

Cyanobacterial production of biopharmaceutical and biotherapeutic proteins

Nico Betterle, Diego Hidalgo Martinez, Anastasios Melis*

Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102, USA

Supplemental and Supporting Information

* Correspondence to: Anastasios Melis, Department of Plant and Microbial Biology, University of California, 111 Koshland Hall, Berkeley, CA 94720-3102, USA; Telephone: 510-642-8166; Fax: 510-642-4995; E-mail: melis@berkeley.edu

Table S1 Sequence of oligonucleotide primers used in the present work.

Oligos name	Oligos DNA sequence
< <i>cpc-us for</i> >	5'- CCATTAGCAAGGCAAATCAAAGAC -3'
< <i>cpcA rev</i> >	5'- GGTGGAAACGGCTTCAGTTAAAG -3'

Table S2 Nucleotide GenBank accession and protein reference numbers of the original DNA sequences and proteins used in this work. The referenced GenBank nucleotide sequences are those prior to codon-use optimization for *Synechocystis* expression. The codon-optimized nucleotide sequences, as expressed in *Synechocystis* for the purposes of this work, and the nucleotide sequences of the full constructs that were synthesized and employed here are shown further below, on pages 3-8 in this Supplemental and Supporting Information file.

Gene name	GenBank Accession #	Protein Reference #
<i>IFN-α2 (H. sapiens)</i>	NM_000605.4	NP_000596.2
<i>cpcB (Synechocystis sp PCC 6803)</i>	NC_020286.1	WP_010871861.1
<i>cpcA (Synechocystis sp PCC 6803)</i>	NC_000911.1	WP_010871860.1
<i>cpcG1 (Synechocystis sp PCC 6803)</i>	NC_007775.1	WP_011429130.1
<i>nptI (E. coli)</i>	NC_012886.1	YP_002995708.1
<i>cmR (E. coli)</i>	NC_005923.1	YP_025721.1

Diagram illustrating the structure of the *cpc* operon. The operon consists of the following genes and their orientations:

- cpcUS* (transcribed left to right)
- IFN* (transcribed left to right)
- cmR* (transcribed left to right)
- cpcA* (transcribed right to left)
- cpcC1* (transcribed left to right)
- cpcC2* (transcribed left to right)
- cpcD* (transcribed left to right)
- cpcDS* (transcribed left to right)

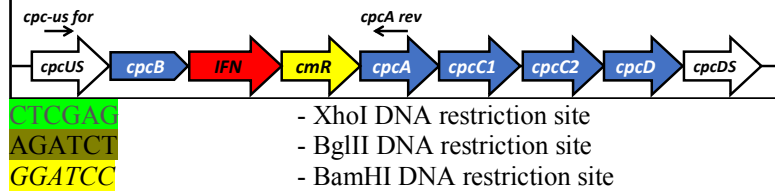
Restriction sites are indicated below the genes:

- XhoI DNA restriction site (TCGAG) is located between *cpcUS* and *IFN*.
- BglII DNA restriction site (AGATCT) is located between *IFN* and *cmR*.
- BamHI DNA restriction site (GGATCC) is located between *cmR* and *cpcA*.

UPPER CASE - codon-optimized human interferon gene for expression in *Synechocystis*
lower case - intergenic sequence
UPPER CASE BOLD - *cmR* (chloramphenicol selection cassette)
lower case - Transcription terminator

CTTTGACCTCTCCCCAGCTGGTATGTTCTCGAG

Construct used for the *cpcB-IFN-cmR* sequence insertion within the *cpc* operon (*IFN*, codon-optimized human interferon gene for expression in *Synechocystis*), between *cpcB* and *cpcA* genes.



