Borneo International Journal of Biotechnology (BIJB)
Vol. 2 (December 2022), 44 – 53

e-ISSN 2716-697X

Design, cloning and expression of a synthetic gene encoding a grouper iridovirus major capsid protein

Fernandes Opook[1], Teoh Peik Lin[1], B. A. V. Maran[2] and Kenneth F. Rodrigues[1*]

[1]Biotechnology Research Institute, Universiti Malaysia Sabah,
Kota Kinabalu, Sabah, Malaysia
[2]Borneo Marine Research Institute, Universiti Malaysia Sabah,
Kota Kinabalu, Sabah, Malaysia
[*]Corresponding author's email: kennethr@ums.edu.my
https://doi.org/10.51200/bijb.v2i.2958
Received: 15 June 2021 | Revised: 20 January 2022 | Accepted: 28 June 2022 | Published: 1 December 2022

ABSTRACT

Synthetic biology approaches can be adopted to successfully redesign, clone and express recalcitrant proteins from viruses in a bacterial expression system. The grouper iridovirus major capsid proteins (GI-MCP) is a surface coat protein which has the potential for application as an antigen for the development of vaccines which confer immunity against GI. The native amino acid coding region of the GI-MCP does not lend itself to expression in the *Escherichia coli* (BL21) (DE3) platform due to the presence of internal motifs which represent the start and stop codons as well as secondary structures. This manuscript reports on the redesigning, cloning and expression of a synthetic GI-MCP in *E. coli*. The native GI-MCP protein coding region was retrieved from the NCBI GenBank. The sequence was codon optimized, internal start and stop codons were modified and potential secondary structures were resolved. Both the native and modified GI-MCP coding DNA sequences were synthesized and ligated onto an expression plasmid, pET22b (+) followed by transformation, cloning and expression in *E. coli* (BL21) (DE3). Induction of protein expression was carried out at 12°C, 20°C and 37°C to assess protein solubility. The gene encoding the native GI-MCP failed to express under any of the culture conditions, however, the modified synthetic gene encoding the GI-MCP expressed the recombinant iridovirus coat protein (rGI-MCP) under all the conditions. Gene design and synthesis offer an ideal solution for the recombinant expression of viral proteins in bacterial expression systems for the industrial production of viral antigens.

Keywords: grouper iridovirus, major capsid protein, codon optimization, synthetic biology

INTRODUCTION

Grouper iridovirus (GI) is a debilitating fish virus which was first isolated from diseased grouper and subsequently identified as the causative agent for haemorrhage and enlargement of the spleen in challenged fishes (Qin et al.,2001). The disease has resulted in economic losses to the aquaculture industry and has spread across the Asia Pacific region to China (Wei et al.,2019). The grouper iridovirus reported from Singapore, designated as the Singapore grouper iridovirus (SGIV) is an icosahedral virus with a diameter of approximately 160 nm and a genome which consists of a double-stranded DNA with a size of 140,131 base pairs which encodes 162 open reading frames (ORFs) (Tsai et al.,2005). Antigenic characterization of the IV isolated from infected fish revealed a fine immune reactive band at 49 kDa which corresponds to the size of the major capsid protein (Qin et al., 2002). This experimental evidence indicates that the host produces specific antibodies against the major capsid protein. The development of vaccines for fish has been necessitated by the rapid development of the aquaculture industry (Munang'andu & Evensen, 2019) and the vaccines for fish are currently in the process of being developed for a wide range of pathogens and deployed in China (Wang et al., 2020), Japan (Matsuura et al., 2019), Korea (Hwang et al., 2020). The development of a vaccine against GIV has been undertaken using bioinformatic approaches (Ou-Yang et al., 2012), biosynthesis of recombinant antigens for application as vaccines (Liu et al., 2015) and the formulation of antigens with nanomaterial such as single-walled carbon nanotubes (SWCN) (Liu et al., 2020) and experimental approaches based on DNA aptamers which bind to the GIV (Li et al., 2014). The current study was directed toward the development of a production system for the mass production of the GI-MCP for application in aquaculture systems using a codon-optimized synthetic gene. Codon optimization has been reported to improve the translation of native genes in bacterial expression systems including *E. coli* (Burgess-Brown et al., 2008) by substituting rare codons with those native to the expression system led to an increase in the yield of the recombinant protein as well as higher fidelity of expression (Menzella et al., 2011). This study was directed towards the optimization of the gene encoding the GI-MCP in *E. coli* by adopting a synthetic biology approach.

