Expression of codon-optimized TgMIC16 in three *Escherichia coli* strains

Hui Sun¹ · Jin Li¹ · Gong-zhen Liu¹ · Kun Yin¹ · Yong Cui¹ · Ting Xiao¹ ·
Chao Xu¹ · Bing-cheng Huang¹ · Qing-kuan Wei¹

Received: 13 June 2017 / Accepted: 11 July 2017 / Published online: 31 July 2017
© Springer-Verlag GmbH Germany 2017

Abstract In a previous study, we found that rabbit anti-*Toxoplasma gondii* serum was capable of recognizing truncated *T. gondii* microneme protein 16 (TgMIC16), indicating that TgMIC16 is an essential antigenic *T. gondii* protein. However, the broad application of this recombinant protein is limited by its low expression level. In this study, we performed codon optimization of *TgMIC16* by changing the codon-adaptation index from 0.22 to 1.0 without altering the amino acid sequence and expressed the optimized gene in three different *Escherichia coli* strains, followed by comparison of soluble recombinant-protein expression and yield. Our results showed that the recombinant protein rTgMIC16 was expressed as inclusion bodies in all three strains following optimization of induction parameters, and western blot analysis revealed the presence of a ~72-kD recombinant protein as a specific band following purification. A shuffle-expression strain was selected to amplify incubation products and induce expression, resulting in an overall rTgMIC16 yield of ~20 mg/L. These findings provide a basis for further investigation of TgMIC16 to elucidate its functions and interaction partners.

Keywords *Toxoplasma gondii* · Microneme protein 16 · Prokaryotic expression · Codon optimization

Hui Sun and Jin Li contributed equally to this work.

✉ Bing-cheng Huang
hbc863@hotmail.com

✉ Qing-kuan Wei
h267310601@126.com

¹ Shandong Academy of Medical Sciences, Shandong Institute of Parasitic Diseases, 11 Taibai Middle Road, Jining 272033, Shandong, China

Introduction

Toxoplasmosis is caused by *Toxoplasma gondii* and represents a source of zoonotic infectious diseases worldwide (Dubey 2008; Tenter et al. 2000). *T. gondii* is an opportunistic pathogen capable of infecting a variety of warm-blooded animals. Usually, an immunocompetent individual will be clinically asymptomatic following *T. gondii* infection; however, in immunocompromised patients, including those with acquired immune deficiency syndrome, those having undergone organ transplantation, or those with malignant tumors, and pregnant women, *T. gondii* infection can cause severe complications (Montoya and Liesenfeld 2004; Flegr 2013; Torrey and Yolken 2013). In addition, previous studies reported relationships between latent toxoplasmosis and serious psychiatric and neurological diseases (Hinze-Selch et al. 2010; Miman et al. 2010; Kusbeci et al. 2011). Furthermore, during the process of animal husbandry, *T. gondii* infection can cause abortions in pregnant animals, resulting in significant economic losses.

Toxoplasma gondii served as a model organism for the study of Apicomplexa and has been studied for many years. Invasion of host cells is the key step necessary for *T. gondii* to sustain survival and propagation (Hoane et al. 2003). Micronemes, rhoptries, and dense granules are three secretory organelles that participate in this process, with microneme proteins (MICs) participating in parasite recognition, reorientation, and invasion. Previous studies showed that at least 27 MICs have been reported in *T. gondii*, with some subsequently used as candidates for genetic engineering vaccines (Zhang et al. 2013, 2015; Jongert et al. 2009). The monomeric form of some MICs play roles in cell infection, whereas others are involved in parasite invasion in the form of MIC complexes (MIC1/4/

6, MIC2/2AP, and MIC3/8) (Reiss et al. 2001; Huynh et al. 2003; Crde et al. 2005). As a key protein involved in invasion, MIC shares similar conserved domains, including epidermal growth-factor-like domain, integrin A-like domain, apple module, thrombospondin-type repeat, galectin-like domain, chitin-binding-like domain, and microneme adhesive repeat domain, with higher eukaryotes (Soldati et al. 2001; Tomley and Soldati 2001).

There is no commercial vaccine for toxoplasmosis in humans; therefore, there is a need to discover an effective and protective antigen against toxoplasmosis. In a previous study, we found that rabbit anti-*T. gondii* serum recognized truncated *T. gondii* microneme protein 16 (TgMIC16), indicating that TgMIC16 is an essential antigenic *T. gondii* protein (Li et al. 2016). However, the broad application of this recombinant protein is limited by its low expression level. Here, we codon-optimized TgMIC16 and expressed the optimized gene in three different *Escherichia coli* strains to compare their abilities to produce soluble protein at high yields.

