Its transcript EDK37994 was identical to A5DFP1 and showed high expression levels (fpkm ranked 11th) (Figure 3 [Figure 3: see original paper]). Thus, the amino acid sequence of the inducible uricase MGU from yeast strain C.G.M.C.C 2.1008 is highly similar to A5DFP1, and its coding sequence should be rapidly clonable from cDNA using PCR.

2.2 Cloning the Coding Sequence of MGU from cDNA

Using cDNA as template and precise primers designed based on the coding sequence for the first 8 residues of A5DFP1 protein and the stop codon plus downstream genomic sequence at the C-terminus, PCR amplification directly yielded a unique product of approximately 900 bp (Figure 4 [Figure 4: see original paper]). The gel-purified PCR product was ligated into the T-vector and sequenced with universal primers, revealing only one nucleotide difference from the A5DFP1 coding sequence at position 435: A5DFP1 has G at this position, while the C.G.M.C.C 2.1008 yeast strain has C, though the corresponding amino acid sequence remained identical (see Genbank accession KY706244.1 for details). Repeated PCR amplification using cDNA as template and sequencing confirmed this position as C. PCR amplification using genomic DNA with the same primers also yielded consistent sequence. Raising the PCR annealing temperature by 3°C still produced the same unique PCR product (data not shown). MALDI-TOF-MS determination of the native enzyme molecular weight differed by nearly 1.0 kDa from that calculated based on the A5DFP1 amino acid sequence, but the N-terminus of native MGU was blocked and could not be sequenced by Edman degradation. Therefore, the MGU amino acid sequence may be identical to A5DFP1 except at the N-terminus, or the native enzyme may possess a large modification group.

2.3 Recombinant Expression and Characterization of MGU

Initial attempts to amplify the unknown N-terminal coding sequence of MGU using 5' -RACE yielded no meaningful new sequences. Recombinant expression of fungal uricase is relatively straightforward, so we first examined whether the recombinant product possessed the same properties as native MGU. Eukaryotic uricase is suitable for fusion expression with a 6His tag to facilitate purification [12]. However, both N- and C-termini of uricase polypeptide chains participate in forming the active center [13]; adding a 6His tag at either end of this fungal uricase might significantly reduce its activity, but it would be impossible to determine whether decreased activity resulted from the tag peptide, cloning errors, or incomplete N-terminal coding sequence. Therefore, we designed expression vectors with N-terminal 6His tag, C-terminal 6His tag, and no tag for expression and comparison.

All three forms of MGU showed significant uricase activity after induced expression (*E. coli* host cells have no endogenous uricase activity). Recombinant uricases pDE1-MGU and pDE2-MGU were purified by Ni-NTA chromatography; tag-free R-MGU was purified by two rounds of DEAE-cellulose chromatography and compared by SDS-PAGE. The results showed minimal differences in peptide molecular weight among the three expression vectors (Figure 5 [Figure 5: see original paper]). R-MGU purification by DEAE-cellulose chromatography was ineffective. However, using native acidic electrophoresis under conditions previously established for MGU separation [7], highly pure peptide bands were obtained but showed no uricase activity.

The enzymatic properties of the three recombinant MGU forms were determined. Specific activities differed significantly among the three forms (Table 1), but this clearly resulted primarily from differences in purification efficiency. No significant differences in Km were observed among the three forms (Table 2). Potassium oxonate and xanthine are representative strong competitive inhibitors of MGU [1,7]; sensitivity to these two competitive inhibitors also showed no significant differences among the three recombinant uricases (Table 2). Remarkably, all three recombinant uricase forms showed higher activity at pH 7.4 than at pH 9.2, similar to native MGU, demonstrating an optimal pH near neutrality.

To compare molecular sizes, R-MGU and native MGU were purified to single protein bands by preparative acidic electrophoresis [7]. Molecular weights of the peptide chains were measured by MALDI-TOF-MS using the same calibration standards. Native MGU had a molecular weight of 17.434 kDa, while R-MGU was 17.429 kDa (Figure 6 [Figure 6: see original paper]), with the difference within the measurement error of MALDI-TOF-MS. Notably, when MGU and R-MGU were mixed in approximately equal amounts and measured by MALDI-TOF-MS, a single peak was still obtained with a peptide molecular weight of 17.428 kDa. Thus, the tag-free recombinant R-MGU and native MGU may have identical amino acid sequences and chemical modifications.