MATERIAL AND METHODS

Gene Design and Synthesis

The nucleotide sequences coding for the Iridovirus major capsid protein (GI-MCP) gene (GenBank Accession Number: AY666015) was retrieved from the NCBI database. The protein coding region was modified by the removal of rare codons followed by substitution with preferred codons of *E. coli* according to *E. coli* codon usage data (Puigbo et al., 2007) from the Kazusa DNA Research Institute. The modification was applied without disrupting the amino acid sequences of the target genes. Both

the native and optimized nucleotide sequences were synthesized (GENEWIZ, USA) with the inclusion of 5'– *Bam*HI and 3'– *Eco*RI restriction enzyme sites. The synthetic genes were cloned into the pUC57 vector by GENEWIZ and delivered in the form of lyophilized plasmids. Native and modified sequences of target genes were analysed by the Countcodon program (http://www.kazusa.or.jp/codon/countcodon.html) where the codons were clustered into codon groups according to the amino acids formed. Gene construction. DNA restriction was performed with the restriction enzymes *Bam*HI, *Eco*RI and *Nco*I (Thermo Scientific, USA). The reaction mixture was prepared in 20 µl in PCR tubes. The plasmid pUC57 from GENEWIZ harbouring the synthetic genes and expression vector pET22b were digested using restriction enzymes *Bam*HI and *Eco*RI. The digestion reaction resulted in sticky ends on both synthetic genes and expression vectors. The restriction reactions were incubated at 37°C for 3 hours (h) followed by 80°C heat deactivation to terminate the enzyme digestion reaction. The products obtained after restriction digestion were resolved by electrophoresis on a 1.5% Tris-Boric-Acid Agarose gel extraction and purification using the DNeasy gel extraction kit (Qiagen) according to the manufacturer's instructions. DNA ligation was done using T4 (Thermo Scientific, USA). Vector and insert DNA were mixed in a molar ratio of 1:6 (restriction product: vector). The ligation process was carried out by incubation at 4°C overnight. The ligation mixtures were then used directly for bacterial transformation into chemically competent *E. coli* strain BL21 (DE3). Individual fresh 5 ml of LB broth containing 50 µg/ml of Ampicillin (LB-amp) was inoculated and cultured at 37°C (180 rpm) overnight. Then, all 5 ml of cultures were collected by centrifugation at 10 621 × g (10 min, 4°C) before being resuspended to a cell density of 1.0 of OD_{600} using LB-amp broth as diluent. After adjustment, 1 ml of uninduced culture was harvested by centrifugation at 15 294 × g for 1 min. The medium was discarded, and samples were stored at −20°C. Fresh LB-amp broth (5 ml) was inoculated with 1% (v/v) or 50 µl of overnight culture. When the cell density reached OD_{600} = 0.4 – 0.6, isopropyl β-D-1-thiogalactopyranoside (IPTG) (Amresco, USA) was added to a final concentration of 0.1 mM to induce recombinant protein expression at 12°C (180 rpm). After 24 hr of incubation, the cells were collected and resuspended with LB-amp broth to obtain a cell density of 1.0 at OD_{600} nm. Then, 1 ml of induced culture was transferred into a new 1.5 ml centrifuge tube and cells were collected at 15,294 × g for 1 min. All samples were stored at −20°C until further use. After that, the incubation temperature for induction was optimized. Three different temperatures were used (12°C, 20°C and 37°C) after the addition of IPTG.

Purification of Recombinant Proteins

The protein was extracted according to the modified method of Palmer and Wingfield 2012). Pelleted cells were thawed on ice and resuspended with 100 µl (10% of the volume harvested) PBS buffer. The cells were sonicated (30%, 20 sec, 1 on, 1 off) on ice. Sonicated samples were centrifuged at 15,294 × g for 1 min to separate insoluble and soluble fractions. After that, the supernatant was collected as a soluble fraction and an insoluble fraction was added with 100 µl PBS buffer for SDS-PAGE.