Sequence for 5'- homologous recombination

UPPER CASE - partial *cpcB*
UPPER CASE - *cpcB-cpcA* intergenic sequence

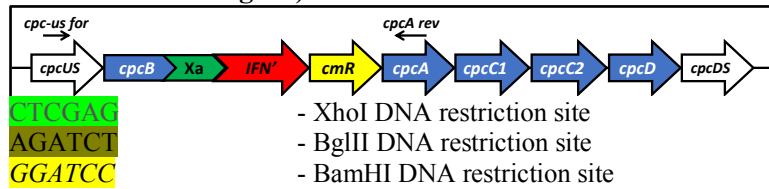
UPPER CASE - codon-optimized human interferon gene for expression in *Synechocystis*
lower case - intergenic sequence
UPPER CASE BOLD - *cmR* (chloramphenicol selection cassette)
lower case - Transcription terminator

Sequence for 3'- homologous recombination

UPPER CASE ITALICS - *cpcB-cpcA* intergenic sequence
UPPER CASE BOLD - *cpcA* (partial)

CTCGAG CCGCATCACCGGTAATGCTTCCGCTATCGTTTCCAACGCTGCTCGTGCTTTGTTTCGCCGAACAGCCCCAAT
TAATCCAACCCGGTGAAACGCCTACACCAGCCGTCGTATGGCTGCTTGTTCGCTGACATGGAAATCATCTCCGC
TATGTTACCTACGCAACCTTACCCGGCGACGCTTCCGTTCTAGAAGATCGTTGCTTGAACGGTCTCCGTGAAACCTAC
GTTGCCCTGGGTGTTCCCGGTGCTTCCGTAGCTGCTGGCGTTCAAAAAATGAAAGAAGCTGCCCTGGACATCGTTAA
CGATCCCAATGGCATCACCCGTGGTGATTGCAGTGCTATCGTTGCTGAAATCGCTGGTTACTTCGACCGCGCCGCTG
CTGCCGTAGCCTAGTCTGGTTATTTTAAAAACCACTTTACTCAGGTTCCATACCCGAGAAAAATCCAGCTT
AAAGCTGACATATCTAGGAAAAATTTTACATTTTACCGGGAGATACCCAGAAATGTGTGACTTGCCTCA
GACGCATTCTTTGGGAAGCCGACGCACACTGATGCTGCTCGCCCAAATGCGCCGGATCTCCTTATTCTCCTGT
CTCAAGGATCGGCATGACTTCGGCTTCCCTCAGGAGGAGTTTGGAATCAGTTCCAAAAGGCCGAAACCATT
CCGGTCTCCATGAAATGATTCAACAGATCTTTAACTTATTTCAGTACCAAAGACAGCAGTGCGGCCTGGGACG
AAACATTACTCGATAAATTCTACACGGAATTATACCAACAGTTGAACGACTTAGAAGCCTGTGTAATCCAAG
GTGTTGGTGTCACTGAGACTCCATTAATGAAAGAAGACTCTATTCTGGCCGTCCGCAAGTATTTCAGCGAAT
CACACTGTATTTGAAAGAGAAAAAGTATTCTCCGTGTGCGTGGGAGGTAGTACGGGCTGAAATCATGCGGTC
CTTCTCTTTAAGCACAAACCTCCAGGAATCTCTGCGCTCCAAAGAATGAAGATCTgcggccgcgtgatcgccacgtaagagg
ttccaacttcaccataatgaaataagatcactaccggcgctatttttgagttatcgagatttcaggagctaaggaagctaaaATGGAGAAAAAATCACTGGA
TATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGTTGCTCAA
TGTACCTATAACCAGACCGTTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAAATAAGCAC
AAGTTTTATCCGGCCTTTATTACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCGGTATGGCA
ATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTACACCGTTTCCATGAGCAAACCT
GAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTTCGCAA
GATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTC
TCAGCCAATCCCTGGGTGAGTTTACCAGTTTGTATTAAACGTGGCCAATATGGACAACCTTCTTCGCC
CCCGTTTTACCATGGGCAAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTACAGGT
TCATCATGCCGTCTGTGATGGCTTCCATGTGCGGAGAATGCTTAATGAATTACAACAGTACTGCGATGA
GTGGCAGGGCGGGGCGTAAATTTTaaaggcagttattggtgccctaaacgcctggGGATCCTCTGGTTATTTTAAAAACCACTTTAC
TCAGGTTCCATACCCGAGAAAAATCCAGCTTAAAGCTGACATATCTAGGAAAAATTTTACATTCTAACGGGAGATACCAG
AACATGAAAACCCCTTTAACTGAAGCCGTTTCCACCGCTGACTCTCAAGGTCGTTTCTGAGCAGCAC
CGAATTGCAAAATTGCTTTCGGTCGTCTACGTCAAGCTAATGCTGGTTTGCAAGCCGCTAAAGCTCTGAC
CGACAATGCCAGAGCTTGGTAAATGGTGCTGCCCAAGCCGTTTATAACAAATTCCCTACACCACCA
AACCCAAGGCAACAACCTTGTGCTGCGGATCAACGGGGTAAAGACAAGTGTGCCCGGACATCGGCTACT
ACCTCCGCATCGTTACCTACTGCTTAGTTGCTGGTGGTACCGGTCCTTTGGATGAGTACTTGATCGCCG
GTATTGATGAAATCAACCGCACCTTTGACCTCTCCCCCAGCTGGTATGTTCTCGAG