Native MGU exhibits a plateau of essentially unchanged activity during the initial stage of thermal inactivation at pH 7.4 with potassium oxonate and at pH 9.2 [7]; none of the three recombinant forms showed this plateau during initial thermal inactivation (Figure 7 [Figure 7: see original paper]). At pH 7.4 without ligand, the thermal inactivation half-life of recombinant R-MGU was shorter than that of native MGU, but showed essentially no difference with ligand present. At pH 9.2 without ligand, the thermal inactivation half-life of C-terminal 6His-MGU was significantly shorter than that of native MGU, though this difference diminished with ligand present.

The results in Figure 5 and Table 1 indicate that tag-free R-MGU is more difficult to purify after recombinant expression in *E. coli* than in its native host cells [1,7], possibly because *E. coli* host cells contain many proteins with similar isoelectric points. Literature reports have successfully expressed eukaryotic uricase with a 6His tag [12]. However, N- and C-terminal sequences of uricase significantly affect its activity and thermal stability [13]. Comparison of recombinant MGU with 6His tags added at the N- or C-terminus showed that the tag position had minimal impact on maximum activity and substrate affinity (Tables 1 and 2), particularly regarding the pH effect on activity (data not shown). Thus, MGU is suitable for fusion expression with a 6His tag to facilitate rapid purification.

SDS-PAGE analysis showed protein molecular weights of approximately 35 kDa for both native and recombinant R-MGU, but possibly due to the polypeptide carrying two positive charges during MALDI-TOF-MS, the peptide chain molecular weight was approximately 17.4 kDa. Overall, our results support

that the amino acid sequence of native MGU is identical to that of recombinant R-MGU expressed in *E. coli* according to the A5DFP1 amino acid sequence. The N-terminus was found to be blocked during Edman degradation sequencing [7], indicating N-terminal chemical modification. Repeated MALDI-TOF-MS measurements of MGU peptide molecular weight were nearly 1.0 kDa larger than that calculated from the A5DFP1 amino acid sequence; such a large difference cannot result from experimental measurement error. The N-terminus of *Arthrobacter globiformis* uricase has a short peptide segment extending beyond its inter-chain -sheet, and this enzyme exhibits excellent thermal stability [13-16]. Deletion of N- or C-terminal sequences from *Bacillus fastidious* uricase significantly reduces its thermal stability [13, 17-18]. Therefore, native MGU may have special modifications related to its thermal stability, and R-MGU can reproduce the N-terminal modification of native MGU, laying the foundation for identifying the chemical modification and its significance in MGU enzymatic properties.

The value of pharmaceutical enzymes depends on their specific activity and thermal stability. The thermal stability of the three recombinant MGUs differed little from native MGU at pH 7.4, but the difference increased at pH 9.2 (Figure 7). The 6His tags at the N- or C-terminus may affect recombinant MGU thermal stability. The low purity of R-MGU and contamination with proteases would certainly reduce its thermal stability through protein degradation. Existing uricases have alkaline optimal pH values, reducing their value as diagnostic tool enzymes and pharmaceutical enzymes. MGU is the first uricase discovered to date with an optimal pH near neutrality. Elucidating the determinants of its optimal pH is undoubtedly valuable for developing pharmaceutical uricases with higher activity under physiological conditions and diagnostic tool enzymes suitable for two-reagent serum uric acid assays [16]. MGU possesses an unidentified chemical modification that may be related to its optimal pH and thermal stability; deciphering the nature of this chemical modification and its significance in determining optimal pH and thermal stability, as well as identifying the determinants of its optimal pH, represent the focus of future research. The peptide information from the initial tandem mass spectrometry analysis commissioned in Shanghai misled the direction of MGU coding sequence cloning; multi-angle cross-validation is essential for obtaining reliable experimental results. It was precisely through multi-angle cross-validation of sequence similarity between MGU and A5DFP1 that this study could rapidly confirm the coding sequence of the neutral-pH-optimal uricase from strain C.G.M.C.C 2.1008, laying the foundation for subsequent research.

References

[1] Wu W M, Zeng Z C, Li X Y, et al. Purification and characterization of the uricase from *Candida utilis*. Journal of Southwest University (Natural Science Edition), 2008, 30(3):84-89.