Identification of Recombinant Proteins

Targeted recombinant protein bands were excised and sent for peptide identification (Proteomics International Pty Ltd, Australia). The protein samples were trypsin digested and peptides extracted according to standard techniques (Bringans et al., 2008). Peptides were analysed by matrix-assisted laser desorption ionization time-of-flight (MALDI TOF-TOF) mass spectrometer using a 4800 Proteomics Analyzer (AB Sciex, USA). Spectra were analysed to identify the protein of interest using Mascot sequence matching software (Matrix Science, UK) with the Ludwig NR Database.

RESULTS AND DISCUSSION

Recombinant Gene Design

The codon modification process was aimed to reduce *E. coli* rare codons in the target sequence. Codon optimization results showed that the GI-MCP encoded a high number of codons that are rare in *E. coli*. The GI-MCP native sequence encoded 94 rare codons and 36 rare codons in the synthetic gene. The analysis of the native sequences showed that the target genes contain a high frequency of some specific codons which are particularly rare in *E. coli*. For example, 16 out of the 17 codons encoding for arginine present in the native sequence of GI-MCP were represented with rare codons with a frequency below 6% in the *E. coli* genome. All 16 of these codons were AGA and AGG, which are common arginine codons in eukaryotes and some viruses (Wu et al., 2014), yet very rare in *E. coli* genes. *E. coli* codon usage shows CGC, CGU and CGG as common arginine codons thus codon modification for IRI replaced AGA and AGG with CGC and CGU for better expression of GI-MCP in *E. coli* host cells. The modified gene encoding the GI-MCP is presented in Figure 1.

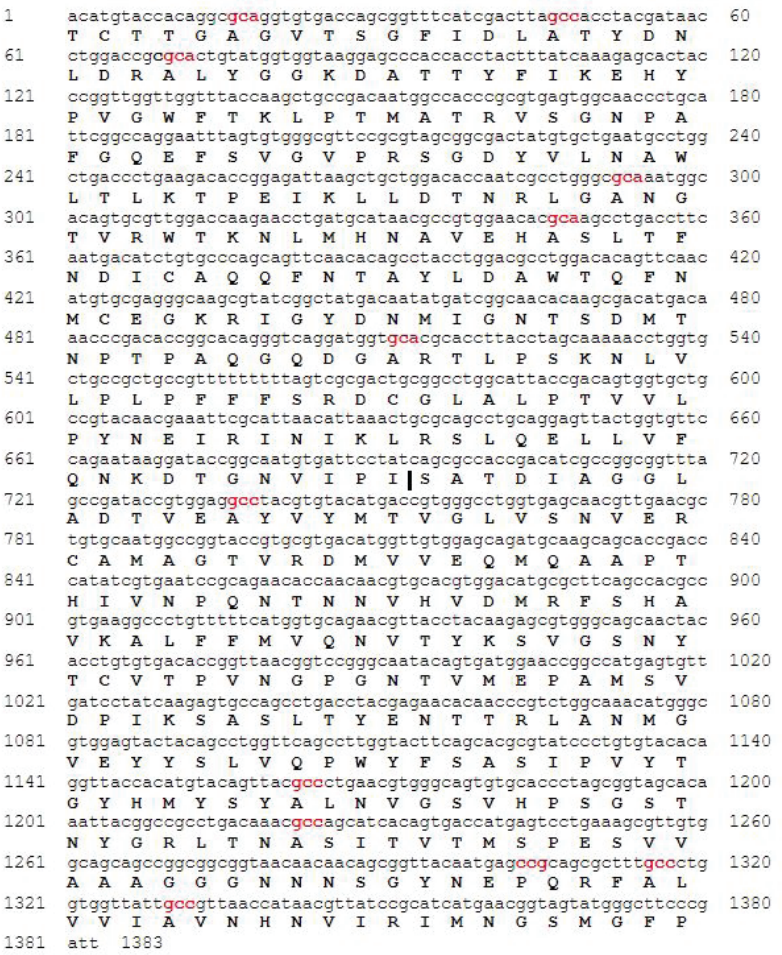

Figure 1 The nucleotide sequence of codon-modified GI-MCP and its deduced amino acid sequence. The amino acid sequence is in bold capital letters. The modified nucleotide sequence after codon optimization has been indicated in red.