Construct used for the *cpcB*IFN'-cmR* sequence insertion within the *cpc* operon (*IFN'*, not-optimized human interferon gene).



Sequence for 5'- homologous recombination

UPPER CASE ITALICS - *cpcB* (without stop codon)

UPPER CASE - Xa factor cleavage site (IEGR)
UPPER CASE - Native human interferon
 lower case - intergenic sequence
UPPER CASE BOLD - *cmR* (chloramphenicol selection cassette)
 lower case - Transcription terminator

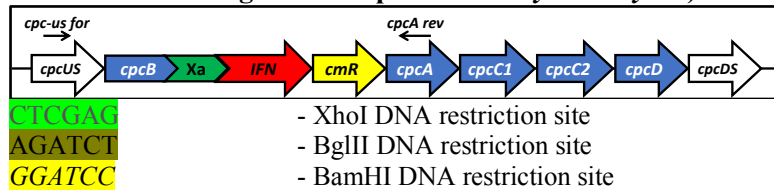
Sequence for 3'- homologous recombination

UPPER CASE ITALICS - *cpcB-cpcA* intergenic sequence

UPPER CASE BOLD - *cpcA* (partial)

CTCGAGATGTTTCGACGTATTCCTCGGGTGTGTTCCCAAGCTGATGCTCGCGGCGAGTACCTCTCTGGTTCTCAGTTA
 GATGCTTTGAGCGCTACCGTTGCTGAAGGCAACAAACGGATTGATTCTGTAAACCGCATCACCGGTAATGCTTCCGCT
 ATCGTTTCCAACGCTGCTCGTGCTTTGTTGCGCCGAACAGCCCCAATTAATCCAACCCGGTGGAACGCCTACACCAG
 CCGTCGTATGGCTGCTTGTGCTGACATGGAATCATCCTCCGCTATGTTACCTACGCAACCTTCACCGGCGACG
 CTTCCGTTCTAGAAGATCGTTGCTTGAACGGTCTCCGTGAAACCTACGTTGCCCTGGGTGTTCCCGGTGCTTCCGTA
 GCTGCTGGCGTTCAAAAAATGAAAGAAGCTGCCCTGGACATCGTTAACGATCCAATGGCATCACCGTGCTGATTG
 CAGTGCTATCGTTGCTGAAATCGCTGGTACTTCGACCGCGCCGCTGCTGCCGTAGCCATCGAAGGGCGATGTGA
 TCTGCCTCAAACCCACAGCCTGGGTAGCAGGAGGACCTTGATGCTCCTGGCACAGATGAGGAGAATCTCTCTT
 TTCTCCTGCTTGAAGGACAGACATGACTTTGGATTTCGCCAGGAGGAGTTTGGAACCCAGTTCCAAAAGGCTG
 AAACCATCCCTGTCTCCATGAGATGATCCAGCAGATCTTCAATCTCTTCAGCACAAAGGACTCATCTGCTGC
 TTGGGATGAGACCCCTCCTAGACAAATTCTACACTGAACCTTACCAGCAGCTGAATGACCTGGAAGCCTGTGTG
 ATACAGGGGGTGGGGGTGACAGAGACTCCCCTGATGAAGGAGGACTCCATTCTGGCTGTGAGGAAATACTTC
 CAAAGAATCACTCTCTATCTGAAAGAGAAGAAATACAGCCCCTGTGCTGGGAGGTTGTGAGAGCAGAAATC
 ATGAGATCTTTTCTTTGTCAACAAACTTGCAAGAAAGTTAAGAAGTAAGGAATGAAGATCTgcggccgctgac
 ggcacgtgaagaggttccaaacttcaccataatgaaataagatcaactaccggcgctattttgagttatcgagatttcaggagctaaggagctaaaATGGAGAAAAAA
 TCACTGGATATAACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAG
 TTGCTCAATGTACCTATAACCAGACCGTTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAA
 ATAAGCACAAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCC
 GTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTACCCCTTGTTACACCGTTTTCCATG
 AGCAAACTGAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATAT
 ATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGT
 TTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAAC
 TCTTCGCCCCCGTTTTTACCACATGGGCAAAATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCG
 ATTCAGGTTTCATCATGCGCTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTAC
 TGCGATGAGTGGCAGGGCGGGGCGTAAATTTtaaggcagttattggtgccttaaacgcctggGGATCCTCTGGTTATTTAAAAAC
 CAACTTTACTCAGGTTCCATACCCGAGAAAAATCCAGCTTAAAGCTGACATATCTAGGAAAAATTTTACATTCTAACGGG
 AGATACCAGAAACAATGAAAACCCCTTTAACTGAAGCCGTTTCCACCGCTGACTCTCAAGGTCGCTTTCTGA
 GCAGCACCGAATTGCAAATTGCTTTCCGGTCGTCTACGTCAAGCTAATGCTGGTTTGCAAGCCGCTAAAG
 CTCTGACCGACAATGCCCAGAGCTTGGTAAATGGTGCTGCCCAAGCCGTTTATAACAAATTTCCCTACA
 CCACCCAAACCCAAGGCAACAACCTTTGCTGCGGATCAACGGGGTAAAGACAAGTGTGCCCGGGACATC
 GGCTACTACCTCCGCATCGTTACCTACTGCTTAGTTGCTGGTGGTACCGGTCCTTTGGATGAGTACTTG
 ATCGCCGGTATTGATGAAATCAACCGCACCTTTGACCTCTCCCCCAGCTGGTATGTTCTCGAG

Construct used for the *cpcB*IFN-cmR* sequence insertion within the *cpc* operon (*IFN*, codon-optimized human interferon gene for expression in *Synechocystis*).



Sequence for 5'- homologous recombination

UPPER CASE ITALICS - *cpcB* (without stop codon)

UPPER CASE - Xa factor cleavage site (IEGR)
UPPER CASE - codon-optimized human interferon gene for expression in *Synechocystis*
 lower case - intergenic sequence
UPPER CASE BOLD - *cmR* (chloramphenicol selection cassette)
 lower case - Transcription terminator

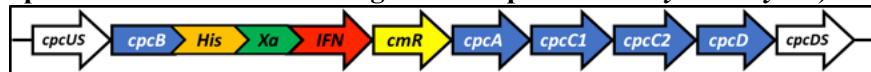
Sequence for 3'- homologous recombination

UPPER CASE ITALICS - *cpcB-cpcA* intergenic sequence

UPPER CASE BOLD - *cpcA* (partial)

