

## SUPPLEMENTARY MATERIAL

Supplementary Figure 1. Accession number and features of <i>Gaussia</i> luciferase in the NCBI Gene database.....	3
Supplementary Figure 2. PSI-PRED secondary structure prediction for the <i>Gaussia</i> luciferase sequence .....	4
Supplementary Figure 3. Folindex© server allows to define disordered areas of folding sequence .....	4
Supplementary Figure 4. Emission spectrum of full-length <i>Gaussia</i> luciferase .....	5
Supplementary Figure 5. Bovine Serum Albumin standard curve calibration.....	5
Supplementary Figure 6. Circular plasmid map of pET28a(+) vector for cloning and expression in <i>E. coli</i> cells.....	6
Supplementary Figure 7. DNA Chromatogram sequencing result from (clone #1) of pET28a(+) GLucFl construct for expression in <i>E. coli</i> cells .....	8
Supplementary Figure 8. Alignment between theoretical coding nucleotide sequence and the sequencing result of pET28a(+) GLucFl construct (clone #1) for expression in <i>E. coli</i> cells.....	9
Supplementary Figure 9. Circular plasmid map of pQE vectors for cloning and expression in <i>E. coli</i> cells.....	10
Supplementary Figure 10. Circular plasmid map of ligation from PCR insert of GLucFlThrom in pQE31 vector.....	11
Supplementary Figure 11. DNA Chromatogram sequencing result from (clone #12) of pQE31GLucFlThrombin construct for expression in <i>E. coli</i> cells .....	13
Supplementary Figure 12. Alignment between theoretical coding nucleotide sequence and the sequencing result of pQE31GLucFlThrombin construct (clone #12) for expression in <i>E. coli</i> cells.....	14
Supplementary Figure 13. DNA Chromatogram sequencing result from (clone #7) of pQE31GLucD1 construct for expression in <i>E. coli</i> cells.....	16
Supplementary Figure 14. Alignment between theoretical coding nucleotide sequence and the sequencing result of pQE31GLucD1 construct (clone #7) for expression in <i>E. coli</i> cells.....	17
Supplementary Figure 15. DNA Chromatogram sequencing result from (clone #125) of pQE30GLucD2 construct for expression in <i>E. coli</i> cells.....	19

---

Supplementary Figure 16. Alignment between theoretical coding nucleotide sequence and the sequencing result of pQE30GLucD2 construct (clone #125) for expression in <i>E. coli</i> cells.....	20
Supplementary Figure 17. Circular map of plasmid vector pFastBac1 for Baculovirus expression in insect cells for <i>in vivo</i> crystallization method.....	21
Supplementary Figure 18. Full-length GLuc sequence for cloning experiments in pFastBac1 vector for crystallization <i>in vivo</i> using <i>Sf21</i> insect cells. ....	22
Supplementary Figure 19. DNA Chromatogram sequencing result from (clone #214) of pFastBac1-ERGLucFl construct for crystallization <i>in vivo</i> using <i>Sf21</i> insect cells.....	23
Supplementary Figure 20. Alignment between theoretical coding nucleotide sequence and the sequencing result of pFastBac1-ERGLucFl construct (clone #214) for crystallization <i>in vivo</i> using <i>Sf21</i> insect cells .....	25
Supplementary Figure 21. Chromatogram sequencing result from (clone #311_1) of pFastBac1-ER-KDELGLucFl construct for crystallization <i>in vivo</i> using <i>Sf21</i> insect cells.....	26
Supplementary Figure 22. Alignment between theoretical coding nucleotide sequence and the sequencing result of pFastBac1-ER-KDELGLucFl construct (clone #311_1) for crystallization <i>in vivo</i> using <i>Sf21</i> insect cells.....	28
Supplementary Figure 23. DNA Chromatogram sequencing result from (clone #414) of pFastBac1-SKLGLucFl construct for crystallization <i>in vivo</i> using <i>Sf21</i> insect cells.....	29
Supplementary Figure 24. Alignment between theoretical coding nucleotide sequence and the sequencing result of pFastBac1-SKLGLucFl construct (clone #414) for crystallization <i>in vivo</i> using <i>Sf21</i> insect cells .....	31

```

Gaussia luciferase features

ACCESSION number in NCBI  AAG54095
FEATURES Location/Qualifiers
Source                1..758
                        /organism="Gaussia princeps (T. Scott, 1894)"
                        /mol_type="mRNA"
                        /db_xref="taxon:148582"
CDS                  30..587
                        /codon_start=1
                        /product="luciferase"
                        /protein_id="AAG54095.1"
                        /db_xref="GI:12621054"

/translation="MGVKVLFALICIAVAEAKPTENNEDFNIVAVASNFATTDLDADRGKLPGKKLPLEVLKEME
ANARKAGCTRGCLICLSHIKCTPKMKKFIPGRCHTYEGDKESAQGGIGEAIVDIPEIPGFKDLEPMEQFIAQVDL
CVDCTTGCLKGLANVQCSDLLKKWLPQRCATFASKIQGQVDKIKGAGGD"

ORIGIN (Gaussia luciferase nucleotide sequence)
   1 ggtactcaaa gtatcttctg gcagggaaaa tgggagtgaa agttcttttt gcccttattt
  61 gtattgctgt ggccgagggc aaaccaactg aaaacaatga agatttcaac attgtagctg
 121 tagctagcaa ctttgctaca acggatctcg atgctgaccg tggtaaattg cccggaaaaa
 181 aattaccact tgaggactc aaagaaatgg aagccaatgc taggaaagct ggctgcacta
 241 ggggatgtct gatatgcctg tcacacatca agtgtacacc caaaatgaag aagtttatcc
 301 caggaagatg ccacacctat gaaggagaca aagaaagtgc acagggagga ataggagagg
 361 ctattgttga cattcctgaa attcctgggt ttaaggattt ggaacccatg gaacaattca
 421 ttgcacaagt tgacctatgt gtagactgca caactggatg cctcaaaggt cttgccaatg
 481 tgcaatgttc tgatttactc aagaaatggc tgccacaaag atgtgcaact tttgctagca
 541 aaattcaagg ccaagtggac aaaataaagg gtgccggtgg tgattaatcc taatagaata
 601 ctgcataact ggatgatgat atactagctt attgctcata aatggccat tttttgtaac
 661 aaatcgagtc tatgtaattc aaaataccta attaattggt aatacatatg taattcctat
 721 aaatataatt tatgcaatcc aaaaaaaaaa aaaaaaaaaa

//

Gaussia luciferase amino acid sequence including signal peptide (1-185aa)
MGVKVLFALICIAVAEAKPTENNEDFNIVAVASNFATTDLDADRGKLPGKKLPLEVLKEMEANARKAGCTRGCLI
CLSHIKCTPKMKKFIPGRCHTYEGDKESAQGGIGEAIVDIPEIPGFKDLEPMEQFIAQVDLCVDCTTGCLKGLAN
VQCSDLLKKWLPQRCATFASKIQGQVDKIKGAGGD

(-17aa) MGVKVLFALICIAVAEAK Signal peptide

Gaussia luciferase amino acid sequence without signal peptide (1-168aa)

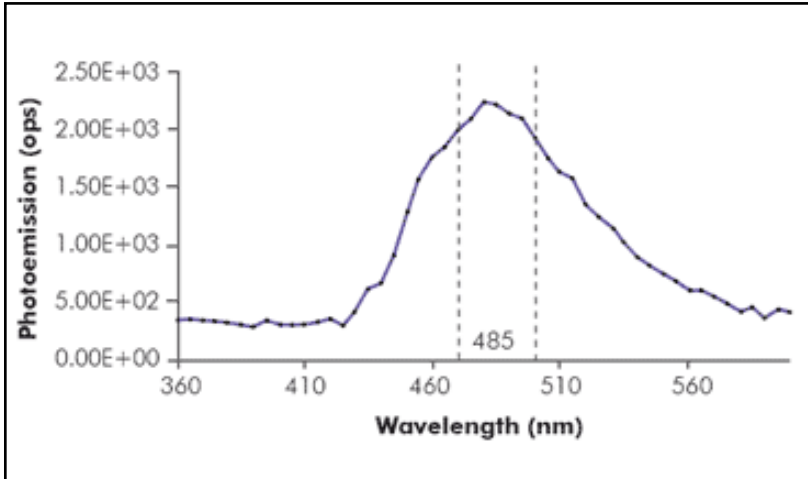
KPTENNEDFNIVAVASNFATTDLDADRGKLPGKKLPLEVLKEMEANARKAGCTRGCLICLSHIKCTPKMKKFIPG
RCHTYEGDKESAQGGIGEAIVDIPEIPGFKDLEPMEQFIAQVDLCVDCTTGCLKGLANVQCSDLLKKWLPQRCAT
FASKIQGQVDKIKGAGGD

```

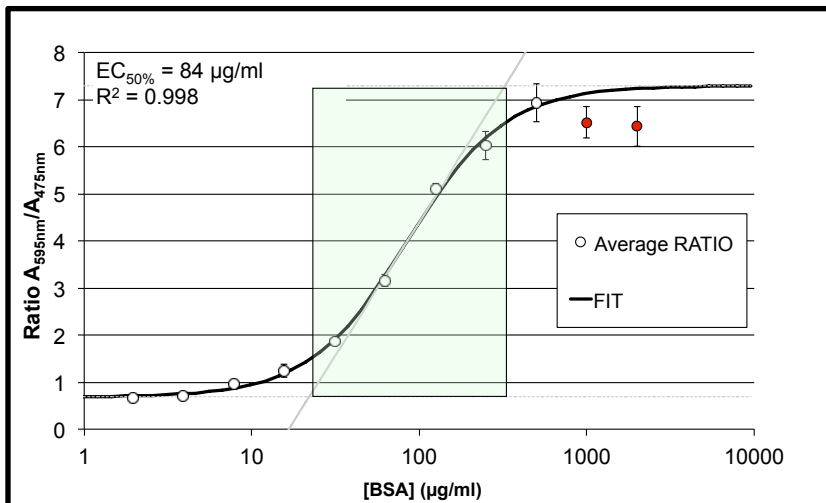
**Supplementary Figure 1. Accession number and features of *Gaussia* luciferase in the NCBI Gene database**

*Gaussia* luciferase protein translation sequence, nucleotide sequence and amino acid sequence with and without signal peptide.



**Supplementary Figure 4. Emission spectrum of full-length *Gaussia luciferase***

Bioluminescence emission spectra of full-length GLuc in the presence of CTZ. Dotted lines show the maximal photoemission from full-length GLuc is at 485 nm.

**Supplementary Figure 5. Bovine Serum Albumin standard curve calibration**

Standard curves from the Coomassie protein assay micro using microplate protocol and BSA (bovine serum albumin, Pierce Rockford, IL.) as reference protein. The curve represents data generated by two experiments with different absorbance wavelengths (595 nm and 475 nm). The plotted parameters reveal an IC<sub>50%</sub> value of 84 µg/ml (r<sup>2</sup> = 0.998). Error bars represent standard deviation (n=3).

## pET-28a-c(+) Vectors

TB074 12/98

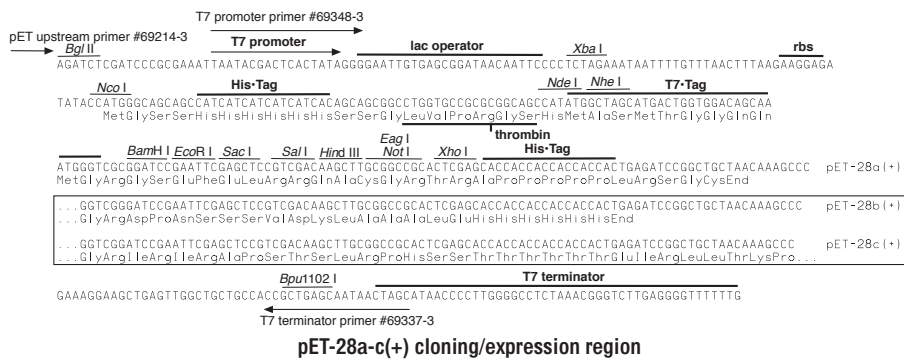
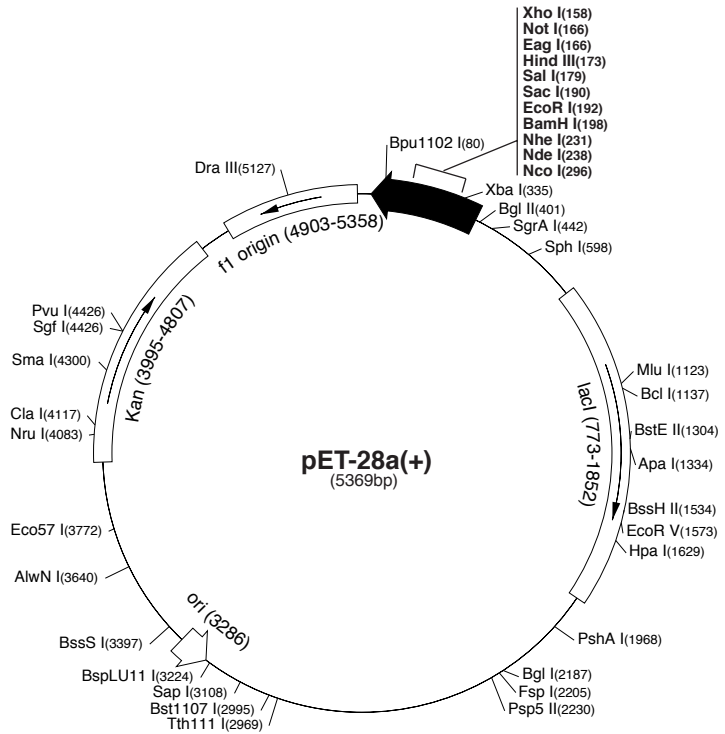
	Cat. No.
pET-28a DNA	69864-3
pET-28b DNA	69865-3
pET-28c DNA	69866-3

The pET-28a-c(+) vectors carry an N-terminal His•Tag<sup>®</sup>/thrombin/T7•Tag<sup>®</sup> configuration plus an optional C-terminal His•Tag sequence. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer (Cat. No. 69337-3).