Recombinant Protein Expression and Optimization

A protein band corresponding to ~56.8 kDa was detected in *E. coli* harbouring pET22b-GI-MCP, while no protein was found at the same position in *E. coli* harbouring the control pET22b plasmid (Figure 2). The molecular weight of the recombinant protein was identical to the predicted result. Sequencing and comparison with the Protein Database confirmed the identity of the protein (Table 1). The recombinant GI-MCP (rGI-MCP) was expressed mainly in inclusion body formation. In the soluble fraction, there was no band detected at the corresponding size of rGI-MCP, indicating that the protein was not present in a soluble form, in either the cytoplasmic or periplasmic regions. Accumulation of recombinant proteins as inclusion bodies was also reported by other studies involving recombinant proteins (Shimmoto et al., 2010; Kim et al.,

2007). However, this result contrasts with the soluble protein produced by Lin et al. (2014) who expressed the native sequence of recombinant GI-MCP that codes a slightly different amino acid sequence (AAV91066) and a different expression vector (pET23a). Apart from solubility, other studies expressing recombinant iridovirus capsid protein also reported molecular size discrepancy when expressed in *E. coli* and *Pichia pastoris* (*P. pastoris*) (Seo et al., 2013). A higher molecular size recombinant capsid protein of rock bream iridovirus (RBIV) (~59 kDa) was expressed in *P. pastoris* compared to the ~44 kDa recombinant capsid protein produced by *E. coli*. The size discrepancy was proven to be the result of glycosylation in *P. pastoris* which is nonetheless extremely rare in *E. coli*, thus expressing a higher molecular size product in *P. pastoris* (Sherlock et al., 2006).

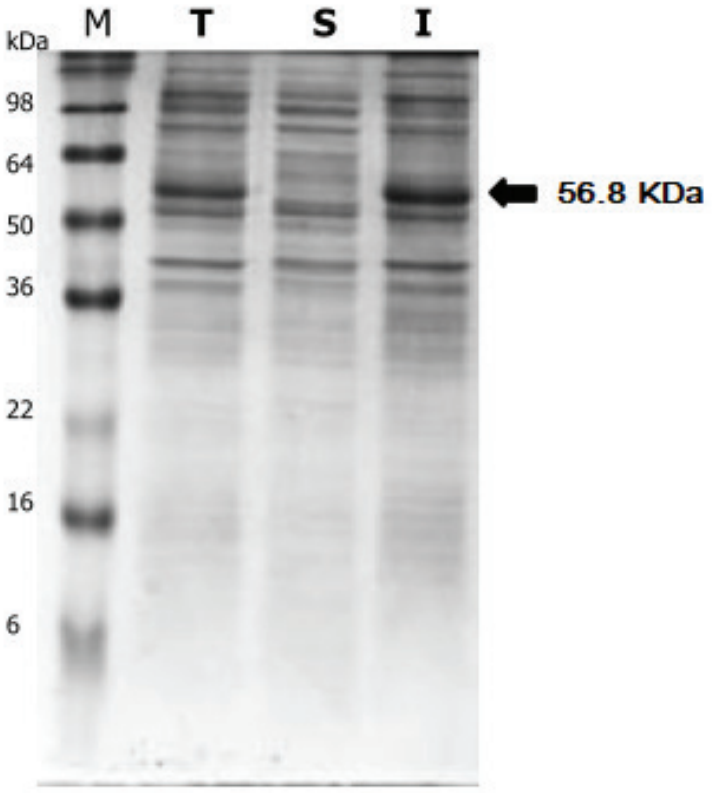

Figure 2 SDS-PAGE profile of *E. coli* BL21 (DE3) transformed with pET22b-GI-MCP. M, protein marker (LC5925) (Invitrogen, USA); (T) represents the total cell protein; (S) the soluble fraction and (I) the insoluble fraction which was recovered as inclusion bodies.