CTCGAGATGTTTCGACGTATTCACCTCGGGTGTTCCTCCAAGCTGATGCTCGCGGCGAGTACCTCTCTGGTTCTCAGTTA
 GATGCTTTGAGCGCTACCGTTGCTGAAGGCAACAAACGGATTGATTCTGTTAACCGCATCACCGGTAATGCTTCCGCT
 ATCGTTTCCAACGCTGCTCGTGCTTTGTTTCGCCGAACAGCCCCAATTAATCCAACCCGGTGAAACGCCTACACAG
 CCGTCGTATGGCTGCTTGTTCGCTGACATGGAATCATCCTCCGCTATGTTACCTACGCAACCTTACCGGCGACG
 CTTCCGTTCTAGAAGATCGTTGCTTGAACGGTCTCCGTGAAACCTACGTTGCCCTGGGTGTTCCCGGTGCTTCCGTA
 GCTGCTGGCGTTCAAAAAATGAAAGAAGCTGCCCTGGACATCGTTAACGATCCCAATGGCATCACCCGTGGTGATTG
 CAGTGCTATCGTTGCTGAAATCGCTGGTTACTTCGACCGCGCCGCTGCTGCCGTAGCC**ATCGAAGGGCGA**TGTGA
 CTTGCCTCAGACGCATTCTTTGGGAAGCCGACGCACACTGATGCTGCTCGCCCAAATGCGCCGGATCTCCTTA
 TTCTCCTGTCTCAAGGATCGGCATGACTTCGGCTTCCCTCAGGAGGAGTTTGGAAATCAGTTCCAAAAGGCCG
 AAACCATTCGGTCTCCATGAAATGATTCAACAGATCTTTAACTTATTCAGTACCAAAGACAGCAGTGCAGG
 CTGGGACGAAACATTACTCGATAAATTCTACACGGAATTATACCAACAGTTGAACGACTTAGAAGCCTGTGT
 AATCCAAGGTGTTGGTGCTACTGAGACTCCATTAATGAAAGAAGACTCTATTCTGGCCGTCCGCAAGTATTTT
 CAGCGAATCACACTGTATTTGAAAGAGAAAAAGTATTCTCCGTGTGCGTGGGAGGTAGTACGGGCTGAAATC
 ATGCGGTCCTTCTCTTTAAGCACAAACCTCCAGGAATCTCTGCGCTCCAAAGAATGA**AGATCT**gcggccgcgttgatcg
 gcacgtaagaggttccaacttcaccataatgaataagatcactaccggcgctatttttgagtatcgagatttcaggagctaaggaagctaaa**ATGGAGAAAAAA**
TCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAG
TTGCTCAATGTACCTATAACCAGACCGTTACGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAA
ATAAGCACAAAGTTTTATCCGGCCTTTATTACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCC
GTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTACCCCTTGTTACACCGTTTTCCATG
AGCAAACCTGAAACGTTTTATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATAT
ATTTCGCAAGATGTGGCGTGTTACGGTGAAAACTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGT
TTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTGTATTTAAACGTGGCCAATATGGACAAC
TCTTCGCCCCCGTTTTTACCATTGGGCAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCG
ATTTCAGGTTTCATCATGCCGTCTGTGATGGCTTCCATGTTCGGCAGAATGCTTAATGAATTACAACAGTAC
TGCGATGAGTGGCAGGGCGGGCGCTAAtttttaaggcagttattggtgccttaaacgcctgg**GGATCC**TCTGGTTATTTAAAAAC
 CAACTTTACTCAGGTTCCATACCCGAGAAAATCCAGCTTAAAGCTGACATATCTAGGAAAAATTTTACATTCTAACGGG
 AGATACCAGAAACAATGAAAACCCCTTTAACTGAAGCCGTTTCCACCGCTGACTCTCAAGGTCGCTTTCTGA
 GCAGCACCGAATTGCAAATTGCTTTTCGGTCGTCTACGTCAAGCTAATGCTGGTTTGCAAGCCGCTAAAG
 CTCTGACCGACAATGCCCAGAGCTTGGTAAATGGTGCTGCCCAAGCCGTTTATAACAAATTTCCCTACA
 CCACCCAAACCCAAGGCAACAACCTTTGCTGCGGATCAACGGGGTAAAGACAAGTGTGCCCGGGACATC
 GGCTACTACCTCCGCATCGTTACCTACTGCTTAGTTGCTGGTGGTACCGGTCCTTTGGATGAGTACTTG
 ATCGCCGGTATTGATGAAATCAACCGCACCTTTGACCTCTCCCCCAGCTGGTATGTT**CTCGAG**

Construct used for the *cpcB*His*Xa*IFN-cmR* sequence insertion within the *cpc* operon (*IFN*, codon-optimized human interferon gene for expression in *Synechocystis*).