### pET-28a(+) sequence landmarks

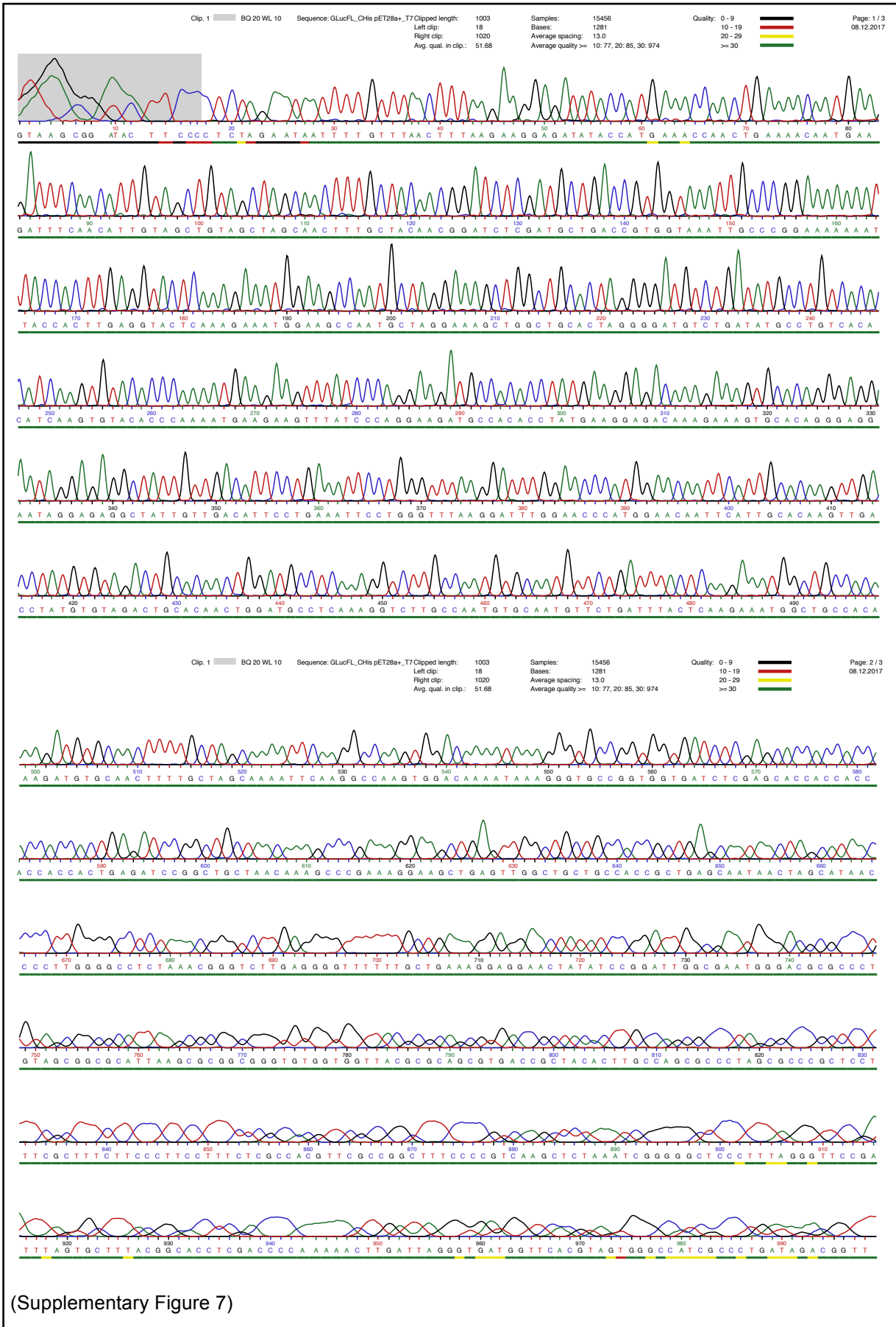
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites	
( <i>Bam</i> H I - <i>Xho</i> I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
<i>lac</i> I coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198.

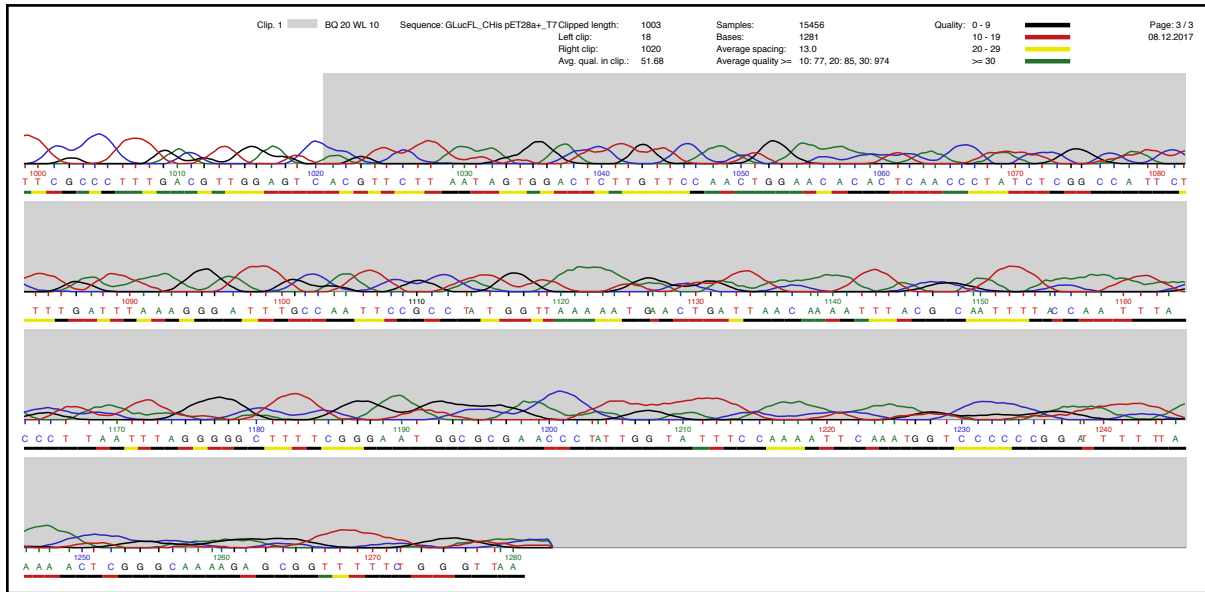


### Supplementary Figure 6. Circular plasmid map of pET28a(+) vector for cloning and expression in *E. coli* cells

Features plasmid map of pET28a(+) vector (Merck, Darmstadt, Germany) used for cloning the full length GLuc sequence for expression in *E. coli* cells.



(Supplementary Figure 7)



**Supplementary Figure 7. DNA Chromatogram sequencing result from (clone #1) of pET28a(+) GLucFL construct for expression in *E. coli* cells**



Alignment of Sequence\_1: [Ligation GLuc FL pET28 a plus C-His.txt.xdna] with  
Sequence\_2: [Sequencing clone #1.xdna]

Similarity : 1140/5739 (19,86 %)

```

P R N * Y D S L * G N C E R I T I P L *
Seq_1 1      tcccgcgaaattaatacgaactcactataggggaattgtgagcggataacaattcccctct 60
Seq_2 1      -----GTAA-GCG-GATACTTCCCTCT 21

K * F C L T L R R R Y T M K P T E N N E
Seq_1 61     agaaataattttgtttaactttaagaaggagatataccatgAAACCAACTGAAAACAATG 120
Seq_2 22     AG-AATAATTTGTTTAACTTTAAGAAGGAGATATACCATGAAACCAACTGAAAACAATG 80

D F N I V A V A S N F A T T D L D A D R
Seq_1 121    AAGATTTCAACATTGTAGCTGTAGCTAGCAACTTTGCTACAACGGATCTCGATGCTGACC 180
Seq_2 81     AAGATTTCAACATTGTAGCTGTAGCTAGCAACTTTGCTACAACGGATCTCGATGCTGACC 140

G K L P G K K L P L E V L K E M E A N A
Seq_1 181    GTGGTAAATGCCCCGAAAAAATTACCACTTGAGGTACTCAAAGAAATGGAAGCCAATG 240
Seq_2 141    GTGGTAAATGCCCCGAAAAAATTACCACTTGAGGTACTCAAAGAAATGGAAGCCAATG 200

R K A G C T R G C L I C L S H I K C T P
Seq_1 241    CTAGGAAAGCTGGCTGCACTAGGGGATGTCTGATATGCCTGTCACACATCAAGTGTACAC 300
Seq_2 201    CTAGGAAAGCTGGCTGCACTAGGGGATGTCTGATATGCCTGTCACACATCAAGTGTACAC 260

K M K K F I P G R C H T Y E G D K E S A
Seq_1 301    CCAAAATGAAGAAGTTTATCCAGGAAGATGCCACACCTATGAAGGAGACAAAGAAAGTG 360
Seq_2 261    CCAAAATGAAGAAGTTTATCCAGGAAGATGCCACACCTATGAAGGAGACAAAGAAAGTG 320

Q G G I G E A I V D I P E I P G F K D L
Seq_1 361    CACAGGGAGGAATAGGAGAGGCTATTGTTGACATTCCTGAAATTCCTGGGTTTAAGGATT 420
Seq_2 321    CACAGGGAGGAATAGGAGAGGCTATTGTTGACATTCCTGAAATTCCTGGGTTTAAGGATT 380

E P M E Q F I A Q V D L C V D C T T G C
Seq_1 421    TGGAACCCATGGAACAATTCATTGCACAAGTTGACCTATGTGTAGACTGCACAACCTGGAT 480
Seq_2 381    TGGAACCCATGGAACAATTCATTGCACAAGTTGACCTATGTGTAGACTGCACAACCTGGAT 440

L K G L A N V Q C S D L L K K W L P Q R
Seq_1 481    GCCTCAAAGGTCTTGCCAATGTGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAA 540
Seq_2 441    GCCTCAAAGGTCTTGCCAATGTGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAA 500

C A T F A S K I Q G Q V D K I K G A G G
Seq_1 541    GATGTGCAACTTTTGTAGCAAAATTC AAGGCCAAGTGGACAAAATAAAGGGTGCCGGTG 600
Seq_2 501    GATGTGCAACTTTTGTAGCAAAATTC AAGGCCAAGTGGACAAAATAAAGGGTGCCGGTG 560

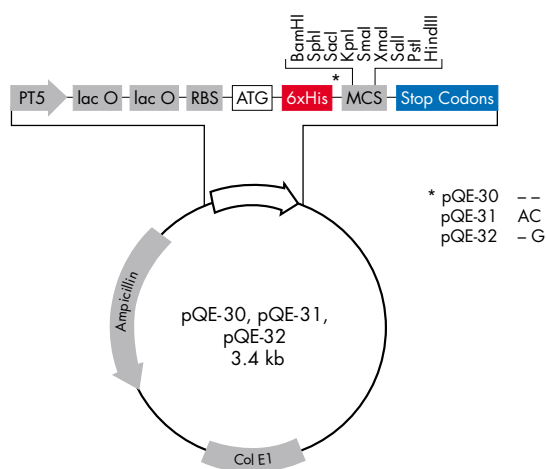
D L E H H H H H H * D P A A N K A R K E
Seq_1 601    GTGATCtcgagcaccaccaccaccactgagatccggctgctaacaagcccgaagg 660
Seq_2 561    GTGATCTCGAGCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAAAGG 620

```

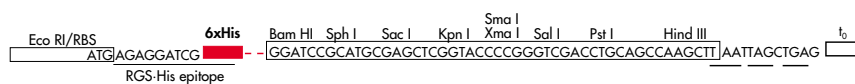
**Supplementary Figure 8. Alignment between theoretical coding nucleotide sequence and the sequencing result of pET28a(+) GLucFI construct (clone #1) for expression in *E. coli* cells**

### pQE-30, pQE-31, and pQE-32 Vectors

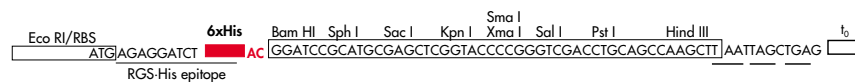
Positions of elements in bases	pQE-30	pQE-31	pQE-32
Vector size (bp)	3461	3463	3462
Start of numbering at <i>Xho</i> I (CTCGAG)	1–6	1–6	1–6
T5 promoter/lac operator element	7–87	7–87	7–87
T5 transcription start	61	61	61
6xHis-tag coding sequence	127–144	127–144	127–144
Multiple cloning site	145–192	147–194	146–193
Lambda <i>t</i> <sub>0</sub> transcriptional termination region	208–302	210–304	209–303
<i>rrnB</i> T1 transcriptional termination region	1064–1162	1066–1164	1065–1163
ColE1 origin of replication	1638	1640	1639
β-lactamase coding sequence	3256–2396	3258–2398	3257–2397