Table 1 Peptide fragments derived from the recombinant GI-MCP identified from MALDITOF-TOF MS are highlighted in yellow. The identity of fragments was confirmed by matching against the MASCOT database (MATRIX SCIENCE)

rGI-MCP	TCTTGAGVTSGFIDLATYDNLDRALYGGKDATTYFIKEHYPVGWFTKLPTMATRVSGNPAFGQEFS-VGVPRSGDYVLNAWLTLKTPEIKLLDTNRLGANGTVRWTKNLMHNAVEHASLTFNDICAQQFNTAY-LDAWTQFNMCEGKRIGYDNMIGNTSDMTNPTPAQGQDGARTLPSKNLVLPLPFFFSRDCGLALPTV-VLPYNEIRINIKLRSLQELLVFQNKDTGNVIPISATDIAGGLADTVEAYVYMTVGLVSNVERCAMAGT-VRDMVVEQMQAAPTHIVNPQNTNNVHVDMRFSHAVKALFFMVQNVTYKSVGSNYTCVTPVNG-PGNTVMEPAMSVDPIKSASLTYENTTRLANMGVEYYSLVQPWYFSASIPVYTGYHMYSYALNVGS-VHPSGSTNYGRLTNASITVTMSPESVVAAAGGGNNNSGYNEPQRFALVVIAVNHNVIRIMNGSMGFPI

The increase of protein degradation as temperature increases was also observed in rGI-MCP. As shown in Figure 3, rGI-MCP was expressed as inclusion bodies at all three different incubation temperatures (12°C, 20°C and 37°C). At 37°C, two smaller bands (denoted by blue arrows) were observed in the inclusion bodies. Similar bands were undetectable in the control of the culture of 12°C and 20°C. Therefore, the two smaller bands were postulated to be the result of protein expressed by *E. coli* under specific temperatures in the presence of the plasmid.

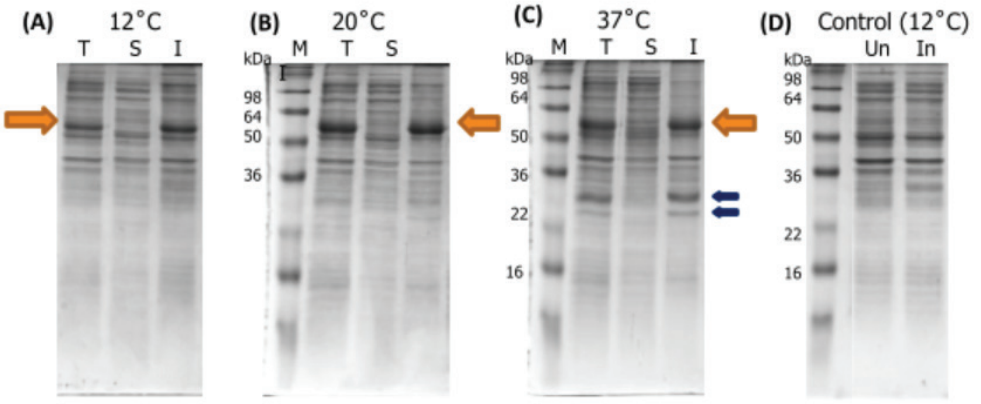

Figure 3 Expression of rGI-MCP at different temperatures in *E. coli* BL21 (DE3) after induction with 0.1 mM IPTG. A, expression of rGI-MCP at 12°C for 24 hours; B, expression of rGI-MCP at 20°C for 12 hours; C, expression of rGI-MCP at 37°C for 12 hours. M, protein marker (LC5925) (Invitrogen, USA); T, total cell protein; S, soluble fraction; I, insoluble fraction. Control, pET22b-transformed *E. coli* BL21 (DE3) as negative control where total cell protein of uninduced cells (Un) and total cell protein of induced cells (In) were used. The orange-coloured arrow indicates the position of the recombinant protein and the two blue arrows indicate the presence of unknown proteins which were detected in the culture which was induced at 37°C and 20°C

CONCLUSIONS

The presence of rare codons within viral native genes represents a challenge to their downstream production for application as vaccines. This study successfully expressed the synthetic version of the recalcitrant GI-MCP by undertaking specific modifications in the coding region and successfully expressed the recombinant protein in *E. coli* (BL21) (DE3). The recombinant protein was expressed as inclusion bodies which can be subjected to linearization and refolding for downstream applications as antigens for vaccine production and diagnostics.

