

APPENDIX – SEQUENCE LISTINGS

SEQ ID NO:1 (Native mature OgLuc)

FTLADFGVDWQQTAGYNQDQVLEQQGGLSSLFQALGVSVTPIQKVVLSENGLKADIHVIIIPYEGLSGFQMGLIEM
IFKVVPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYPGIAVFDGKQITVTGTLWNGNKIYDERLINPDGSLL
FRVTINGVTGWRLCENILA

SEQ ID NO:2 (OgLuc nt sequence) Synthetic wt nucleotide seq

atgggtgtttaccttggcagatcttcgttgagactggcaacagacagctggatacaaccaagatcaagtgttagaaca
aggaggattgtctagctctgttccaagccctgggagtgtcagtcacccaatccagaaagttgtgctgtctggggaga
atgggttaaaagctgatattcatgtcatcatcccttacgagggactcagtggttttcaaagtggtctgattgaaatg
atcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgttattgacgg
tgtgacaccaaacatgattgactactttggacgcccttacccctggaattgctgtgtttgacggcaagcagatcacag
ttactggaactctgtggaacggcaacaagatctatgatgagcgccctgatcaaccagatggttcactcctcttccgc
gttactatcaatggagtaccggatggcgccctttgcgagaacattcttgcc

SEQ ID NO: 3 (1PMP)

SNKFLGTWKLVSSENFDEYMKALGVGLATRKLGNLAKPRVVISKKGDIITIRTESPFKNTEISFKLGQEFEEETTA
DNRKTKSTVTTLARGSLNQVQKWNGNETTIKRKLVDGKMMVECKMKDVVCTRIYEKV

SEQ ID NO:4 (1CRB)

PVDFNGYWKMLSNENFEEYLRALDVNVALRKIANLLKPDKEIVQDGDHMIIRTLSTFRNYIMDFQVGKEFEEDLT
GIDDRKCMTTVSWDGDKLQCVQKGEKEGRGWTQWIEGDELHLEMRAEGVTCKQVFKKVH

SEQ ID NO:5 (1VYF)

GSMSSFLGKWKLSSESHNFDAVMSKLGVSATRQIGNTVTPTVTFTMDGDKMTMLTESTFKNLSCTFKFGEEFDEK
TSDGRNVKSVVEKNSESKLTQTQVDPKNTTVIVREVDGDTMKTTVTVGDTVTAIRNYKRLS

SEQ ID NO:6 (1VYF)

GSMSSFLGKWKLSSESHNFDAVMSKLGVSATRQIGNTVTPTVTFTMDGDKMTMLTESTFKNLSCTFKFGEEFDEK
TSDGRNVKSVVEKNSESKLTQTQVDPKNTTVIVREVDGDTMKTTVTVGDTVTAIRNYKRLS

SEQ ID NO:7 (OgLuc with residues 1-2 and 168-169 omitted)

LADFGVDWQQTAGYNQDQVLEQQGGLSSLFQALGVSVTPIQKVVLSENGLKadihviIPYEGLSGFQMgliEMIF
KVvypvddhhfkiilhygtlvidGVTPNMIDYFGRPYPGIAVFDGKQITVTGTLWNGNKIYDERLINPDGSLLFR
VTINGVTGWRLCENI

SEQ ID NO:8 (1PMP)

SNKFLGTWKLVSSENFDEYMKALGVGLATRKLGNLAKPRVVISKKGDIITIRTESPFKNTEISFKLGQEFEEETTA
DNRKTKSTVTTLARGSLNQVQKWNGNETTIKRKLVDGKMMVECKMKDVVCTRIYEKV

SEQ ID NO:9 (1CRB)

PVDFNGYWKMLSNENFEEYLRALDVNVALRKIANLLKPDKEIVQDGDHMIIRTLSTFRNYIMDFQVGKEFEEDLT
GIDDRKCMTTVSWDGDKLQCVQKGEKEGRGWTQWIEGDELHLEMRAEGVTCKQVFKKVH

SEQ ID NO: 10 (Native protein seq (with signal sequence))

MAYSTLFIIALTAVVTQASSTQKSNLTFTLADFGVDWQQTAGYNQDQVLEQQGGLSSLFQALGVSVTPIQKVVLSE
ENGLKADIHVIIIPYEGLSGFQMGLIEMIFKVVPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYPGIAVFDGK
QITVTGTLWNGNKIYDERLINPDGSLLFRVTINGVTGWRLCENILA

SEQ ID NO: 11 (1PMP;):

SNKFLGTWKLVSSENFDEYMKALGVGLATRKLGNLAKPRVVISKKGDIITIRTESPFKNTEISFKLGQEFEEETTA
DNRKTKSTVTTLARGSLNQVQKWNGNETTIKRKLVDGKMMVECKMKDVVCTRIYEKV