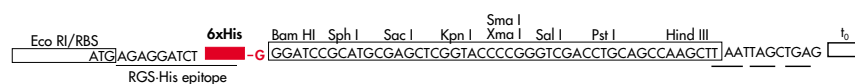
#### pQE-30



#### pQE-31

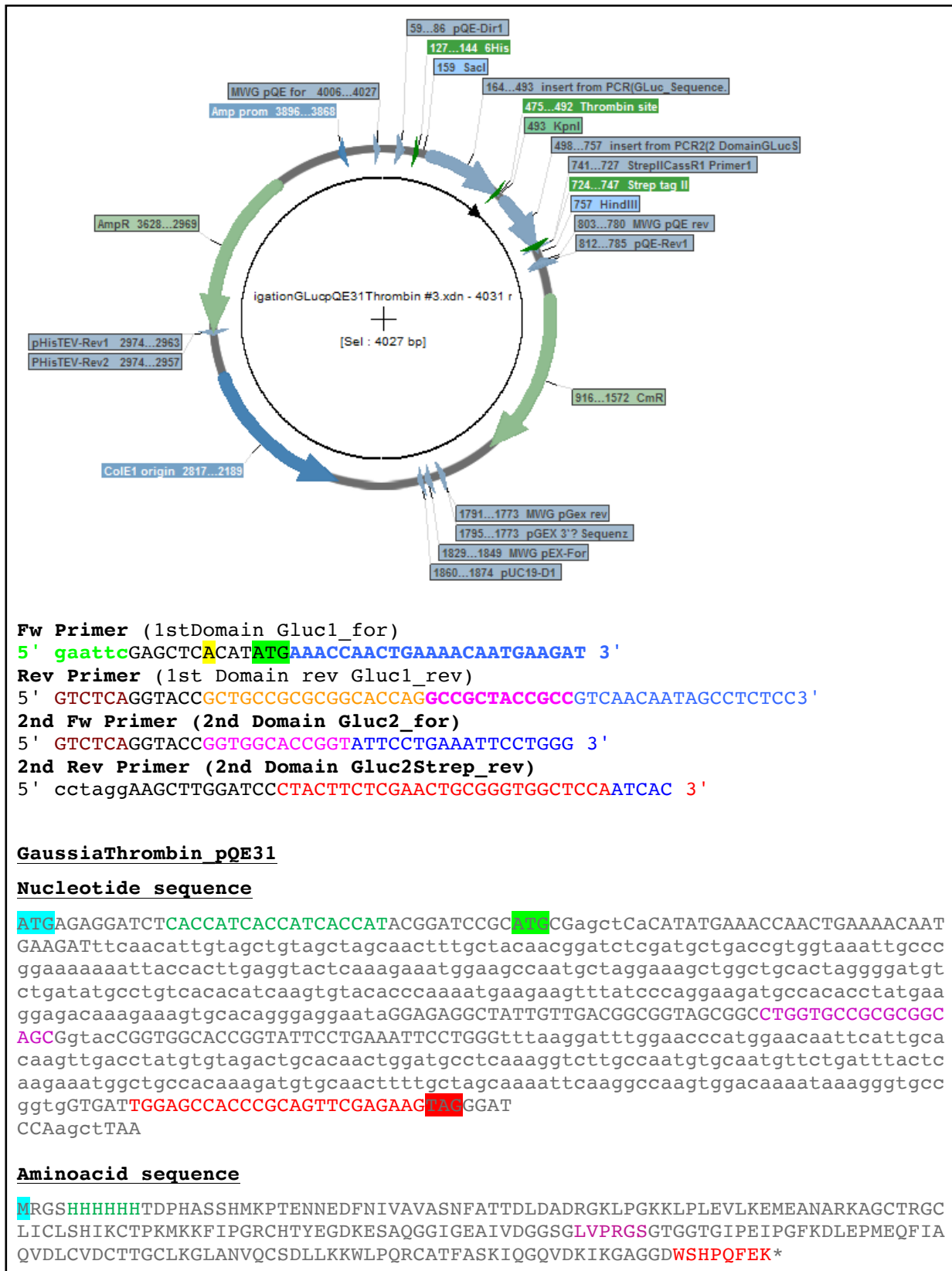


#### pQE-32



### Supplementary Figure 9. Circular plasmid map of pQE vectors for cloning and expression in *E. coli* cells

Features plasmid map of pQE vectors (Qiagen, Hilden, Germany) used for cloning the full-length GLuc and the two catalytic domains sequences for expression in *E. coli* cells.

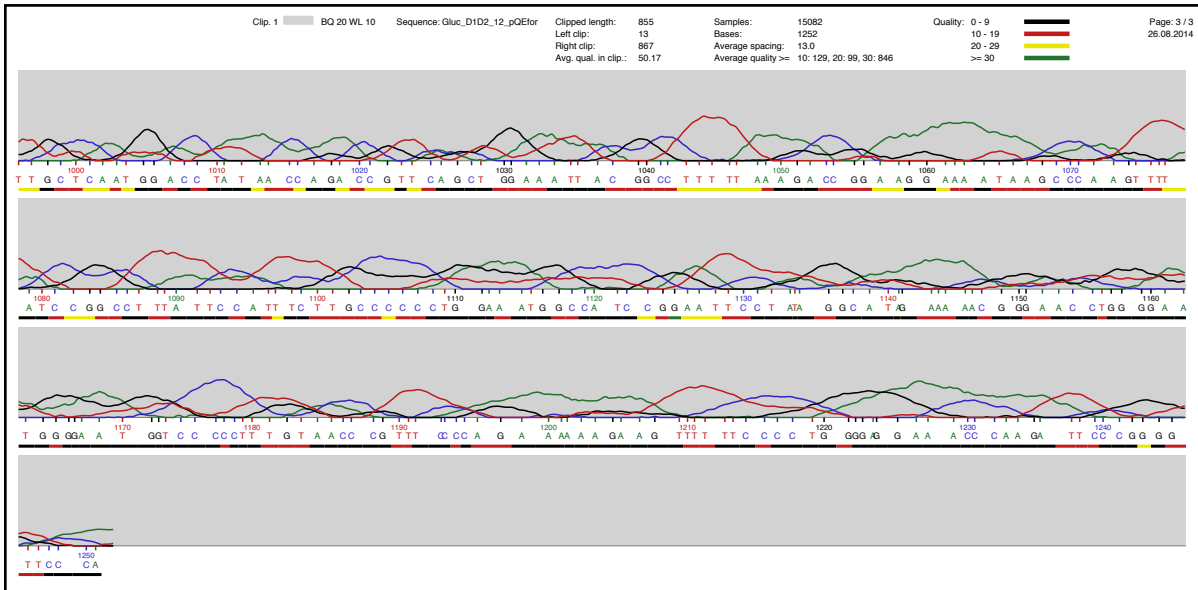


**Supplementary Figure 10. Circular plasmid map of ligation from PCR insert of GLucFIThrom in pQE31 vector**

Ligation of the construct from GLucFIThrombin and the individual two PCR inserts corresponding to the two catalytic domains in pQE31 vector expression in *E. coli* cells. The vector imparts ampicillin and chloramphenicol resistance genes.



(Supplementary Figure 11)



**Supplementary Figure 11. DNA Chromatogram sequencing result from (clone #12) of pQE31GLucFIThrombin construct for expression in *E. coli* cells**

Alignment of Sequence\_1: [LigationGLucpQE31Thrombin #3.xdna] with sequence\_2:[Sequencing result clone #12.xdna]

Similarity : 854/4031 (21,19 %)

```

Seq_1 1      CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCA 60
                |||
Seq_2 1      -----AAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCA 45
                K I Y L L C E R I T I I I D S

Seq_1 61      ATTGTGAGCGGATAACAATTTACACAGAATTCATTAAGAGGAGAAATTAACATGAGA 120
                |||
Seq_2 46      ATTGTGAGCGGATAACAATTTACACAGAATTCATTAAGAGGAGAAATTAACATGAGA 105
                I V S G * Q F H T E F I K E E K L T M R

Seq_1 121     GGATCTCACCATCACCATCACCATACGGATCCGCATGCGagctCaCATATGAAACCAACT 180
                |||
Seq_2 106     GGATCTCACCATCACCATCACCATACGGATCCGCATGCGAGCTCACATATGAAACCAACT 165
                G S H H H H H H T D P H A S S H M K P T

Seq_1 181     GAAAACAATGAAGATttcaacattgtagctgtagctagcaactttgctacaacggatctc 240
                |||
Seq_2 166     GAAAACAATGAAGATTTCAACATTGTAGCTGTAGCTAGCAACTTTGCTACAACGGATCTC 225
                E N N E D F N I V A V A S N F A T T D L

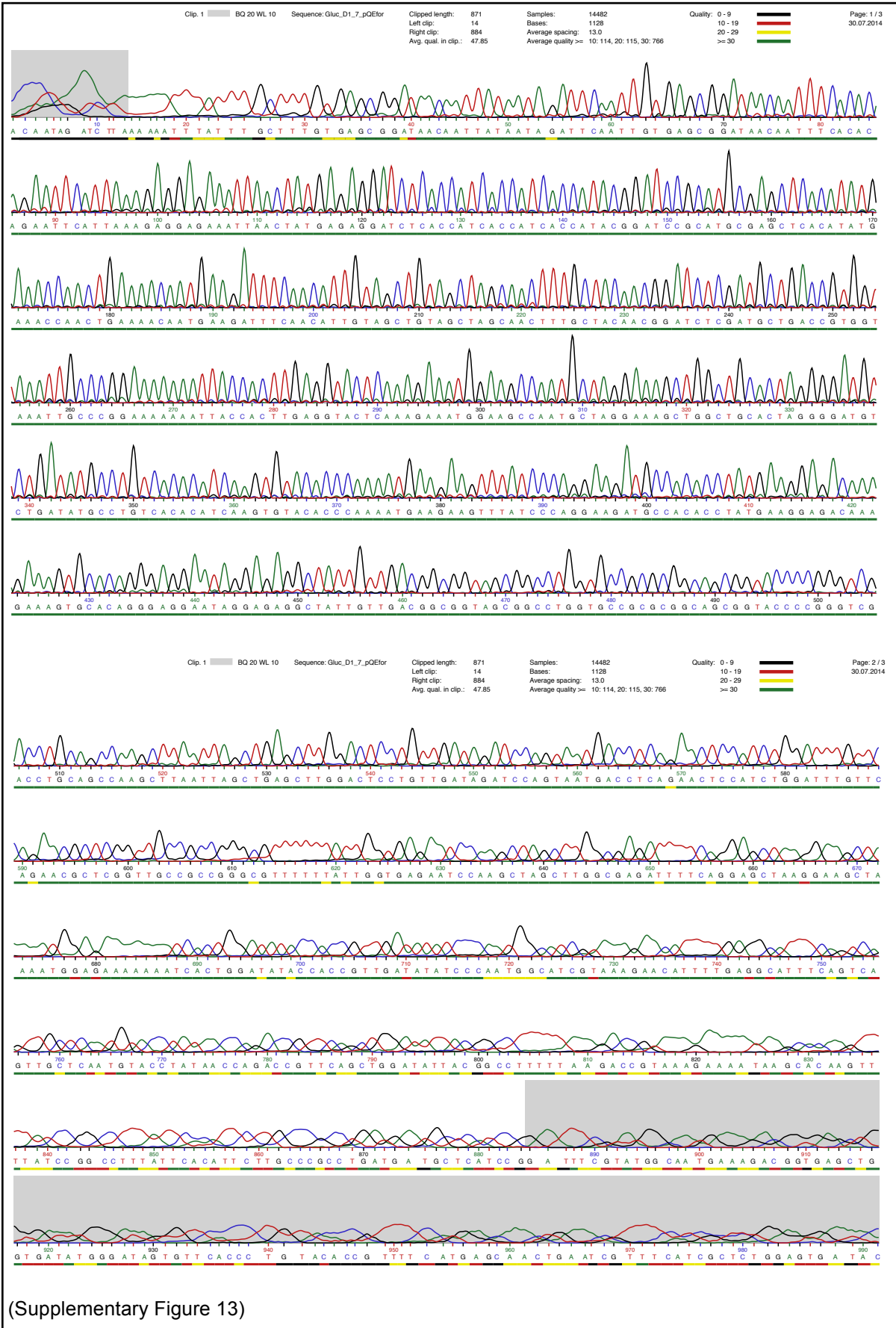
Seq_1 241     gatgctgaccgtggtaaaattgcccggaaaaaaattaccacttgagggtactcaagaaatg 300
                |||
Seq_2 226     GATGCTGACCGTGGTAAATTGCCCGGAAAAAAATTACCACTTGAGGTACTCAAAGAAATG 285
                D A D R G K L P G K K L P L E V L K E M

Seq_1 301     gaagccaatgctaggaagctggctgcactaggggatgtctgatatgcctgtcacacatc 360
                |||
Seq_2 286     GAAGCCAATGCTAGGAAAGCTGGCTGCACTAGGGGATGTCTGATATGCCTGTCACACATC 345
                E A N A R K A G C T R G C L I C L S H I
    
```

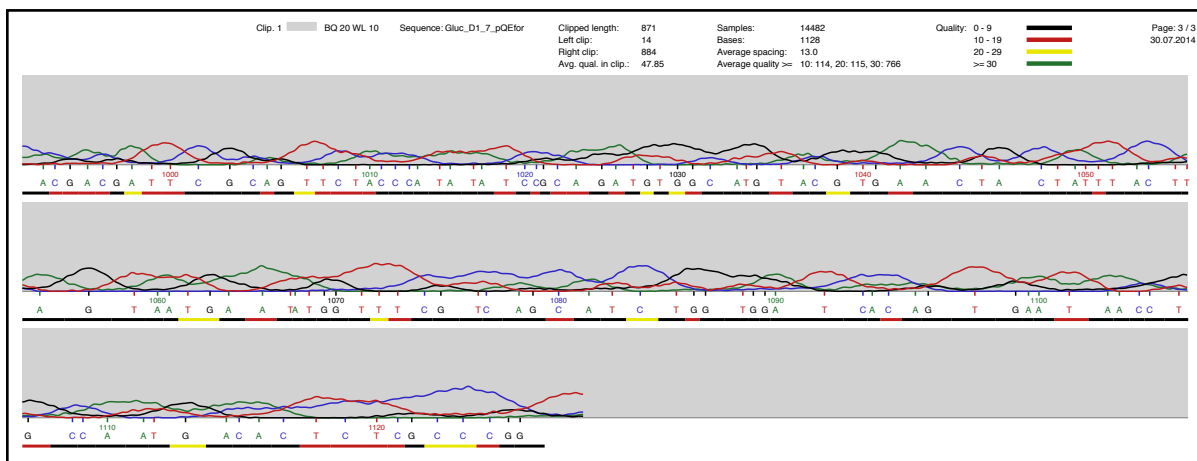
(Supplementary Figure 12)