CTCGAG - XhoI DNA restriction site
 AGATCT - BglII DNA restriction site
 GGATCC - BamHI DNA restriction site

Sequence for 5'- homologous recombination

UPPER CASE ITALICS - *cpcB* (without stop codon)

UPPER CASE - Histag (6x)
 UPPER CASE - Xa factor cleavage site (IEGR)
 UPPER CASE - codon-optimized human interferon gene for expression in *Synechocystis*
 lower case - intergenic sequence
 UPPER CASE BOLD - *cmR* (chloramphenicol selection cassette)
 lower case - Transcription terminator

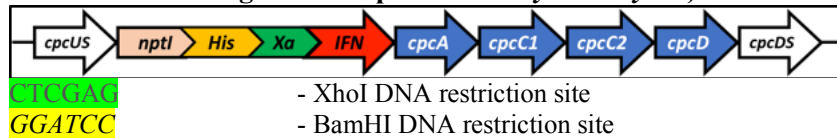
Sequence for 3'- homologous recombination

UPPER CASE ITALICS - *cpcB-cpcA* intergenic sequence

UPPER CASE BOLD - *cpcA* (partial)

CTCGAGATGTTTCGACGTATTCACTCGGGTTGTTTCCCAAGCTGATGCTCGCGGCGAGTACCTCTCTGGTTCTCAGTTA
 GATGCTTTGAGCGCTACCGTTGCTGAAGGCAACAAACGGATTGATTCTGTTAACCGCATCACCGGTAATGCTTCCGCT
 ATCGTTTCCAACGCTGCTCGTCTTGTTCGCGCAACAGCCCCAATTAATCCAACCCGGTGGAACGCCTACACCAG
 CCGTCGTATGGCTGCTTGTGCTGACATGGAATCATCTCCGCTATGTTACCTACGCAACCTTCACCGGCGACG
 CTTCGTTCTAGAAGATCGTTGCTTGAACGGTCTCCGTGAAACCTACGTTGCCCTGGGTGTTCCCGGTGCTTCCGTA
 GCTGCTGGCGTTCAAAAAATGAAAGAAGCTGCCCTGGACATCGTTAACGATCCCAATGGCATCACCCGTGGTGATTG
 CAGTGCTATCGTTGCTGAAATCGCTGGTTACTTCGACCGCGCCGCTGCTGCCGTAGCCACCATCACCATCACCAT
 ATCGAAGGGCGATGTGACTTGCCTCAGACGCATTCTTTGGGAAGCCGACGCACACTGATGCTGCTCGCCCAA
 ATGCGCCGGATCTCCTTATTCTCCTGTCTCAAGGATCGGCATGACTTCGGCTTCCCTCAGGAGGAGTTTGAA
 ATCAGTTCCAAAAGGCCGAAACATTCCGGTCTCCATGAAATGATTCAACAGATCTTTAACTTATTTCAGTAC
 CAAAGACAGCAGTGCAGCCTGGGACGAAACATTACTCGATAAATTCTACACGGAATTATACCAACAGTTGAA
 CGACTTAGAAGCCTGTGTAATCCAAGGTGTTGGTGTCACTGAGACTCCATTAATGAAAGAAGACTCTATTCTG
 GCCGTCCGCAAGTATTTCCAGCGAATCACACTGTATTTGAAAGAGAAAAAGTATTCTCCGTGTGCGTGGGAG
 GTAGTACGGGCTGAAATCATGCGGTCTTCTCTTAAGCACAAACCTCCAGGAATCTCTGCGCTCCAAAGAAT
 GAAGATCTcgggccgcgttgatggcagcgaagaggttccaactttcaccataatgaaataagatcactaccggcgatatttttgagtatcgagatttcaggagcgaaggaa
 gctaaaATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTT
 TGAGGCATTTTCAGTCAGTTGCTCAATGTACCTATAAACCAGACCGTTTCAGCTGGATATTACGGCCTTTTT
 AAAGACCGTAAAGAAAAAATAAGCACAAAGTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAA
 TGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCCACCCTT
 GTTACACCGTTTTCCATGAGCAAACCTGAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCC
 GGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAG
 GGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGT
 GGCCAATATGGACAACCTTCTCGCCCCCGTTTTTACCATGGGCAAAATATTATACGCAAGGCGACAAGGT
 GCTGATGCCGCTGGCGATTCAAGTTTCATCATGCCGTCTGTGATGGCTTCCATGTGCGGCAGAATGCTTAA
 TGAATTACAACAGTACTGCGATGAGTGGCAGGGGCGGGCGTAAttttttaaggcagttattggtgccctaaacgcctggGGATC
 CTCTGGTTATTTAAAAACCAACTTTACTCAGGTTCCATACCCGAGAAAAATCCAGCTTAAAGCTGACATATCTAGGAAA
 ATTTTACATTCTAACGGGAGATACCAGAACAAATGAAAACCCCTTTAACTGAAGCCGTTTCCACCGCTGACTC
 TCAAGGTCGCTTTCTGAGCAGCACCAGAAATTGCAAAATGCTTTTCGGTCGTCTACGTCAAGCTAATGCTGG
 TTTGCAAGCCGCTAAAGACTCTGACCGACAATGCCAGAGCTTGGTAAATGGTGCTGCTGCCCCAAGCCGTTT
 ATAACAAATTCCCCTACACCACCAAAACCCAAGGCAACAACCTTGTGCTGCGGATCAACGGGGTAAAGAC
 AAGTGTGCCCGGGACATCGGCTACTACCTCCGCATCGTTACCTACTGCTTAGTTGCTGGTGGTACCGGT
 CCTTTGGATGAGTACTTGATCGCCGGTATTGATGAAATCAACCGCACCTTTGACCTCTCCCCCAGCTGG
 TATGTTCTCGAG