PATENTS

A patent for the DNA sequence was filed at the Malaysian Intellectual Property Organization (MyIPO). The title of the patent is "A Polynucleotide Sequence Coding an Iridovirus Polypeptide and a Use Thereof". **Application No:** PI 2013701136

ACKNOWLEDGEMENTS

This study was funded by the research grant SCF00074 (Agridana, Ministry of Agriculture, Government of Malaysia).

REFERENCES

Bringans, S., Eriksen, S., Kendrick, T., Gopalakrishnakone, P., Livk, A., Lock, R., & Lipscombe, R. (2008). Proteomic analysis of the venom of *Heterometrus longimanus* (Asian black scorpion). *Proteomics*, *8* (5), 1081 – 1096. https://doi.org/10.1002/pmic.200700948

Burgess-Brown, N. A., Sharma, S., Sobott, F., Loenarz, C., Oppermann, U., & Gileadi, O. (2008). Codon optimization can improve the expression of human genes in *Escherichia coli*: A multi-gene study. *Protein Expression and Purification*, *59* (1), 94 – 102. https://doi.org/10.1016/j.pep.2008.01.008

Hwang, J. Y., Kwon, M. G., Seo, J. S., Hwang, S. D., Jeong, J. M., Lee, J. H., & Jee, B. Y. (2020). Current use and management of commercial fish vaccines in Korea. *Fish and Shellfish Immunology*, *102*, 20 – 27. https://doi.org/10.1016/j.fsi.2020.04.004

Kim, M., Qiao, Z., Yu, J., Montefiori, D., & Reinherz, E. L. (2007). Immunogenicity of recombinant human immunodeficiency virus type 1-like particles expressing gp41 derivatives in a pre-fusion state. *Vaccine*, *25* (27), 5102 – 5114. https://doi.org/10.1016/j.vaccine.2006.09.071

Li, P., Yan, Y., Wei, S., Wei, J., Gao, R., Huang, X., & Qin, Q. (2014). Isolation and characterization of a new class of DNA aptamers specific binding to Singapore grouper iridovirus (SGIV) with antiviral activities. *Virus Research*, *188*, 146 – 154. https://doi.org/10.1016/j.virusres.2014.04.010

Lin, H. Y., Liou, C. J., Cheng, Y. H., Hsu, H. C., Yiu, J. C., Chiou, P. P., & Lai, Y. S. (2014). Development and application of a monoclonal antibody against grouper iridovirus (GIV) major capsid protein. *Journal of Virological Methods*, *205*, 31 – 37. https://doi.org/10.1016/j.jviromet.2014.04.013

Liu, H. I., Chiou, P. P., Gong, H. Y., & Chou, H. Y. (2015). Cloning of the major capsid protein (MCP) of grouper iridovirus of Taiwan (TGIV) and preliminary evaluation of a recombinant MCP vaccine against TGIV. *International Journal of Molecular Sciences*, *16* (12), 28647 – 28656. https://doi.org/10.3390/ijms161226118

Liu, G. Y., Wang, E. L., Qu, X. Y., Yang, K. C., Zhang, Z. Y., Liu, J. Y., & Wang, G. X. (2020). Single-walled carbon nanotubes enhance the immune protective effect of a bath subunit vaccine for pearl gentian grouper against Iridovirus of Taiwan. *Fish and Shellfish Immunology*, *106*, 510 – 517. https://doi.org/10.1016/j.fsi.2020.08.003

Matsuura, Y., Terashima, S., Takano, T., & Matsuyama, T. (2019). Status of fish vaccines in Japan. *Fish and Shellfish Immunology*, *95*, 236 – 247. https://doi.org/10.1016/j.fsi.2019.09.031

Menzella, H. G. (2011). Comparison of two codon optimization strategies to enhance recombinant protein production in Escherichia coli. *Microbial Cell Factories*, *10* (1), 1 – 8. https://doi.org/10.1186/1475-2859-10-15