Seq_1	361	aagtgtacacccaaaaatgaagaagtttatcccaggaagatgccacacctatgaaggagac	420
Seq_2	346	AAGTGTACACCCAAAATGAAGAAGTTTATCCCAGGAAGATGCCACACCTATGAAGGAGAC	405
		K C T P K M K K F I P G R C H T Y E G D	
Seq_1	421	aaagaaagtgcacagggaggaataGGAGAGGCTATTGTTGACGGCGGTAGCGGCCCTGGTG	480
Seq_2	406	AAAGAAAGTGCACAGGAGGAATAGGAGAGGCTATTGTTGACGGCGGTAGCGGCCCTGGTG	465
		K E S A Q G G I G E A I V D G G S G L V	
Seq_1	481	CCGCGCGGCAGCGgtacCGGTGGCACCGGTATTCTCGAAATTCCTGGGttaaaggatttg	540
Seq_2	466	CCGCGCGGCAGCGGTACCGGTGGCACCGGTATTCTCGAAATTCCTGGGTTTAAGGATTTG	525
		P R G S G T G G T G I P E I P G F K D L	
Seq_1	541	gaacccatggaacaattcattgcacaagttgacctatgtgtagactgcacaactggatgc	600
Seq_2	526	GAACCCATGGAACAATTCATTGCACAAGTTGACCTATGTGTAGACTGCACAACCTGGATGC	585
		E P M E Q F I A Q V D L C V D C T T G C	
Seq_1	601	ctcaaaggtcttgccaatgtgcaatgttctgatttactcaagaaatggctgccacaaaga	660
Seq_2	586	CTCAAAGGTCTTGCCAATGTGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAAGA	645
		L K G L A N V Q C S D L L K K W L P Q R	
Seq_1	661	tgtgcaacttttgctagcaaaaattcaaggccaagtggacaaaataaagggtgccgggtgGT	720
Seq_2	646	TGTGCAACTTTTGCTAGCAAAAATTCAAGGCCAAGTGGACAAAATAAAGGGTGCCGGTGGT	705
		C A T F A S K I Q G Q V D K I K G A G G	
Seq_1	721	GATtggAGCCACCCGAGTTCGAGAAGTAGGGATCCAagctTAATTAGCTGAGCTTGGAC	780
Seq_2	706	GATTGGAGCCACCCGAGTTCGAGAAGTAGGGATCCAAGCTTAATTAGCTGAGCTTGGAC	765
		D W S H P Q F E K * G S K L N * L S L D	
Seq_1	781	TCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTTCAGAACGCTCG	840
Seq_2	766	TCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTTCAGAACGCTCG	825
		S C * * I Q * * P Q N S I W I C S E R S	
Seq_1	841	GTTGCCGCCGGCGTTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGA	900
Seq_2	826	GTTGCCGCCGGCGTTTTTTATTGGTGA-AA-----	855
		V A A G R F L L V K	

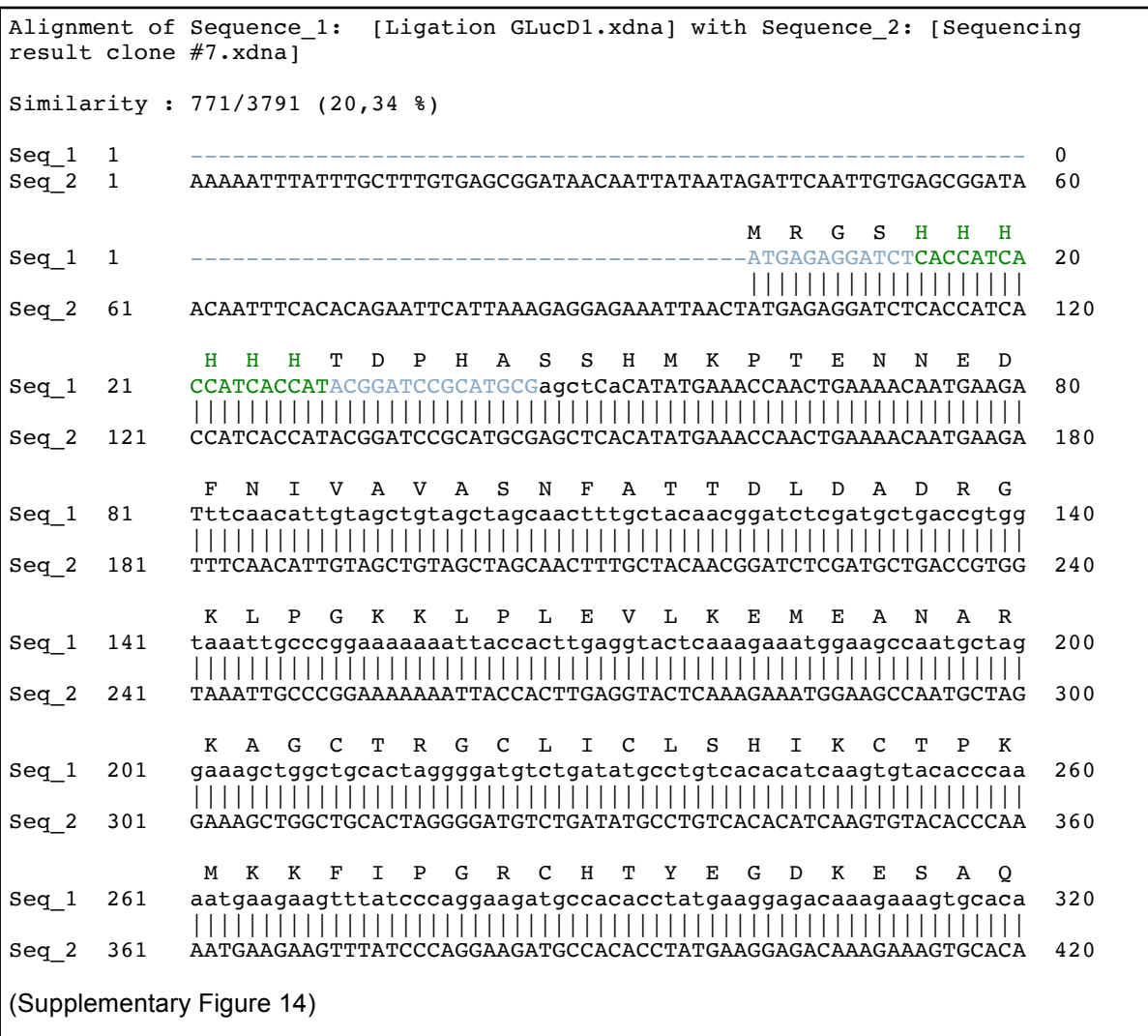
**Supplementary Figure 12. Alignment between theoretical coding nucleotide sequence and the sequencing result of pQE31GLucFITHrombin construct (clone #12) for expression in *E. coli* cells**



(Supplementary Figure 13)



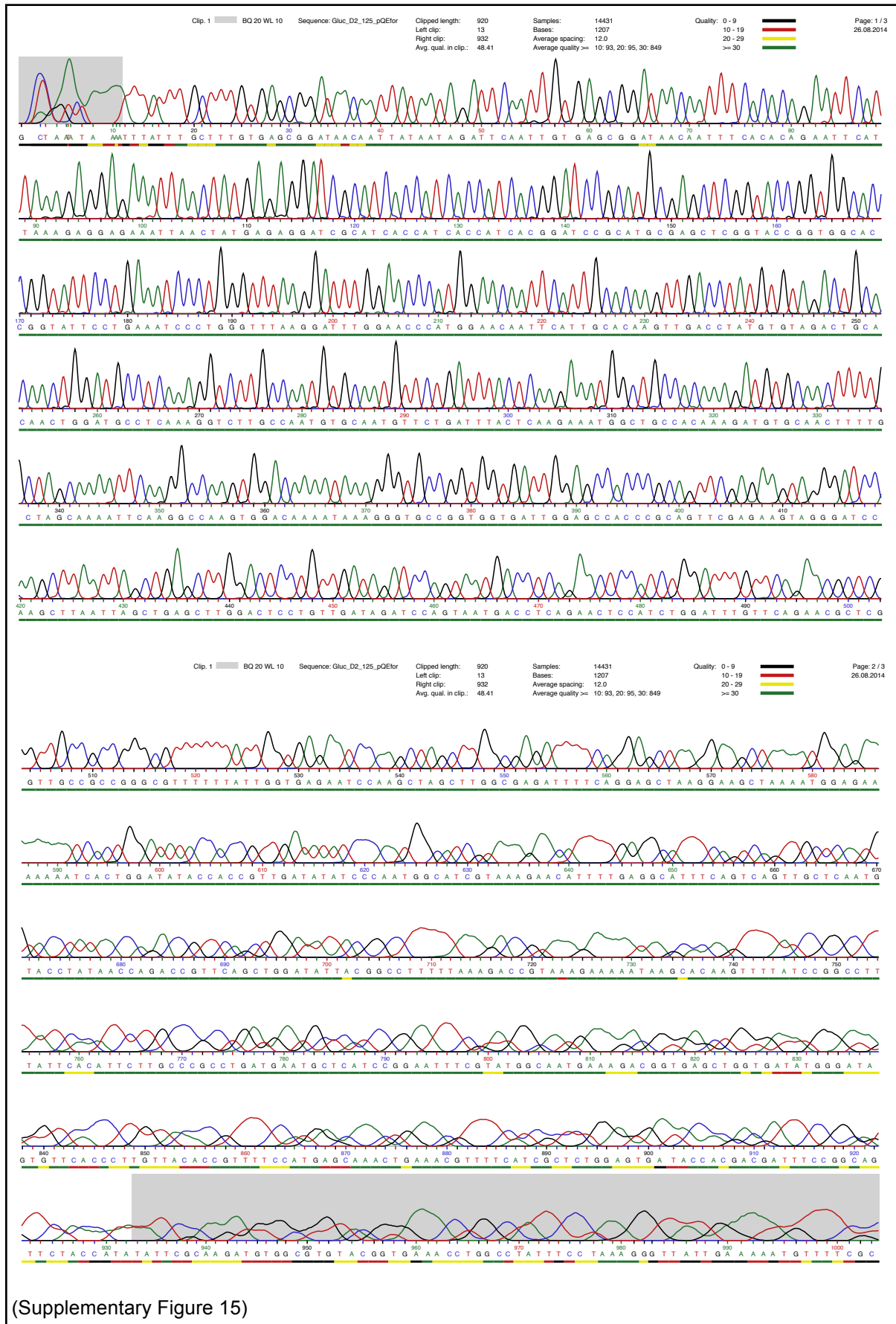
**Supplementary Figure 13. DNA Chromatogram sequencing result from (clone #7) of pQE31GLucD1 construct for expression in *E. coli* cells**



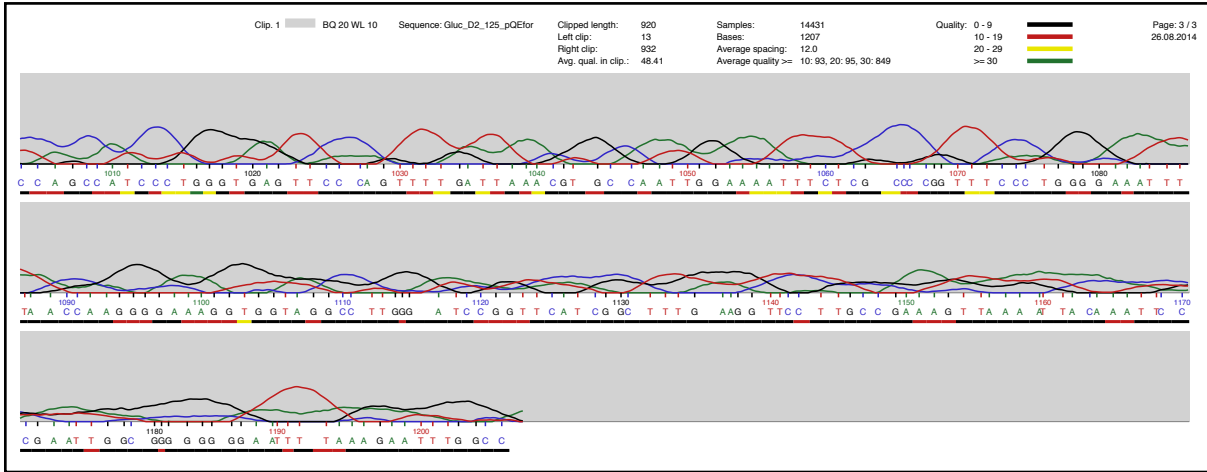


Seq_1	321	G G I G E A I V D G G S G L V P R G S G gggaggaataGGAGAGGCTATTGTTGACGGCGGTAGCGCCTGGTGCCGCGCGGCAGCGg	380
Seq_2	421	GGGAGGAATAGGAGAGGCTATTGTTGACGGCGGTAGCGCCTGGTGCCGCGCGGCAGCGG	480
Seq_1	381	T P G R P A A K L N * L S L D S C * * I tacCCCGGGTTCGACCTGCAGCCAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGAT	440
Seq_2	481	TACCCGGGTTCGACCTGCAGCCAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGAT	540
Seq_1	441	Q * * P Q N S I W I C S E R S V A A G R CCAGTAATGACCTCAGAACTCCATCTGGATTTGTTTCAGAACGCTCGGTTGCCGCCGGGCG	500
Seq_2	541	CCAGTAATGACCTCAGAACTCCATCTGGATTTGTTTCAGAACGCTCGGTTGCCGCCGGGCG	600
Seq_1	501	F L L V R I Q A S L A R F S G A K E A K TTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAA	560
Seq_2	601	TTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAA	660
Seq_1	561	M E K K I T G Y T T V D I S Q W H R K E AATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCAATGGCATCGTAAAGA	620
Seq_2	661	AATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCAATGGCATCGTAAAGA	720
Seq_1	621	H F E A F Q S V A Q C T Y N Q T V Q L D ACATTTTGAGGCATTTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTTCAGCTGGA	680
Seq_2	721	ACATTTTGAGGCATTTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTTCAGCTGGA	780
Seq_1	681	I T A F L K T V K K N K H K F Y P A F I TATTACGGCCTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTTAT	740
Seq_2	781	TATTACGGCCTTTTAA-GACCGTAAAG-AAAATAAGCACAAGTTTTATCCGGCCTTTTAT	838
Seq_1	741	H I L A R L M N A H P E F R M A M K D G TCACATTCCTTGCCCGCTGATGAATGCTCATCCGGAATTTTCGTATGGCAATGAAAGACGG	800
Seq_2	839	TCACATTCCTTGCCCGCTGATGA-TGCTCATCGG-----	871