Construct used for the *nptI*His*Xa*IFN* sequence insertion within the *cpc* operon (*IFN*, codon-optimized human interferon gene for expression in *Synechocystis*).



Sequence for 5'- homologous recombination

UPPER CASE ITALICS - *cpc* operon upstream region

UPPER CASE - *nptI* (kanamycin selection cassette)
 UPPER CASE - Histag (6x)
 UPPER CASE - Xa factor cleavage site (IEGR)
 UPPER CASE - codon-optimized human interferon gene for expression in *Synechocystis*
 lower case - Transcription terminator

Sequence for 3'- homologous recombination

UPPER CASE ITALICS - *cpcB-cpcA* intergenic sequence

UPPER CASE BOLD - *cpcA* (partial)

CTCGAG GGAAAGTAGGCTGTGGTTCCTAGGCAACAGTCTTCCCTACCCCACTGGAAACTAAAAAACGAGAAAAGT
 TCGCACCGAACATCAATTGCATAATTTAGCCCTAAACATAAGCTGAACGAAACTGGTTGTCTTCCCTTCCCAATCCA
 GGACAATCTGAGAATCCCCTGCAACATTACTTAAACAAAAAGCAGGAATAAAATTAACAAGATGTAAACAGACATAAGTC
 CCATCACCGTTGTATAAAGTTAACTGTGGGATTGCAAAAGCATTCAAGCCTAGGCGCTGAGCTGTTTGAGCATCCCGG
 TGGCCCTTGTGCTGCTCCGTGTTTCTCCCTGGATTATTTAGGTAATATCTCTCATAAATCCCCGGGTAGTTAACGA
 AAGTTAATGGAGATCAGTAACAATAACTCTAGGGTCATTACTTTGGACTCCCTCAGTTTATCCGGGGGAATTGTGTTTA
 AGAAAATCCCAACTCATAAAGTCAAGTAGGAGATTAATTCAATGAGTCACATCCAGAGAGAACTAGTTGTTCCCGACC
 TCGTTTGAATAGCAATATGGATGCAGATCTGTACGGATATAAATGGGCGCGAGATAACGTAGGCCAATCTGGGGCCA
 CTATTATCGGTTATATGGCAAAACAGATGCTCCCGAACTGTTTCTCAAAACATGGCAAAGGGTCTGTGGCCAATGATG
 TTACCGATGAAATGGTGGGTTGAAGTGGTACAGAAATTTATGCCCCTCCCGACCATCAAAATTTATCAGGACTC
 CAGACGATGCATGGCTATTAACTACGGCCATTCTTGGGAAAACTGCCTTTCAGGTGTTGGAAGAAATATCCCGATTCTG