Materials and methods

Strains and plasmids

Escherichia coli TOP10 cells were used as the host for recombinant DNA manipulation, and BL21 (DE3), ArcticExpress (DE3), and Shuffle cells were used for gene-expression experiments. In addition, plasmid pET-30a(+) was used for cloning and expression. All hosts and plasmids were stored in our central laboratory.

Codon optimization and gene synthesis

The wild-type TgMIC16 gene was selected for codon optimization according to the a codon-adaptation tool (<http://www.jcat.de>) (Grote et al. 2005) and OptimumGene (GenScript, Nanjing, China). To clone the gene into the pET-30a(+) vector, restriction sites *Nde*I and *Hind*III were introduced at the 5' and the 3' ends of the optimized sequence, respectively. The gene was synthesized by GenScript and cloned into the pEASY-blunt vector (TransGen Biotech, Beijing, China), resulting in pEASY-mTgMIC16.

Subcloning of the optimized TgMIC16 gene into the pET-30a(+) vector

The pET-30a(+) vector and the pEASY-mTgMIC16 were double digested with *Nde*I and *Hind*III, and the TgMIC16 gene was ligated into the linearized plasmid pET-30a(+) to attain the prokaryotic expression vector pET30-mTgMIC6.

The resulting plasmid was transformed into *E. coli* TOP10 cells, and generated transformants were screened with *Nde*I and *Hind*III and confirmed by DNA sequencing. The plasmids positive for the recombinant gene were purified and transformed into *E. coli* BL21 (DE3), ArcticExpress (DE3), and Shuffle cells.

Parameter optimization and protein expression

Strains expressing the recombinant gene were grown at 37 °C with shaking at 220 rpm in Luria-Bertani medium containing 100 µg/mL kanamycin until an optical density at 600 nm of 0.5. Expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.8 mM, and recombinant protein expression was evaluated at different incubation temperatures (15 and 37 °C) and different induction times (16 and 4 h).

Estimation of recombinant protein-expression levels

Each culture sample was harvested at various time points, and the medium was discarded. Following centrifugation (8000×g for 20 min), the pellet was disrupted by sonication with the sample in an ice bath. Following another round of centrifugation, the pellet was resuspended in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer, and the supernatant and precipitant were subjected to 12% SDS-PAGE analysis after boiling the sample for 5 min.

rTgMIC16 purification by affinity chromatography

The His-tagged fusion protein was purified by nickel-affinity chromatography according to manufacturer instructions (Qiagen, Hilden, Germany). In brief, cells were disrupted by sonication, and the supernatant was separated by centrifugation at 4 °C and 10,000×g for 20 min. The supernatant was then transferred to an Ni-NTA column, and unbound contaminant proteins were removed after washing with buffer [20 mM Tris-HCl (pH 8.0), 20 mM imidazole, and 500 mM NaCl]. His-tagged proteins were eluted with elution buffer [20 mM Tris-HCl (pH 8.0), 250 mM imidazole, and 500 mM NaCl]. The inclusion bodies containing the recombinant protein were resuspended in buffer [20 mM Tris-HCl (pH 8.0), 20 mM imidazole, 8 M urea, and 500 mM NaCl] and purified by Ni-NTA chromatography. The eluted fractions containing the pure fusion protein were dialyzed overnight in phosphate-buffered saline (pH 7.4). The concentration of purified rTgMIC16 was determined using a BCA protein assay kit (Sigma-Aldrich, St. Louis, MO, USA), and rTgMIC16 purity was evaluated by SDS-

PAGE and quantified using Quantity One software (Bio-Rad, Hercules, CA, USA).

Western blot analysis

All protein samples were separated by 12% reducing SDS-PAGE and transferred to a polyvinylidene fluoride membrane. The membrane was blocked with 2.5% (w/v) skim milk (BD Biosciences, Franklin Lakes, NJ, USA), and successively reacted with an anti-His monoclonal antibody and a horseradish peroxidase-conjugated rabbit anti-mouse antibody for 1 h at 37 °C. After washing, the protein bands were developed by chromogenic reaction.

Results

Design and bioinformatics analysis

In this study, we expressed the *TgMIC16* gene from the start codon to the transmembrane domain. The open reading frame of the *TgMIC16* gene is 1884 bp, has a G+C content of 53.18%, and encodes a protein of 668 amino acids. The codon-adaptation index was optimized from 0.22 for wild-type *TgMIC16* to 1.0 in the optimized form, mTgMIC16 (Fig. 1). Sequence alignment showed that the mTgMIC16 sequence shared 74.6% identity with TgMIC16, with identical amino acid sequences.