**Supplementary Figure 14. Alignment between theoretical coding nucleotide sequence and the sequencing result of pQE31GLucD1 construct (clone #7) for expression in *E. coli* cells**



(Supplementary Figure 15)



**Supplementary Figure 15. DNA Chromatogram sequencing result from (clone #125) of pQE30GLuc2D construct for expression in *E. coli* cells**

Alignment of Sequence\_1: [GLuc\_2D\_125\_pQEfor.seq] with Sequence\_2: [Sequencing GLuc\_2Domain.xdna]

Similarity : 918/1100 (83,45 %)

```

Seq_1 1      TTGTCAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAATTGTGAG 60
          |||
Seq_2 1      -----TTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAATTGTGAG 49
          Y L L C E R I T I I I D S I V S

Seq_1 61      CCGATAACAATTTTCACACAGAATTCATTAAGAGGAGAAAATTAACTATGAGAGGATCGCA 120
          |||
Seq_2 50      CCGATAACAATTTTCACACAGAATTCATTAAGAGGAGAAAATTAACTATGAGAGGATCGCA 109
          G * Q F H T E F I K E E K L T M R G S H

Seq_1 121     TCACCATCACCATCACGGATCCGCATGCGAGCTCGGTACCGGTGGCACCGGTATTCCTGA 180
          |||
Seq_2 110     TCACCATCACCATCACGGATCCGCATGCGAGCTCGGTACCGGTGGCACCGGTATTCCTGA 169
          H H H H H G S A C E L G T G G T G I P E

Seq_1 181     AATCCCTGGGTTTAAGGATTTGGAACCCATGGAACAATTCATTGCACAAGTTGACCTATG 240
          |||
Seq_2 170     AATCCCTGGGTTTAAGGATTTGGAACCCATGGAACAATTCATTGCACAAGTTGACCTATG 229
          I P G F K D L E P M E Q F I A Q V D L C

Seq_1 241     TGTAGACTGCACAACCTGGATGCCTCAAAGGTCTTGCCAATGTGCAATGTTCTGATTTACT 300
          |||
Seq_2 230     TGTAGACTGCACAACCTGGATGCCTCAAAGGTCTTGCCAATGTGCAATGTTCTGATTTACT 289
          V D C T T G C L K G L A N V Q C S D L L

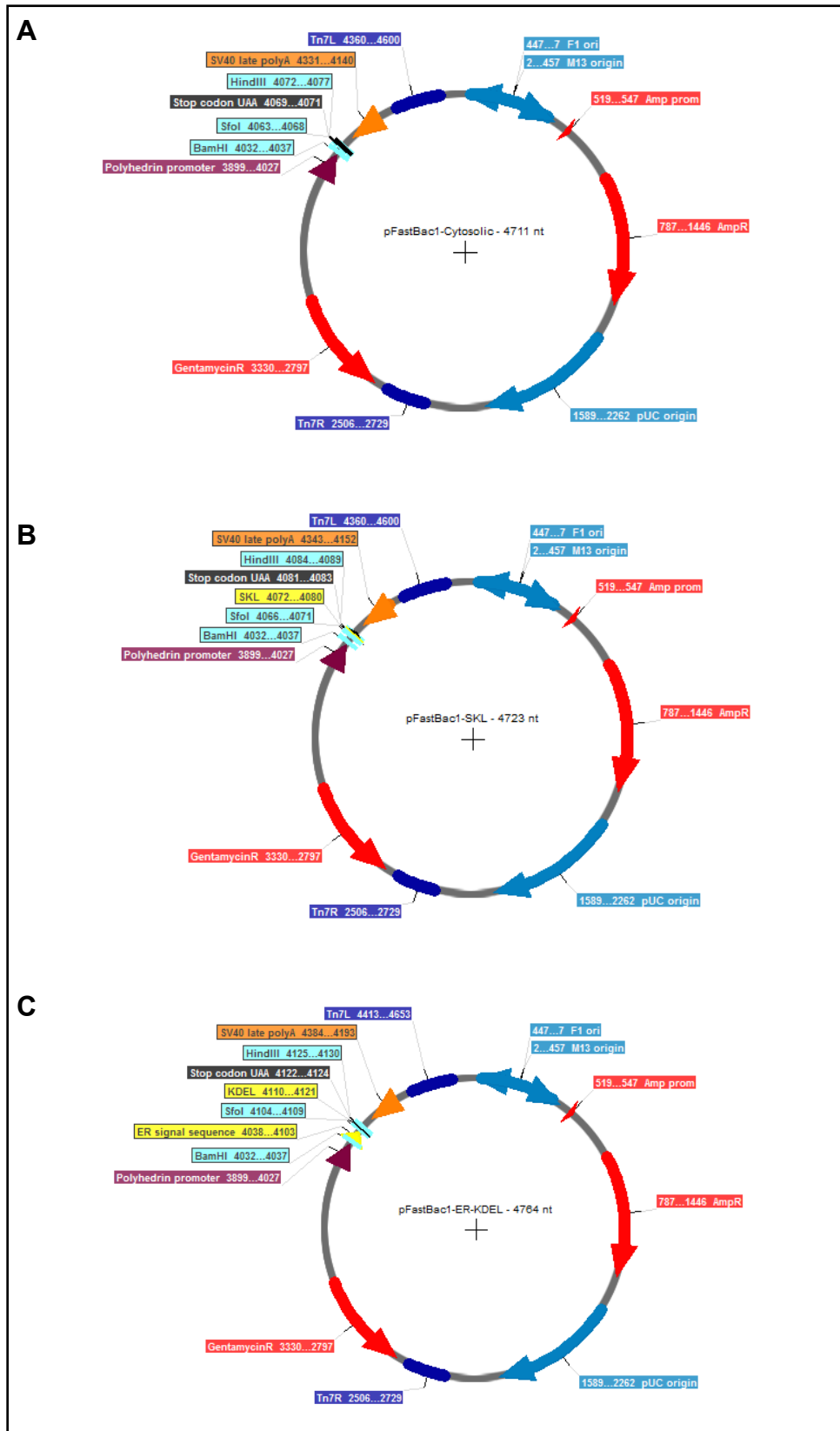
Seq_1 301     CAAGAAATGGCTGCCACAAAGATGTGCAACTTTTGCTAGCAAATTC AAGGCCAAGTGG A 360
          |||
Seq_2 290     CAAGAAATGGCTGCCACAAAGATGTGCAACTTTTGCTAGCAAATTC AAGGCCAAGTGG A 349
          K K W L P Q R C A T F A S K I Q G Q V D

Seq_1 361     CAAAATAAAGGGTGCCGGTGGTGATGGAGCCACCCGAGTTCGAGAAGTAGGGATCCAA 420
          |||
Seq_2 350     CAAAATAAAGGGTGCCGGTGGTGATGGAGCCACCCGAGTTCGAGAAGTAGGGATCCAA 409
          K I K G A G G D W S H P Q F E K * G S K
    
```

(Supplementary Figure 16)

Seq_1	421	GCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCAT	480
Seq_2	410	GCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCAT	469
		L N * L S L D S C * * I Q * * P Q N S I	
Seq_1	481	CTGGATTTGTTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAGAATCCAAGC	540
Seq_2	470	CTGGATTTGTTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAGAATCCAAGC	529
		W I C S E R S V A A G R F L L V R I Q A	
Seq_1	541	TAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAATCACTGGATA	600
Seq_2	530	TAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAATCACTGGATA	589
		S L A R F S G A K E A K M E K K I T G Y	
Seq_1	601	TACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGT	660
Seq_2	590	TACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGT	649
		T T V D I S Q W H R K E H F E A F Q S V	
Seq_1	661	TGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTAAAGACCGT	720
Seq_2	650	TGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTAAAGACCGT	709
		A Q C T Y N Q T V Q L D I T A F L K T V	
Seq_1	721	AAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCTGATGAA	780
Seq_2	710	AAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCTGATGAA	769
		K K N K H K F Y P A F I H I L A R L M N	
Seq_1	781	TGCTCATCCGGAATTTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGT	840
Seq_2	770	TGCTCATCCGGAATTTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGT	829
		A H P E F R M A M K D G E L V I W D S V	
Seq_1	841	TCACCTTGTTACACCGTTTTCCATGAACAACTGAAACGTTTTTCATCGCTCTGGAATGA	900
Seq_2	830	TCACCTTGTTACACCGTTTTCCATGAGCAAACCTGAAACGTTTTTCATCGCTCTGGAGTG-	888
		H P C Y T V F H E Q T E T F S S L W S D	
Seq_1	901	ATACCACGACGATTTCCGGCAGTTTCTACCATATATCCCAAGATGTGGCGGGTACGGTA	960
Seq_2	889	ATACCACGACGATTTCCGGCAGTT-CTACCATA-----	920
		T T T I S G S S T I	

**Supplementary Figure 16. Alignment between theoretical coding nucleotide sequence and the sequencing result of pQE30GLucD2 construct (clone #125) for expression in *E. coli* cells**



**Supplementary Figure 17. Circular map of plasmid vector pFastBac1 for Baculovirus expression in insect cells for *in vivo* crystallization method**

A) pFastBac1-Cytosolic (4711 nt), B) pFastBac1-SKL (4723 nt) and C) pFastBac1-ER-KDEL (4764 nt) vector maps (Invitrogen, USA). The graphics show the restriction sites specific for cloning. The vector imparts ampicillin and gentamycin resistance genes. It contains a polyhedrin promoter P(PH), allows a high-level of expression of the recombinant protein in insect cells.

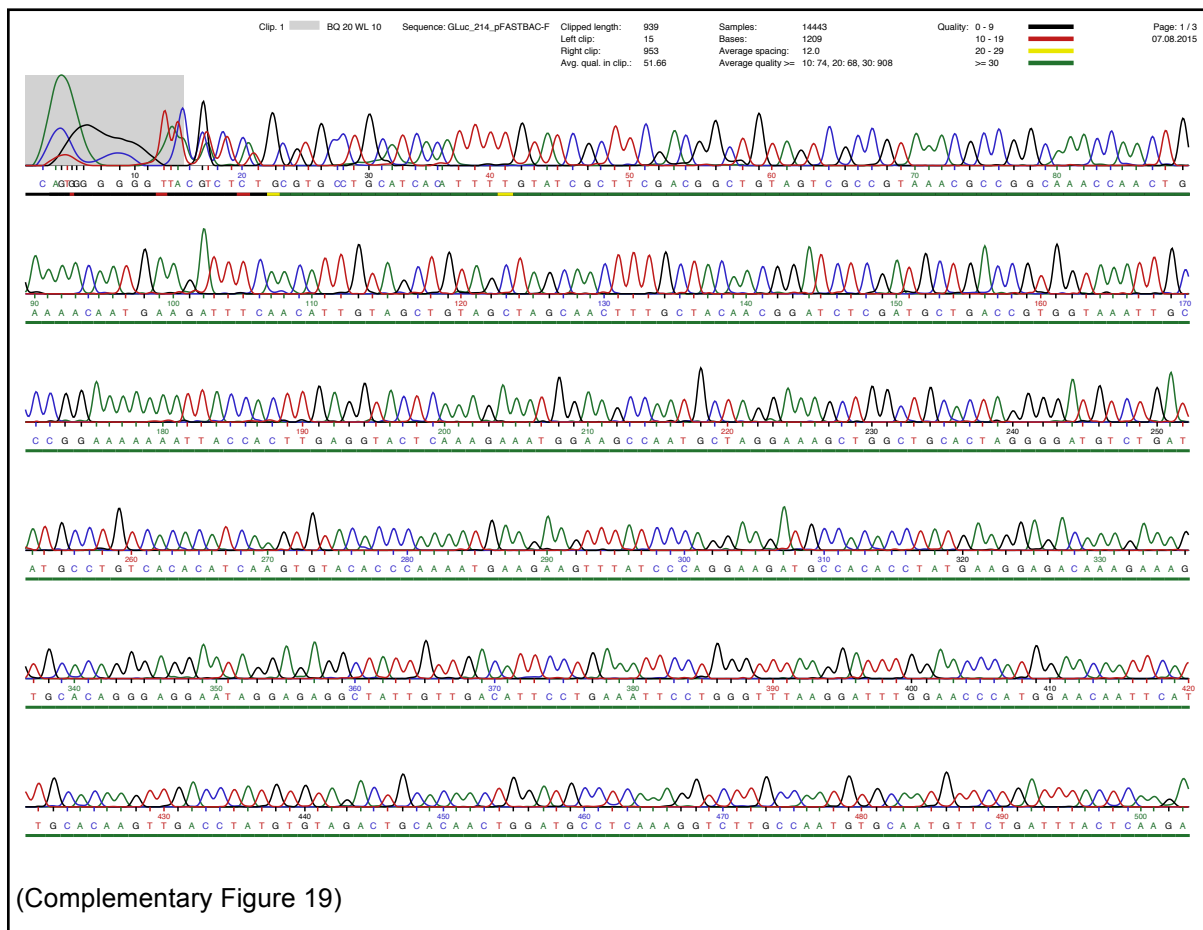
GLucFL nucleotide (510 bp)