[2] Ishizaka N, Ishizaka Y, Toda E, et al. Association between serum uric acid,

metabolic syndrome, and carotid atherosclerosis in Japanese individuals. *Arterioscler Thromb Vasc Biol*, 2005, 25(5):1038-1044.

[3] Duncan P, Gochman N, Bayse D, et al. A candidate reference method for uric acid in serum. II. Interlaboratory testing. *Clinical Chemistry*, 1982, 28(2):291-293.

[4] Koyama Y, Ichikawa T, Nakano E. Cloning, Sequence Analysis, and Expression in *Escherichia coli* of the Gene Encoding the *Candida utilis* Urate Oxidase (Uricase). *Journal of Biochemistry*, 1996, 120(5):969-973.

[5] Pui C H, Relling M V, Lascombes F, et al. Urate oxidase in prevention and treatment of hyperuricemia associated with lymphoid malignancies. *Leukemia*, 1997, 11(11):1813-1816.

[6] Zhu X J, Liu J G, Li G X. Cloning and expression of urate oxidase and its application in serum uric acid analysis. *Chinese Journal of Biotechnology*, 2001, 17(1):68-72.

[7] Wu J, Yang X, Wang D, et al. A numerical approach for kinetic analysis of the nonexponential thermoinactivation process of uricase. *Protein Journal*, 2016, 35(4):1-12.

[8] Zhao Y S, Bu Y Q, Liao F. Rapid cloning gene sequence of the uricase from *Bacillus fastidious* by degenerated primer. *Journal of Chongqing Medical University*, 2017, 42(02):219-224.

[9] Zhao Y S, Zhao L N, Yang G Q, et al. Characterization of an extracellular uricase from *Bacillus fastidious* and its application in direct kinetic assay of serum uric acid. *Acta Medica Universitatis Scientiae et Technologiae Huazhong*, 2007, 36(2):239-242.

[10] Liao F, Zhu X Y, Wang Y M, et al. The comparison of the estimation of enzyme kinetic parameters by fitting reaction curve to the integrated Michaelis-Menten rate equations of different predictor variables. *Journal of Biochemical & Biophysical Methods*, 2005, 62(1):13-24.

[11] Liu X, Wen M, Li J, et al. High-yield expression, purification, characterization, and structure determination of tag-free *Candida utilis* uricase. *Applied Microbiology & Biotechnology*, 2011, 92(3):529-537.

[12] Wu S L, Chen B, Liu C Q, et al. Expression in *Escherichia coli*, purification and enzymatic properties of porcine urate oxidase. *Chinese Journal of Biotechnology*, 2009, 25(11):1664-1670.

[13] Feng J, Wang L, Liu H, et al. Crystal structure of *Bacillus fastidious* uricase reveals an unexpected folding of the C-terminus residues crucial for thermostability under physiological conditions. *Applied Microbiology & Biotechnology*, 2015, 99(19):7973-7986.

[14] Chen D, Han J, Shi H N, et al. Expression, purification and activity of

Arthrobacter globiformis uricase. *Chinese Journal of Biologicals*, 2014, 27(3):351-355.

[15] Suzuki K, Sakasegawa S I, Misaki H, et al. Molecular cloning and expression of uricase gene from *Arthrobacter globiformis* in *Escherichia coli* and characterization of the gene product. *Journal of Bioscience & Bioengineering*, 2004, 98(3):153-158.

[16] Huang Y, Chen Y, Yang X, et al. Optimization of pH values to formulate the bireagent kit for serum uric acid assay. *Biotechnology & Applied Biochemistry*, 2015, 62(1):137.

[17] Zhang C, Yang X, Feng J, Yuan Y, Li X, Bu Y, Xie Y, Yuan H, Liao F. Effects of modification of amino groups with poly(ethylene glycol) on a recombinant uricase from *Bacillus fastidiosus*. *Biosci Biotechnol Biochem*. 2010; 74(6):1298-1301.

[18] Zhang C, Yang X, Gao A, Hu X, Pu J, Liu H, Feng J, Liao J, Li Y, Liao F. Comparison of modification of a bacterial uricase with N-hydroxysuccinimide esters of succinate and carbonate of monomethoxyl poly(ethylene glycol). *Biotechnol Appl Biochem*. 2014; 61(6):683-690.

Note: Figure translations are in progress. See original paper for figures.