Construction of expression transformants

The mTgMIC16 product removed from the pEASY-mTgMIC16 vector was ligated into the pET-30a(+) vector and verified for *NdeI/HindIII* digestion and sequencing (Fig. 2). The pET30-mTgMIC16 vector was transformed into BL21 (DE3), ArcticExpress (DE3), and Shuffle cells, resulting in the expressions of the recombinant protein rTgMIC16 in strains BMIC16, AMIC16, and SMIC16, respectively. Positive transformants were selected for further optimization and expression.

Optimization of rTgMIC16 expression

Induction temperature and time requires optimization for effective protein expression. To determine their effects on rTgMIC16 expression, BMIC16, AMIC16, and SMIC16 were inoculated in the same medium and using the same IPTG concentration. Our results showed that all three expression strains were effective at 15 °C for 16 h and 37 °C for 4 h, with noninduced strains used as negative controls (Fig. 3).

Solubility analysis of rTgMIC16

To investigate the variability in yield according to the induction parameters, we determined rTgMIC16 solubility under different conditions. The supernatant and cell lysis fractions were analyzed using SDS-PAGE following sonication, with results showing that the proteins expressed by BMIC16, AMIC16, and SMIC16 mainly existed as inclusion bodies, with little or no recombinant protein present in the soluble form (Fig. 4).

Purification and detection of the fusion protein

Protein purification and enrichment was performed to determine levels of soluble protein. Soluble proteins and inclusion bodies were purified by Ni-NTA, and affinity-purified proteins were analyzed by SDS-PAGE and western blot. Western blot analysis revealed the presence of a ~72-kD recombinant protein as a specific band in the eluted fraction, with the recombinant rTgMIC16 protein existing as part of an inclusion body and not as a soluble protein in any of the three expression strains (Fig. 5). SMIC16 was subsequently selected for culture incubation and expression induction, with overall yields of rTgMIC16 at ~20 mg/L and a final protein concentration of 0.8 mg/mL after refolding (purity >85%).

Discussion

The heterologous expression of recombinant proteins plays an important role in the research of protein structure, function, and drug–target selection, resulting in its wide use in basic science research, structural biology, and pharmacology. In 1977, Itakura et al. successfully expressed mammalian somatostatin in *E. coli*, representing the first expression of an exogenous gene in prokaryotic cells (Itakura et al. 1977). The *E. coli*-expression system has been widely used due its being a well-characterized host, a time-saving method, and being the low-cost and simple purification process associated with protein expression (Choi and Lee 2004). The use of prokaryotic expression vectors, host strains, and the correct induction conditions are important in terms of recombinant protein expression. In a previous study, we used two expression vectors [pET-28a(+) and pET-32a(+)] and expression strains (*E. coli* BL21 and Rosetta cells) to express the *TgMIC16* gene, but little or no recombinant protein was attained (Li et al. 2016). Therefore, we hypothesized that codon usage between *T. gondii* and the *E. coli* might be one explanation for this outcome.

Because codon usage is degenerate, more than one codon can specify the same amino acid, with codon

Fig. 1 Sequence comparison between the original and optimized forms of the *TgMIC16* gene

Fig. 2 Verification of the recombinant plasmid pET30a-mTgMIC16. Lane 1 pET30a-mTgMIC16 digested with *Nde*I and *Hind*III. Lane M DNA marker

preference present in both prokaryotic and eukaryotic organisms. Therefore, rare codon usage can greatly reduce protein expression and potentially halt protein synthesis (Tuller et al. 2010). To address this problem, the TgMIC16 gene sequence was initially redesigned by replacing less frequently used codons in *E. coli* with preferred codons corresponding to the same amino acids as a method to upgrade the codon-adaptation index from 0.22 to 1.0. Codons, such as CGA (Arg), CGG (Arg), AGG (Arg), AGA (Arg), AUA (Ile), CUA (Leu), CCC (Pro), and GGG (Gly), are usually considered rare codons in *E. coli*; therefore, 81% (43/53), 37% (10/28), and 47% (16/34) of the codons for Arg, Pro, and Gly, respectively, were altered here to optimize expressions in *E. coli*.