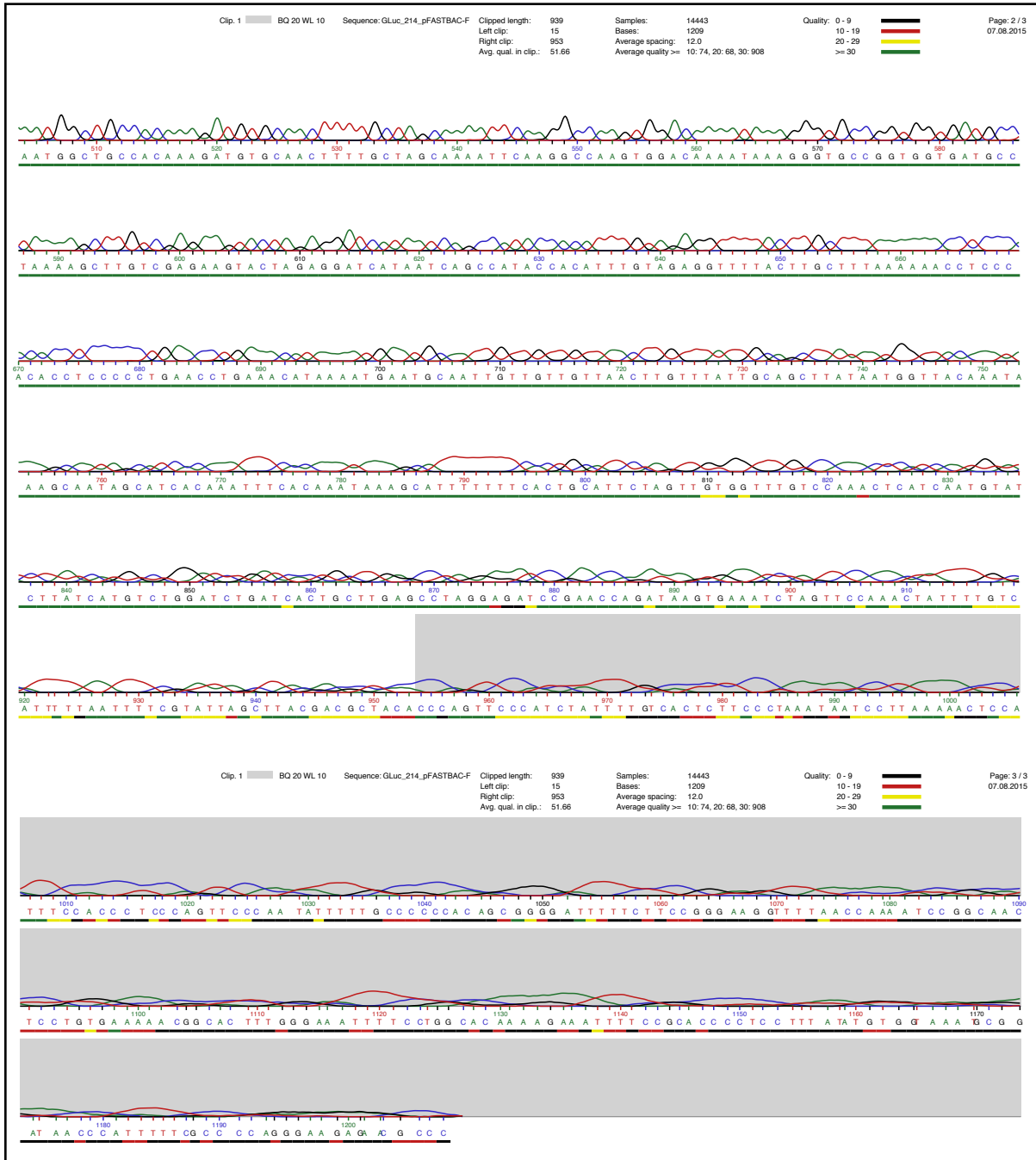
ATGAAACCAACTGAAAAACAATGAAGATTTCAACATTGTAGCTGTAGCTAGCAACTTTGCTACAACGGATCTCGAT  
 GCTGACCGTGGTAAATTGCCCGGAAAAAATTACCACCTTGAGGTACTCAAAGAAATGGAAGCCAATGCTAGGAAA  
 GCTGGCTGCACTAGGGGATGTCTGATATGCCTGTACACATCAAGTGTACACCCAAAATGAAGAAGTTTATCCCA  
 GGAAGATGCCACACCTATGAAGGAGACAAAAGAAAGTGCACAGGGAGGAATAGGAGAGGGCTATTGTTGACATTCC  
 GAAATTCCTGGGTTAAGGATTTGGAACCCATGGAACAATTCATTGCACAAGTTGACCTATGTGTAGACTGCACA  
 ACTGGATGCCTCAAAGGCTTTGCCAATGTGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAAGATGTGCA  
 ACTTTTGTAGCAAAAATTC AAGGCCAAGTGGACAAAATAAAGGGTGCCGGTGGTGATTAG

GLucFL amino acid sequence (169 aa)

MKPTENNEDFNIVAVASNFAATDLDADRGLKLPKGLPLEVLKEMEANARKAGCTRGCLICLSHIKCTPKMKKFIP  
 GRCHTYEGDKESAQGGIGEAIVDIPDIPGFKDLPEMQFIAQVDLCVDCCTTGCLKGLANVQCSDLLKKWLPQRCA  
 TFASKIQGVQDKIKGAGGD

**Supplementary Figure 18. Full-length GLuc sequence for cloning experiments in pFastBac1 vector for crystallization *in vivo* using Sf21 insect cells.**





**Supplementary Figure 19. DNA Chromatogram sequencing result from (clone #214) of pFastBac1-ERGLucFl construct for crystallization *in vivo* using Sf21 insect cells**

Alignment of Sequence\_1: [GLuc\_214\_pFASTBAC-F.seq] with Sequence\_2:  
[LigationGLucpFastBaclER.xdna]

Similarity : 1059/5256 (20,15 %)

```

Seq_1 1 ----- 0
Seq_2 1 AAAATGATAACCATCTCGCAAATAAATAAGTATTTACTGTTTTCGTAACAGTTTTGTAA 60
      N D N H L A N K * V F Y C F R N S F V I

Seq_1 1 -----CA-GTGGGGGGTACG 16
Seq_2 61 TAAAAAACCTATAAATATTCGGGATTATTCATACCGTCCCACCATCGGGCGGgatcca 120
      K K P I N I P D Y S Y R P T I G R G S M

Seq_1 17 TCTC----TGCGTGCCTGCATCACATTTTGTATCGCTTCGACGGCTGTAGTCGCCGTAA 71
      |
Seq_2 121 tgcattctcatgcgtgcctgcatcacatTTTgtatcgcttcgacggctgtagtcgccgtaa 180
      H L M R A C I T F C I A S T A V V A V N

Seq_1 72 ACGCCGGCAAACCAACTGAAAACAATGAAGATTTCAACATTGTAGCTGTAGCTAGCAACT 131
      |
Seq_2 181 acgccggcAAACCAACTGAAAACAATGAAGATTTCAACATTGTAGCTGTAGCTAGCAACT 240
      A G K P T E N N E D F N I V A V A S N F

Seq_1 132 TTGCTACAACGGATCTCGATGCTGACCGTGGTAAATTGCCCGGAAAAAATTACCACCTTG 191
      |
Seq_2 241 TTGCTACAACGGATCTCGATGCTGACCGTGGTAAATTGCCCGGAAAAAATTACCACCTTG 300
      A T T D L D A D R G K L P G K K L P L E

Seq_1 192 AGGTACTCAAAGAAATGGAAGCCAATGCTAGGAAAGCTGGCTGCACTAGGGGATGTCTGA 251
      |
Seq_2 301 AGGTACTCAAAGAAATGGAAGCCAATGCTAGGAAAGCTGGCTGCACTAGGGGATGTCTGA 360
      V L K E M E A N A R K A G C T R G C L I

Seq_1 252 TATGCCTGTACACATCAAGTGTACACCCAAAATGAAGAAGTTTATCCCAGGAAGATGCC 311
      |
Seq_2 361 TATGCCTGTACACATCAAGTGTACACCCAAAATGAAGAAGTTTATCCCAGGAAGATGCC 420
      C L S H I K C T P K M K K F I P G R C H

Seq_1 312 ACACCTATGAAGGAGACAAAGAAAGTGACAGGGAGGAATAGGAGAGGCTATTGTTGACA 371
      |
Seq_2 421 ACACCTATGAAGGAGACAAAGAAAGTGACAGGGAGGAATAGGAGAGGCTATTGTTGACA 480
      T Y E G D K E S A Q G G I G E A I V D I

Seq_1 372 TTCCTGAAATTCCTGGGTTTAAGGATTTGGAACCCATGGAACAATTCATTGCACAAGTTG 431
      |
Seq_2 481 TTCCTGAAATTCCTGGGTTTAAGGATTTGGAACCCATGGAACAATTCATTGCACAAGTTG 540
      P E I P G F K D L E P M E Q F I A Q V D

Seq_1 432 ACCTATGTGTAGACTGCACAACCTGGATGCCTCAAAGGTCTTGCCAATGTGCAATGTTCTG 491
      |
Seq_2 541 ACCTATGTGTAGACTGCACAACCTGGATGCCTCAAAGGTCTTGCCAATGTGCAATGTTCTG 600
      L C V D C T T G C L K G L A N V Q C S D

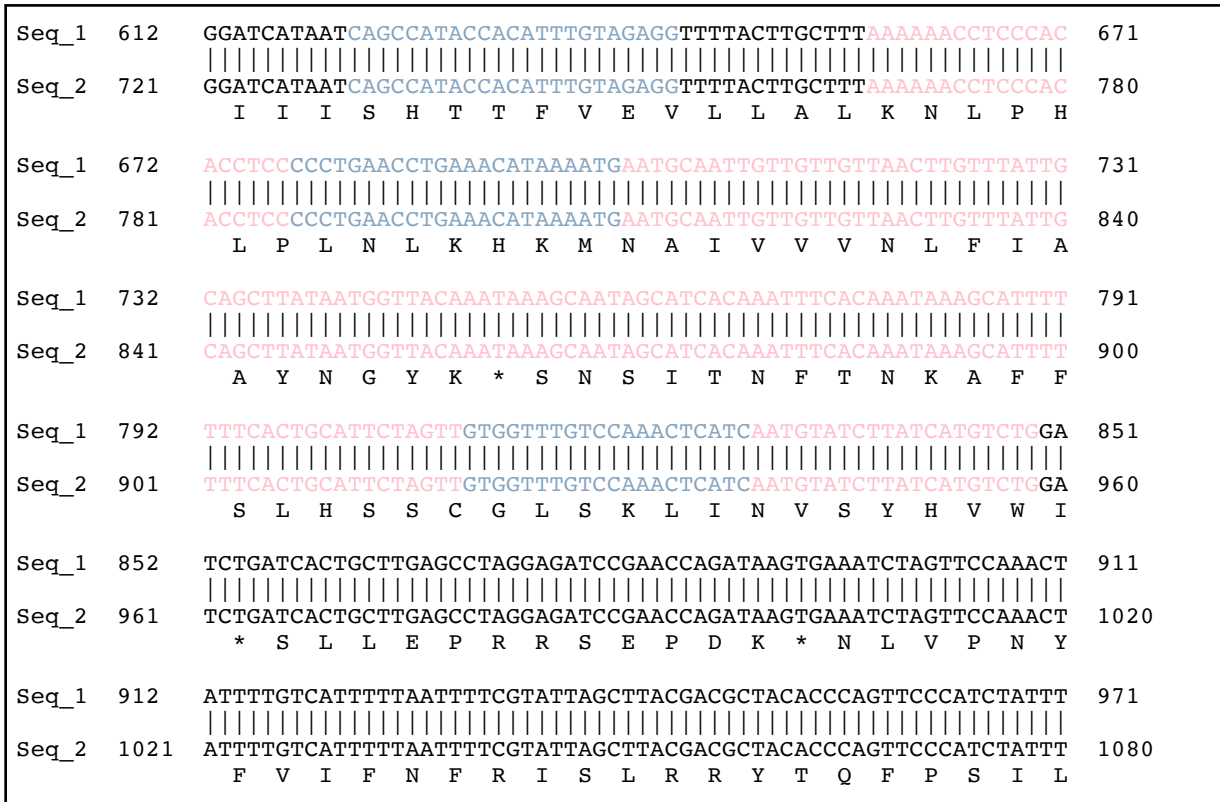
Seq_1 492 ATTTACTCAAGAAATGGCTGCCACAAAGATGTGCAACTTTTGCTAGCAAAATTC AAGGCC 551
      |
Seq_2 601 ATTTACTCAAGAAATGGCTGCCACAAAGATGTGCAACTTTTGCTAGCAAAATTC AAGGCC 660
      L L K K W L P Q R C A T F A S K I Q G Q

Seq_1 552 AAGTGGACAAAATAAAGGGTGCCGGTGGTGTATGCCTAAAAGCTTGTGCGAGAAGTACTAGA 611
      |
Seq_2 661 AAGTGGACAAAATAAAGGGTGCCGGTGGTGTATgcctaaaagctTGTGCGAGAAGTACTAGA 720
      V D K I K G A G G D A * K L V E K Y * R

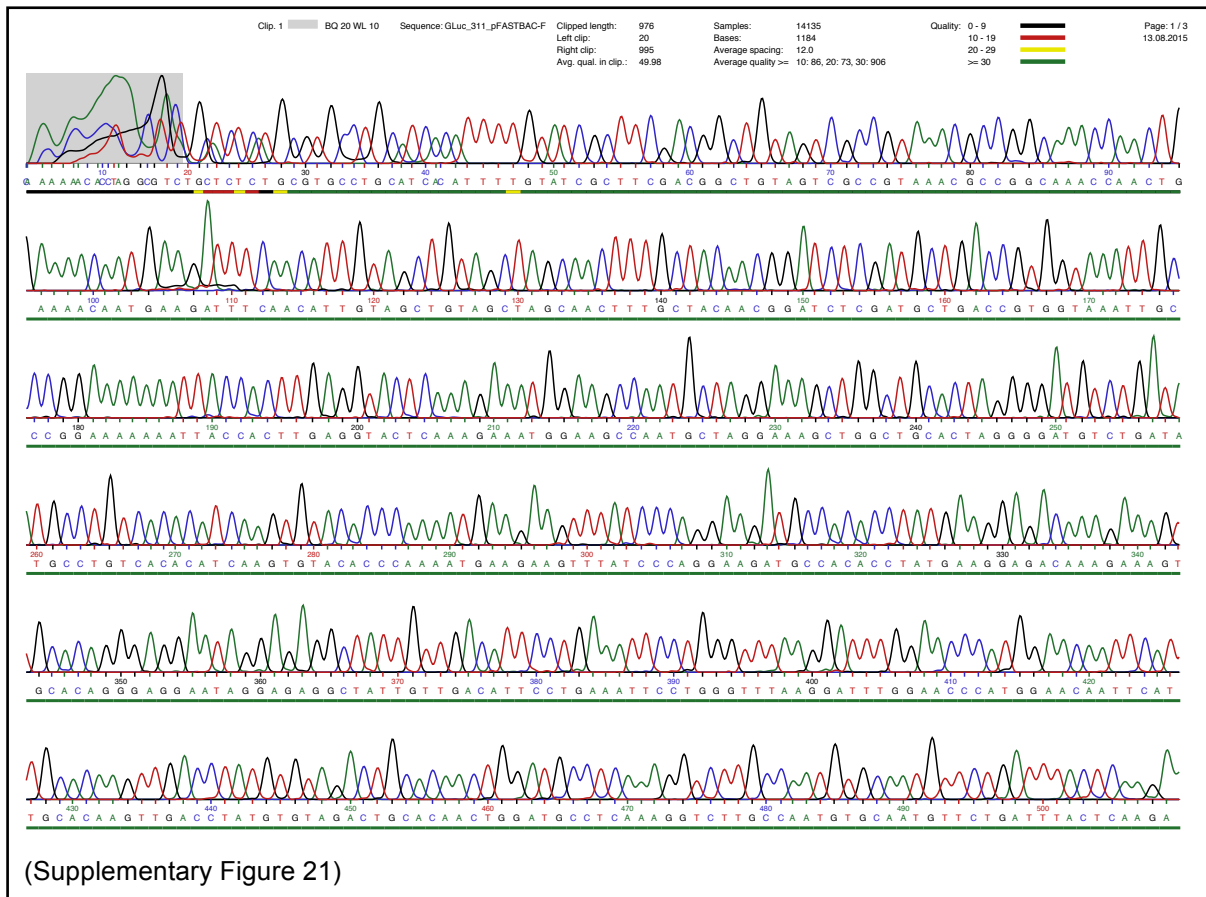
```