 GTGAGAATATCGTCGATGCGTTAGCGGTTTTCTAAGACGTCTACATAGCATTCCCGTTTGCAATTGTCCCTTTAATTC
 GGACCGGGTGTTCGCTTGGCGCAGGCTCAGTCCCGGATGAATAACGGTTTGGTAGATGCCTCGGACTTTGATGAT
 GAACGGAACGGCTGGCCCGTTGAACAGGTTTGGAAAGAGATGCATAAGCTGCTGCCCTTCTCCCCCGACAGCGTTG
 TTAATCATGGAGATTTTCTCTCGATAATCTGATTTTCGACGAAGGCAAGCTAATTGGCTGTATCGATGTGGGACGGGT
 AGGGATTGCGGACCGGTATCAAGACCTAGCAATTTTGTGGAAGTGCCTAGGTGAATTTTCCCCCAGCCTACAAAAACG
 GCTGTTTCAAAAATACGGAATCGATAATCCCGACATGAACAAATTACAATTCATCTGATGCTAGATGAGTTCTTTCAC
 CATCACCATCACCATATCGAAGGGCGATGTGACTTGCCTCAGACGCATTCTTTGGGAAGCCGACGCACACTGA
 TGCTGCTCGCCCAATGCGCCGGATCTCCTTATTCTCTGTCTCAAGGATCGGCATGACTTCGGCTTCCCTCAG
 GAGGAGTTTGGAAATCAGTTCCAAAAGGCCGAAACATTCCGGTCTCCATGAAATGATTCAACAGATCTTTA
 ACTTATTCAGTACCAAAGACAGCAGTGCAGGCTGGGACGAAACATTACTCGATAAATTCTACACGGAATTAT
 ACCAACAGTTGAACGACTTAGAAGCCTGTGTAATCCAAGGTGTTGGTGTCACTGAGACTCCATTAATGAAAG
 AAGACTCTATTCTGGCCGTCCGCAAGTATTTCCAGCGAATCACACTGTATTTGAAAGAGAAAAAGTATTCTCC
 GTGTGCGTGGGAGGTAGTACGGGCTGAAATCATGCGGTCTTCTCTTAAAGCACAAACCTCCAGGAATCTCTG
 CGCTCCAAAGAATGATTTTtaaggcagttattggtgcccttaaacgcctggGGATCCCTTGTTATTTTAAAAACCAACTTTACTCAGGT
 TCCATACCCGAGAAAAATCCAGCTTAAAGCTGACATATCTAGGAAAAATTTTACATTCTAACGGGAGATACCAGAACAAT
 GAAAAACCCCTTTAACTGAAGCCGTTTCCACCGCTGACTCTCAAGGTCGCTTTTCTGAGCAGCACCCGAATT
 GCAAATTGCTTTCGGTTCGTCTACGTCAAGCTAATGCTGGTTTGAAGCCGCTAAAGCTCTGACCGACAA
 TGCCAGAGCTTGGTAAATGGTGTGCTGCCCCAAGCCGTTTATAACAAATTCCCCTACACCACCCAAACCCA
 AGGCAACAACCTTTGCTGCGGATCAACGGGGTAAAGACAAGTGTGCCCCGGGACATCGGCTACTACCTCC
 GCATCGTTACCTACTGCTTAGTTGCTGGTGGTACCGGTCTTTGGATGAGTACTTGATCGCCGGTATTG
 ATGAAATCAACCGCACCTTTGACCTCTCCCCAGCTGGTATGTTCTCGAC