Munang'andu, H. M., & Evensen, Ø. (2019). Correlates of protective immunity for fish vaccines. *Fish and Shellfish Immunology*, *85*, 132 – 140. https://doi.org/10.1016/j.fsi.2018.03.060

Ou-Yang, Z., Wang, P., Huang, Y., Huang, X., Wan, Q., Zhou, S., & Qin, Q. (2012). Selection and identification of Singapore grouper iridovirus vaccine candidate antigens using bioinformatics and DNA vaccination. *Veterinary Immunology and Immunopathology*, *149* (1 – 2), 38 – 45. https://doi.org/10.1016/j.vetimm.2012.05.021

Palmer, I., & Wingfield, P. T. (2012). Preparation and extraction of insoluble (inclusion-body) proteins from *Escherichia coli*. *Current Protocols in Protein Science*, *70* (1), 6 – 3. https://doi.org/10.1002/0471140864.ps0603s70

Puigbo, P., Guzman, E., Romeu, A., & Garcia-Vallve, S. (2007). OPTIMIZER: A web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Research*, *35* (Suppl 2), W126 – W131. https://doi.org/10.1093/nar/gkm219

Qin, Q. W., Lam, T. J., Sin, Y. M., Shen, H., Chang, S. F., Ngoh, G. H., & Chen, C. L. (2001). Electron microscopic observations of a marine fish iridovirus isolated from brown-spotted grouper, *Epinephelus tauvina*. *Journal of Virological Methods*, *98* (1), 17 – 24. https://doi.org/10.1016/S0166-0934(01)00350-0

Qin, Q. W., Shi, C., Gin, K. Y., & Lam, T. J. (2002). Antigenic characterization of a marine fish iridovirus from grouper, *Epinephelus* spp. *Journal of Virological Methods*, *106* (1), 89 – 96. https://doi.org/10.1016/S0166-0934(02)00139-8

Sherlock, O., Dobrindt, U., Jensen, J. B., Vejborg, R. M., & Klemm, P. (2006). Glycosylation of the self-recognizing *Escherichia coli* Ag43 autotransporter protein. *Journal of Bacteriology*, *188* (5), 1798 – 1807. https://doi.org/10.1128/JB.188.5.1798-1807.2006

Seo, J. Y., Chung, H. J., & Kim, T. J. (2013). Codon-optimized expression of fish iridovirus capsid protein in yeast and its application as an oral vaccine candidate. *Journal of Fish Diseases*, *36* (9), 763 – 768. https://doi.org/10.1111/jfd.12037

Shimmoto, H., Kawai, K., Ikawa, T., & Oshima, S. I. (2010). Protection of red sea bream *Pagrus major* against red sea bream iridovirus infection by vaccination with a recombinant viral protein. *Microbiology and Immunology*, *54* (3), 135 – 142. https://doi.org/10.1111/j.1348-0421.2010.00204.x

Tsai, C. T., Ting, J. W., Wu, M. H., Wu, M. F., Guo, C., & Chang, C. Y. (2005). Complete genome sequence of the grouper iridovirus and comparison of genomic organization with those of other iridoviruses. *Journal of Virology*, *79* (4), 2010 – 2023. https://doi.org/10.1128/JVI.79.4.2010-2023.2005

Wang, Q., Ji, W., & Xu, Z. (2020). Current use and development of fish vaccines in China. *Fish and Shellfish Immunology*, *96*, 223 – 234. https://doi.org/10.1016/j.fsi.2019.12.010

Wei, J., Huang, Y., Zhu, W., Li, C., Huang, X., & Qin, Q. (2019). Isolation and identification of Singapore grouper iridovirus Hainan strain (SGIV-HN) in China. *Archives of Virology*, *164* (7), 1869 – 1872. https://doi.org/10.1007/s00705-019-04268-z

Wu, X., Jörnvall, H., Berndt, K. D., & Oppermann, U. (2004). Codon optimization reveals critical factors for high level expression of two rare codon genes in Escherichia coli: RNA stability and secondary structure but not tRNA abundance. *Biochemical and Biophysical Research Communications*, *313* (1), 89 – 96. https://doi.org/10.1016/j.bbrc.2003.11.091