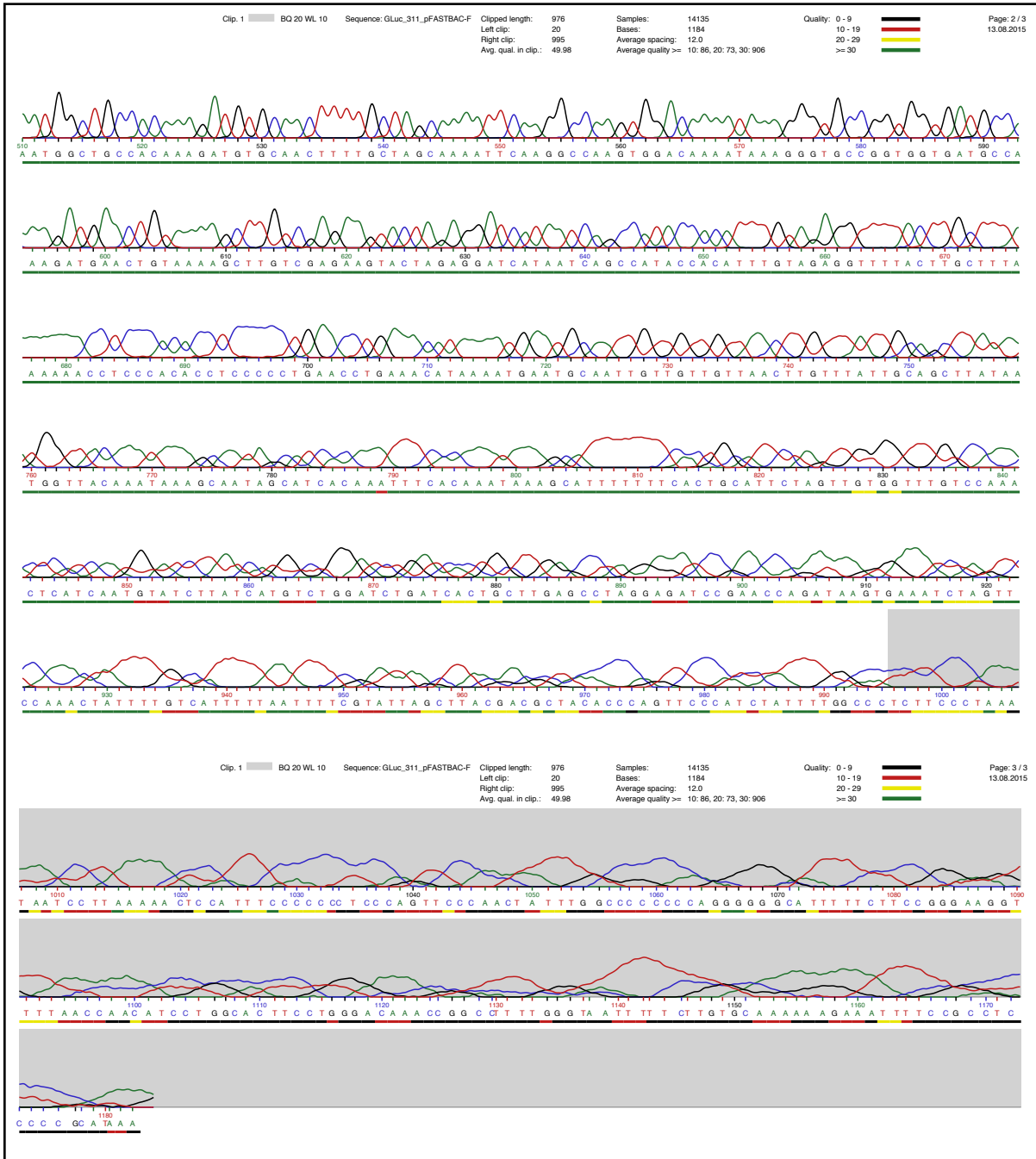
(Supplementary figure 20)





**Supplementary Figure 20. Alignment between theoretical coding nucleotide sequence and the sequencing result of pFastBac1-ERGLucFl construct (clone #214) for crystallization *in vivo* using Sf21 insect cells**





Supplementary Figure 21. Chromatogram sequencing result from (clone #311\_1) of pFastBac1-ER-KDELGLucFl construct for crystallization *in vivo* using Sf21 insect cells

Alignment of Sequence\_1: [Ligation GLucpFastBac1\_ER-KDEL.xdna] with Sequence\_2: [GLuc\_311\_1\_Polyhed\_fwd.seq]