Recombinant protein production was also promoted by optimizing induction conditions, such as temperature and time (Narciandi et al. 1996). At higher temperatures, cells grow and proliferate rapidly, resulting in possible aggregation of recombinant proteins into inclusion

Fig. 3 SDS-PAGE analysis of rTgMIC16 expressed in three *Escherichia coli* strains. Lane M protein marker; lane NC noninduced whole-cell lysates; lane 1 whole-cell lysates induced at 15 °C for 16 h; lane 2 whole-cell lysates induced at 37 °C for 4 h

Fig. 4 Solubility analysis of rTgMIC16. Lane M protein marker; lane NC1 noninduced supernatant; lane NC2 noninduced precipitant; lane 1 supernatant of culture induced at 15 °C for 16 h; lane 2 precipitant

culture induced at 15 °C for 16 h; lane 3 supernatant of culture induced at 37 °C for 4 h; lane 4 precipitant of culture induced at 37 °C for 4 h

Fig. 5 Western blot analysis of purified rTgMIC16. Lane M protein marker; lanes 1 and 3 purified soluble protein reacted with His-tag antibody; lanes 2 and 4 purified insoluble protein reacted with His-tag antibody

bodies. Incubation temperature reduction can be an effective approach to promote the proper folding of recombinant proteins and prevent the formation of inclusion bodies. However, in this study, all three expression strains were unable to express the recombinant protein as a soluble product at either 37 or 15 °C. Formation of inclusion bodies represent a complex and dynamic process affected by various factors, with no universal approach for the efficient production of soluble forms of aggregation-prone recombinant proteins (Baneyx and Mujacic 2004; Sørensen and Mortensen 2005; Grigoroudis et al. 2015). Several strategies, including introduction of the expression of molecular chaperones in host bacteria, modulation of target-protein expression levels, optimization of culture medium, and addition of exogenous chaperones, might potentially prevent rTgMIC16 recombinant aggregation and will be tested in future studies.

In conclusion, our results indicated successful expression of the mTgMIC16 gene in three different *E. coli* strains, with the protein product expressed as a part of inclusion bodies following codon optimization. The highest productivity of rTgMIC16 was achieved (~20 mg/L) following systematic optimization of the expression conditions. The strategies described here and outside of engineered hosts or temperature should be tested to determine their usefulness in production of soluble rTgMIC16 for further study.

Acknowledgements This work was supported by the grant from the National Natural Science Foundation of China to HS (No. 81501770), the grant from the Natural Science Foundation of Shandong Province to HS (No. ZR2015YL051), and the grant from The Innovation Project of Shandong Academy of Medical Sciences.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

- Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol 22:1399–1408. doi:10.1038/nbt1029
- Cérède O, Dubremetz JF, Soète M, Deslée D, Vial H, Bout D, Lebrun M (2005) Synergistic role of micronemal proteins in *Toxoplasma gondii* virulence. J Exp Med 201:453–463. doi:10.1084/jem.20041672
- Choi JH, Lee SY (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl Microbiol Biotechnol 64:625–635. doi:10.1007/s00253-004-1559-9
- Dubey JP (2008) The history of *Toxoplasma gondii*—the first 100 years. J Eukaryot Microbiol 55:467–475. doi:10.1111/j.1550-7408.2008.00345.x
- Flegr J (2013) How and why *Toxoplasma* makes us crazy. Trends Parasitol 29:156–163. doi:10.1016/j.pt.2013.01.007
- Grigoroudis AI, McInnes C, Premnath PN, Kontopidis G (2015) Efficient soluble expression of active recombinant human cyclin A2 mediated by *E. coli* molecular chaperones. Protein Expr Purif 113:8–16. doi:10.1016/j.pep.2015.01.013 [Epub 2015 May 6]
- Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, Jahn D (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. doi:10.1093/nar/gki376
- Hinze-Selch D, Däubener W, Erdag S, Wilms S (2010) The diagnosis of a personality disorder increases the likelihood for seropositivity to *Toxoplasma gondii* in psychiatric patients. Folia Parasitol 57:129–135
- Hoane JS, Carruthers VB, Striemen B, Morrison DP, Entzeroth R, Howe DK (2003) Analysis of the *Sarcocystis neuromus* microneme protein SnMIC10: protein characteristics and expression during intracellular development. Int J Parasitol 33:671–679. doi:10.1016/S0020-7519(03)00031-6
- Huynh MH, Rabenau KE, Harper JM, Beatty WL, Sibley LD, Carruthers VB (2003) Rapid invasion of host cells by *Toxoplasma* requires secretion of the MIC2–M2AP adhesive protein complex. EMBO J 22:2082–2090. doi:10.1093/emboj/cdg217
- Itakura K, Hirose T, Crea R, Riggs A, Heyneker H, Bolivar F, Boyer H (1977) Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science 198:1056–1063. doi:10.1126/science.412251
- Jongert E, Roberts CW, Gargano N, Förster-Wald E, Petersen E (2009) Vaccines against *Toxoplasma gondii*: challenges and opportunities. Mem Inst Oswaldo Cruz 104:252–266. doi:10.1590/S0074-02762009000200019