SEQ ID NO: 12 (True wt nucleotide seq (with signal peptide))

atggcgtaactccactctgttcataattgcattgaccgccgttgctactcaagcttcctcaactcaaaaatctaac
ctaacttttacgttggcagatcttcgttgagactggcaacagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgtctagctctgttccaagccctgggagtgtcagtcacgcccatacagaaagttgtactgtctggg
gagaatgggttaaaagctgatattcatgtcataataccttacgagggactcagtggttttcaaagtggtcctaatt
gaaatgatcttcaaagttgtttaccccgtagatgatcatcatttcaagattattctccattatggtacactcgtt
attgacggtgtaacaccaaacatgattgactactttggaagacccttacccctggaattgctgtatttgacggcaag

cagatcacagttactggaactctgtggaacggcaacaagatctatgatgagaggctaataaccctgatggttca
ctcctcttcagagttactatcaatggagtcacgggatggaggctttgcgagaacattcttgcc

SEQ ID NO: 13 (Synthetic wt protein seq)

MVFTLADFGVDWQQTAGYNQDQVLEQGGGLSSLFQALGVSVTPIQKVVLSGENGLKADIHV IIPYEGLSGFQMGLI
EMIFKVVYPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYPGIAVFDGKQITVTGTLWNGNKIYDERLINPDGS
LLFRVTINGVTGWRLCENILA

SEQ ID NO: 14 (N166R nucleotide seq)

atggtggtttacattggcagatcttgcgttgagactggcagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgtctagtctgttccaaagcctgggagtgctcagtcaccccaatccagaaaattgtgctgtctggg
gagaatgggtttaaagctgatattcatgtcatcatcccttacgagggactcagtggttttcaaattgggtctgatt
gaaatgatcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgtt
attgacggtgtgacaccaaacatgattgactactttggacgcccttacccctggaattgctgtgtttgacggcaag
cagatcacagttactggaactctgtggaacggcaacaagatctatgatgagcgctgatcaaccagatggttca
ctcctcttccgcgttactatcaatggagtcacccggatggcgccctttgcgagcgtattcttgcc

SEQ ID NO: 15 (N166R protein seq)

MVFTLADFGVDWQQTAGYNQDQVLEQGGGLSSLFQALGVSVTPIQKVVLSGENGLKADIHV IIPYEGLSGFQMGLI
EMIFKVVYPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYPGIAVFDGKQITVTGTLWNGNKIYDERLINPDGS
LLFRVTINGVTGWRLCERILA

SEQ ID NO: 16 (C1+A4E nucleotide seq)

atggtggtttacattggaggatcttgcgttgagactggcggcagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgtctagtctgttccaaaagctgggagtgctcagtcaccccaatccagaaaattgtgctgtctggg
gagaatgggtttaaatttgatattcatgtcatcatcccttacgagggactcagtggttttcaaattgggtctgatt
gaaatgatcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgtt
attgacggtgtgacaccaaacatgattgactactttggacgcccttacgagggaattgctgtgtttgacggcaag
aagatcacagttactggaactctgtggaacggcaacaagatcattgatgagcgctgatcaaccagatggttca
ctcctcttccgcgttactatcaatggagtcacccggatggcgccctttgcgagcgtattcttgcc

SEQ ID NO: 23 (C1+A4E protein seq)

MVFTLEDFVGDWRQTAGYNQDQVLEQGGGLSSLFQKLGVSVTPIQKIVLSGENGLKF DIHV IIPYEGLSGFQMGLI
EMIFKVVYPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKI I DERLINPDGS
LLFRVTINGVTGWRLCERILA

SEQ ID NO: 24 (C1 nucleotide seq)

atggtggtttacattggcagatcttgcgttgagactggcggcagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgtctagtctgttccaaaagctgggagtgctcagtcaccccaatccagaaaattgtgctgtctggg
gagaatgggtttaaatttgatattcatgtcatcatcccttacgagggactcagtggttttcaaattgggtctgatt
gaaatgatcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgtt
attgacggtgtgacaccaaacatgattgactactttggacgcccttacgagggaattgctgtgtttgacggcaag
aagatcacagttactggaactctgtggaacggcaacaagatcattgatgagcgctgatcaaccagatggttca
ctcctcttccgcgttactatcaatggagtcacccggatggcgccctttgcgagcgtattcttgcc

SEQ ID NO: 25 (C1 protein seq)

MVFTLADFGVDWRQTAGYNQDQVLEQGGGLSSLFQKLGVSVTPIQKIVLSGENGLKF DIHV IIPYEGLSGFQMGLI
EMIFKVVYPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKI I DERLINPDGS
LLFRVTINGVTGWRLCERILA

SEQ ID NO: 26 (C1+C2 nucleotide seq)

atggtggtttacattggcagatcttgcgttgagactggcggcagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgtctagtctgttccaaaagctgggagtgctcagtcaccccaatccagaaaattgagctgtctggg
gagaatgggtttaaatttgatattcatgtcatcatcccttacgagggactcagtggttttcaaattgggtctgatt
gaaatgatcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgtt
attgacggtgtgacacctaatacatgattgactactttggacgcccttacgagggaattgctgtgtttgacggcaag
aagatcacagttactggaactctgtggaacggcaagaagatcattgagagcgctgatcaaccagatggttca
ctcctcttccgcgttactatcaatggagtcacccggatggcgccctttgcgagcgtgttcttgcc

SEQ ID NO: 27 (C1+C2 protein seq)

MVFTLADFVGDWQRQTAGYNQDQVLEQGGLSSLFQKLGVSVTPIQKIELSGENGLKFEDIHVIIIPYEGLSGFQMGLI
EMIFKVVYPVDDHHFKIILHYGTLVIDGVTNLNMIDYFGRPYEGIAVFDGKKITVTGTLWNGKKIIERLINPDGS
LLFRVTINGVTGWRLCERVLA

SEQ ID NO: 28 (C1+C2+A4E nucleotide seq)

atggtggtttacattggaggatttcggttgagactggcggcagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgtctagtctgttccaaaagctgggagtgtcagtcaccccaatccagaaaattgagctgtctggg
gagaatgggttataaaatttgatattcatgtcatcatcccttacgagggactcagtggttttcaaattgggtctgatt
gaaatgatcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgtt
attgacggtgtgacacttaacatgattgactactttggacgcccttacgagggaattgctgtgtttgacggcaag
aagatcacagttactggaactctgtggaacggcaagaagatcattgaggagcgctgatcaaccagatggttca
ctctcttccgcgttactatcaatggagtcaccggatggcgccctttgcgagcgtgttcttgcc

SEQ ID NO: 29 (C1+C2+A4E protein seq)

MVFTLEDFVGDWQRQTAGYNQDQVLEQGGLSSLFQKLGVSVTPIQKIELSGENGLKFEDIHVIIIPYEGLSGFQMGLI
EMIFKVVYPVDDHHFKIILHYGTLVIDGVTNLNMIDYFGRPYEGIAVFDGKKITVTGTLWNGKKIIERLINPDGS
LLFRVTINGVTGWRLCERVLA

SEQ ID NO: 30 (C1+C3+A4E nucleotide seq)

atggtggtttacattggaggatttcggttgagactggcggcagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgcctagtctgttccaaaagatgggagtgtcagtcaccccaatccagaaaattgtgctgtctggg
gagaatgtgttataaaatttgatattcatgtcatcatcccttacgagggactcagtggttttcaaattgggtctgatt
gaaatgatcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgtt
gttgacggtgtgacaccaaacatgattgactactttggacgcccttacgagggaattgctgtgtttgacggcaag
aagatcacagttactggaactctgtggaacggcaacaagatcattgatgagcgctgtctcaaccagatggttca
ctctcttccgcgttactatcaatggagtcaccggatggcgccctttgcgagcgtattcttgcc

SEQ ID NO: 31 (C1+C3+A4E protein seq)

MVFTLEDFVGDWQRQTAGYNQDQVLEQGGLPSLFQKMGVSVTPIQKIVLSGENVLKFEDIHVIIIPYEGLSGFQMGLI
EMIFKVVYPVDDHHFKIILHYGTLVVDGVTNPMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLLNPDGS
LLFRVTINGVTGWRLCERILA

SEQ ID NO: 32 (T2T nt sequence)

atggtggtttacattggcagatttcggttgagactggcaacagacagctggatacaaccaagatcaagtgttagaa
caaggaggattgtctagtctgttccaagccctgggagtgtcagtcaccccaatccagaaaattgtgctgtctggg
gagaatgggttataaaagctgatattcatgtcatcatcccttacgagggactcagtggttttcaaattgggtctgatt
gaaatgatcttcaaagttgtttaccagtggtgatcatcatttcaagattattctccattatggtacactcgtt
attgacggtgtgacaccaaacatgattgactactttggacgcccttaccctggaattgctgtgtttgacggcaag
cagatcacagttactggaactctgtggaacggcaacaagatctatgatgagcgctgatcaaccagatggttca
ctctcttccgcgttactatcaatggagtcaccggatggcgccctttgcgagaacattcttgcc