- Kusbeci OY, Miman O, Yaman M, Aktepe OC, Yazar S (2011) Could *Toxoplasma gondii* have any role in Alzheimer disease? *Alzheimer Dis Assoc Disord* 25:1–3. doi:[10.1097/WAD.0b013e3181f73bc2](https://doi.org/10.1097/WAD.0b013e3181f73bc2)
- Li J, Cui Y, Yin K, Liu G-Z, Xiao T, Xu C, Huang B-C, Wei Q-K, Sun H (2016) Prokaryotic expression, purification and immunological characterization of micronemal protein 16 of *Toxoplasma gondii*. *Chin J Parasitol Parasit Dis* 34(3):198–202 (in chinese)
- Miman O, Kusbeci OY, Aktepe OC, Cetinkaya Z (2010) The probable relation between *Toxoplasma gondii* and Parkinson's disease. *Neurosci Lett* 475:129–131. doi:[10.1016/j.neulet.2010.03.057](https://doi.org/10.1016/j.neulet.2010.03.057)
- Montoya JG, Liesenfeld O (2004) Toxoplasmosis. *Lancet* 363(9425):1965–1976. doi:[10.1016/S0140-6736\(04\)16412-X](https://doi.org/10.1016/S0140-6736(04)16412-X)
- Narciandi RE, Mörbe FJ, Riesenberg D (1996) Maximizing the expression of recombinant Kringle 1 (Streptokinase) synthesized in *Escherichia coli*. Influence of culture and induction conditions. *Biotechnol Lett* 18(11):1261–1266. doi:[10.1007/BF00129951](https://doi.org/10.1007/BF00129951)
- Reiss M, Viebig N, Brecht S, Fourmaux MN, Soete M, Di Cristina M, Dubremetz JF, Soldati D (2001) Identification and characterization of an escorter for two secretory adhesins in *Toxoplasma gondii*. *J Cell Biol* 153:563–578
- Soldati D, Dubremetz JF, Lebrun M (2001) Microneme proteins: structural and functional requirements to promote adhesion and invasion by the apicomplexan parasite *Toxoplasma gondii*. *Int J Parasitol* 31:1293–1302. doi:[10.1016/S0020-7519\(01\)00257-0](https://doi.org/10.1016/S0020-7519(01)00257-0)
- Sørensen HP, Mortensen KK (2005) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact* 4:1. doi:[10.1186/1475-2859-4-1](https://doi.org/10.1186/1475-2859-4-1)
- Tenter AM, Heckeroth AR, Weiss LM (2000) *Toxoplasma gondii*: from animals to humans. *Int J Parasitol* 30:1217–1258. doi:[10.1016/S0020-7519\(00\)00124-7](https://doi.org/10.1016/S0020-7519(00)00124-7)
- Tomley FM, Soldati DS (2001) Mix and match modules: structure and function of microneme proteins in apicomplexan parasites. *Trends Parasitol* 17:81–88. doi:[10.1016/S1471-4922\(00\)01761-X](https://doi.org/10.1016/S1471-4922(00)01761-X)
- Torrey EF, Yolken RH (2013) *Toxoplasma* oocysts as a public health problem. *Trends Parasitol* 29:380–384. doi:[10.1016/j.pt.2013.06.001](https://doi.org/10.1016/j.pt.2013.06.001)
- Tuller T, Waldman YY, Kupiec M, Ruppin E (2010) Translation efficiency is determined by both codon bias and folding energy. *Proc Natl Acad Sci USA* 107:3645–3650. doi:[10.1073/pnas.0909910107](https://doi.org/10.1073/pnas.0909910107)
- Zhang N-Z, Chen J, Wang M, Petersen E, Zhu X-Q (2013) Vaccines against *Toxoplasma gondii*: new developments and perspectives. *Expert Rev Vaccines* 12:1287–1299. doi:[10.1586/14760584.2013.844652](https://doi.org/10.1586/14760584.2013.844652)
- Zhang N-Z, Wang M, Xu Y, Petersen E, Zhu X-Q (2015) Recent advances in developing vaccines against *Toxoplasma gondii*: an update. *Expert Rev Vaccines* 14:1609–1621. doi:[10.1586/14760584.2015.1098539](https://doi.org/10.1586/14760584.2015.1098539)