Similarity : 1020/5268 (19,36 %)

```

Seq_1  1  -----TAAATATTCGGGATTATTCA  20
          |||
Seq_2  1  AATCACTGTTTTCTGACAGTTTTGTAATAAAAAAACCTATAAATATTCGGGATTATTCA  60
          N S L F S Y S F V I K K P I N I P D Y S

Seq_1  21  TACCGTCCCACCATCGGGCGGgatccatgcatctcatgctgcctgcatcacattttgt  80
          |||
Seq_2  61  TACCGTCCCACCATCGGGCGGGATCCATGCATCTCATGCGTGCCTGCATCACATTTTGT  120
          Y R P T I G R G S M H L M R A C I T F C

Seq_1  81  atcgcttcgacggctgtagtcgccgtaaacgccggcAAACCAACTGAAAAAATGAAGAT  140
          |||
Seq_2  121  ATCGCTTCGACGGCTGTAGTCGCCGTAAACGCCGGCAAACCAACTGAAAAAATGAAGAT  180
          I A S T A V V A V N A G K P T E N N E D

Seq_1  141  TTCAACATTGTAGCTGTAGCTAGCAACTTTGCTACAACGGATCTCGATGCTGACCGTGGT  200
          |||
Seq_2  181  TTCAACATTGTAGCTGTAGCTAGCAACTTTGCTACAACGGATCTCGATGCTGACCGTGGT  240
          F N I V A V A S N F A T T D L D A D R G

Seq_1  201  AAATGCCCCGAAAAAAATTACCACCTTGAGGTACTIONAAAGAAATGGAAGCCAATGCTAGG  260
          |||
Seq_2  241  AAATGCCCCGAAAAAAATTACCACCTTGAGGTACTIONAAAGAAATGGAAGCCAATGCTAGG  300
          K L P G K K L P L E V L K E M E A N A R

Seq_1  261  AAAGTGGCTGCAC TAGGGGATGTCTGATATGCCTGTCACACATCAAGTGTACACCCAAA  320
          |||
Seq_2  301  AAAGTGGCTGCAC TAGGGGATGTCTGATATGCCTGTCACACATCAAGTGTACACCCAAA  360
          K A G C T R G C L I C L S H I K C T P K

Seq_1  321  ATGAAGAAGTTTATCCAGGAAGATGCCACACCTATGAAGGAGACAAAGAAAGTGCACAG  380
          |||
Seq_2  361  ATGAAGAAGTTTATCCAGGAAGATGCCACACCTATGAAGGAGACAAAGAAAGTGCACAG  420
          M K K F I P G R C H T Y E G D K E S A Q

Seq_1  381  GGAGGAATAGGAGAGGCTATTGTTGACATTCCTGAAATTCCTGGGTTTAAGGATTTGGAA  440
          |||
Seq_2  421  GGAGGAATAGGAGAGGCTATTGTTGACATTCCTGAAATTCCTGGGTTTAAGGATTTGGAA  480
          G G I G E A I V D I P E I P G F K D L E

Seq_1  441  CCCATGGAACAATTCATTGCACAAGTTGACCTATGTGTAGACTGCACAACCTGGATGCCTC  500
          |||
Seq_2  481  CCCATGGAACAATTCATTGCACAAGTTGACCTATGTGTAGACTGCACAACCTGGATGCCTC  540
          P M E Q F I A Q V D L C V D C T T G C L

Seq_1  501  AAAGTCTTGCCAATGTGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAAGATGT  560
          |||
Seq_2  541  AAAGTCTTGCCAATGTGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAAGATGT  600
          K G L A N V Q C S D L L K K W L P Q R C

Seq_1  561  GCAACTTTTGTAGCAAAAATTCAAGGCCAAGTGGACAAAATAAAGGGTGCCGGTGGTGAT  620
          |||
Seq_2  601  GCAACTTTTGTAGCAAAAATTCAAGGCCAAGTGGACAAAATAAAGGGTGCCGGTGGTGAT  660
          A T F A S K I Q G Q V D K I K G A G G D

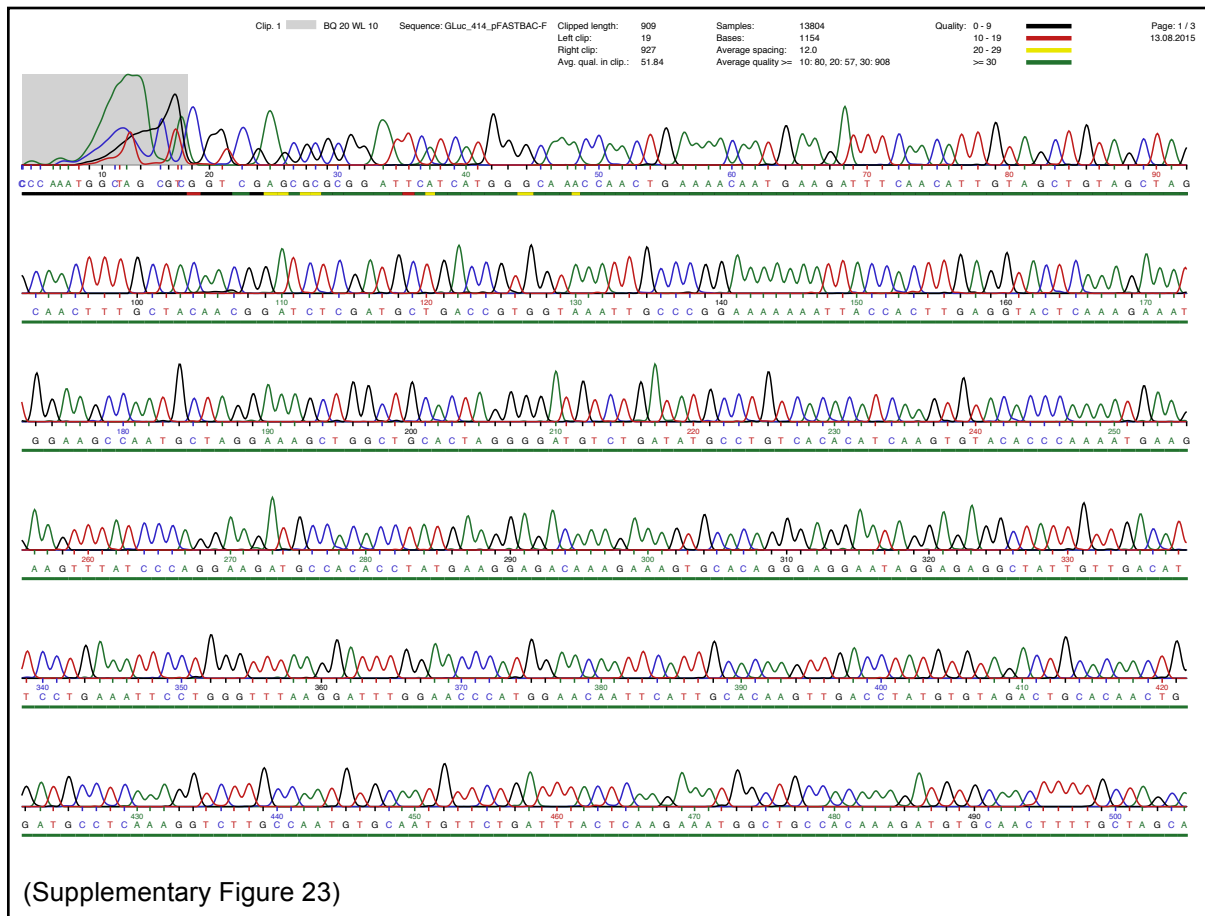
Seq_1  621  gccaaagatgaactgtaaaagctTGTCGAGAAGTACTAGAGGATCATAATCAGCCATAACC  680
          |||
Seq_2  661  GCCAAAGATGAACGTAAAAGCTTGTCGAGAAGTACTAGAGGATCATAATCAGCCATAACC  720
          A K D E L * K L V E K Y * R I I I S H T

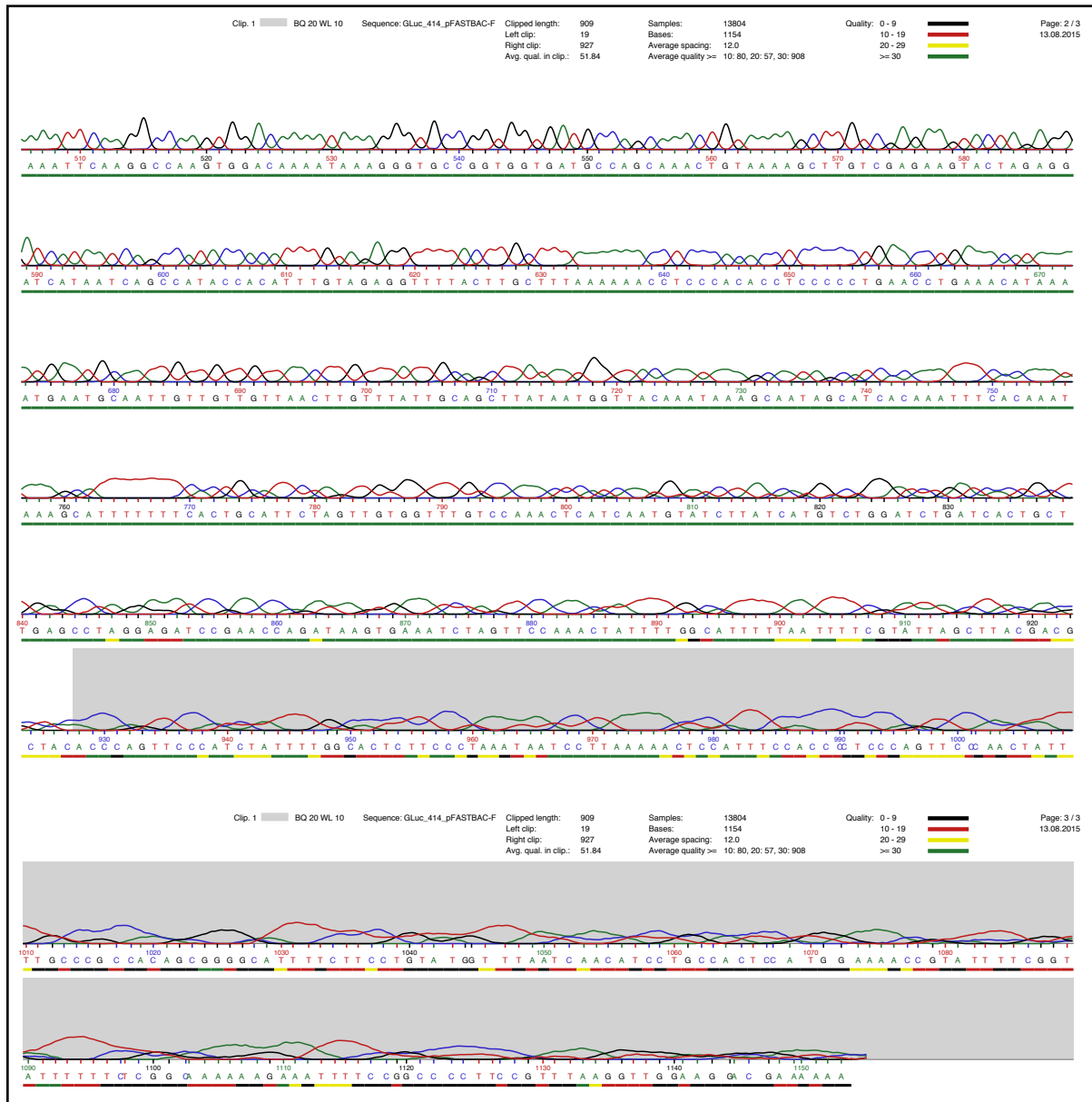
```

(Supplementary Figure 22)

Seq_1	681	ACATTGTAGAGGTTTACTTGCTTTAAAAACCTCCCACACCTCCCCCTGAACCTGAAA	740
Seq_2	721	ACATTGTAGAGGTTTACTTGCTTTAAAAACCTCCCACACCTCCCCCTGAACCTGAAA	780
		T F V E V L L A L K N L P H L P L N L K	
Seq_1	741	CATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA	800
Seq_2	781	CATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA	840
		H K M N A I V V V N L F I A A Y N G Y K	
Seq_1	801	TAAAGCAATAGCATCACAAAATTTACAAAATAAAGCATTTTTTTTCAGTGCATTCTAGTTGT	860
Seq_2	841	TAAAGCAATAGCATCACAAAATTTACAAAATAAAGCATTTTTTTTCAGTGCATTCTAGTTGT	900
		* S N S I T N F T N K A F F S L H S S C	
Seq_1	861	GGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTCGATCTGATCACTGCTTGAGCCT	920
Seq_2	901	GGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTCGATCTGATCACTGCTTGAGCCT	960
		G L S K L I N V S Y H V W I * S L L E P	

**Supplementary Figure 22. Alignment between theoretical coding nucleotide sequence and the sequencing result of pFastBac1-ER-KDELGLucFl construct (clone #311\_1) for crystallization *in vivo* using Sf21 insect cells**





**Supplementary Figure 23. DNA Chromatogram sequencing result from (clone #414) of pFastBac1-SKLGLucFl construct for crystallization *in vivo* using *Sf21* insect cells**

Alignment of Sequence\_1: [GLuc\_414\_pFASTBAC-F.seq] with Sequence\_2: [Ligation GLucpFastBac1\_SKL.xdna]

Similarity : 1051/5227 (20,11 %)

```

Seq_1 1 -----CCCCAAAT-GGCTAGCGTCCGGTCGAGC 26
                               | | | |
Seq_2 1 TATAAATATTCCGGATTATTCATACCGTCCCACCATCGGGCGCGgatcccggtccgaagc 60
      * I F R I I H T V P P S G A D P G P K R

Seq_1 27 GCGCGG-ATTCATCATGGGCAAACCAACTGAAAACAATGAAGATTTCAACATTGTAGCTG 85
      |||||
Seq_2 61 gcgcggaattcatcatgggcAAACCAACTGAAAACAATGAAGATTTCAACATTGTAGCTG 120
      A E F I M G K P T E N N E D F N I V A V

Seq_1 86 TAGCTAGCAACTTTGCTACAACGGATCTCGATGCTGACCGTGGTAAATTGCCCGGAAAAA 145
      |||||
Seq_2 121 TAGCTAGCAACTTTGCTACAACGGATCTCGATGCTGACCGTGGTAAATTGCCCGGAAAAA 180
      A S N F A T T D L D A D R G K L P G K K

Seq_1 146 AATTACCACTTGAGGTACTCAAAGAAATGGAAGCCAATGCTAGGAAAGCTGGCTGCACTA 205
      |||||
Seq_2 181 AATTACCACTTGAGGTACTCAAAGAAATGGAAGCCAATGCTAGGAAAGCTGGCTGCACTA 240
      L P L E V L K E M E A N A R K A G C T R

Seq_1 206 GGGGATGTCTGATATGCCTGTCACACATCAAGTGTACACCCAAAATGAAGAAGTTTATCC 265
      |||||
Seq_2 241 GGGGATGTCTGATATGCCTGTCACACATCAAGTGTACACCCAAAATGAAGAAGTTTATCC 300
      G C L I C L S H I K C T P K M K K F I P

Seq_1 266 CAGGAAGATGCCACACCTATGAAGGAGACAAAGAAAGTGCACAGGGAGGAATAGGAGAGG 325
      |||||
Seq_2 301 CAGGAAGATGCCACACCTATGAAGGAGACAAAGAAAGTGCACAGGGAGGAATAGGAGAGG 360
      G R C H T Y E G D K E S A Q G G I G E A

Seq_1 326 CTATTGTTGACATTCTGAAATTCTGGGTTTAAGGATTTGGAACCCATGGAACAATTCA 385
      |||||
Seq_2 361 CTATTGTTGACATTCTGAAATTCTGGGTTTAAGGATTTGGAACCCATGGAACAATTCA 420
      I V D I P E I P G F K D L E P M E Q F I

Seq_1 386 TTGCACAAGTTGACCTATGTGTAGACTGCACAACCTGGATGCCTCAAAGGTCTTGCCAATG 445
      |||||
Seq_2 421 TTGCACAAGTTGACCTATGTGTAGACTGCACAACCTGGATGCCTCAAAGGTCTTGCCAATG 480
      A Q V D L C V D C T T G C L K G L A N V

Seq_1 446 TGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAAGATGTGCAACTTTTGTAGCA 505
      |||||
Seq_2 481 TGCAATGTTCTGATTTACTCAAGAAATGGCTGCCACAAAGATGTGCAACTTTTGTAGCA 540
      Q C S D L L K K W L P Q R C A T F A S K

Seq_1 506 AAATTCAAGCCAAGTGACAAAATAAAGGGTGCCGGTGGTGATGCCAGCAAAGTGTAAA 565
      |||||
Seq_2 541 AAATTCAAGCCAAGTGACAAAATAAAGGGTGCCGGTGGTGATgccagcaaaactgtaaa 600
      I Q G Q V D K I K G A G G D A S K L * K

Seq_1 566 AGCTTGTGCGAGAAGTACTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTAC 625
      |||||
Seq_2 601 agctTGTCGAGAAGTACTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTAC 660
      L V E K Y * R I I I S H T T F V E V L L

Seq_1 626 TTGCTTTAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTG 685
      |||||
Seq_2 661 TTGCTTTAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTG 720
      A L K N L P H L P L N L K H K M N A I V

```

(Supplementary Figure 24)

Seq_1	686	TTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAA	745
Seq_2	721	TTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAA	780
		V  V  N  L  F  I  A  A  Y  N  G  Y  K  *  S  N  S  I  T  N	
Seq_1	746	ATTTCCAAATAAAGCATTFTTTTCTACTGCATTCTAGTTGTGGTTTGTCCAACTCATCA	805
Seq_2	781	ATTTCCAAATAAAGCATTFTTTTCTACTGCATTCTAGTTGTGGTTTGTCCAACTCATCA	840
		F  T  N  K  A  F  F  S  L  H  S  S  C  G  L  S  K  L  I  N	
Seq_1	806	ATGTATCTTATCATGTCTGGATCTGATCACTGCTTGAGCCTAGGAGATCCGAACCAGATA	865
Seq_2	841	ATGTATCTTATCATGTCTGGATCTGATCACTGCTTGAGCCTAGGAGATCCGAACCAGATA	900
		V  S  Y  H  V  W  I  *  S  L  L  E  P  R  R  S  E  P  D  K	
Seq_1	866	AGTGAATCTAGTTCCAAATATTTTGGCATTFTTTAATTTTCGTATTAGCTTACGACGCT	925
Seq_2	901	AGTGAATCTAGTTCCAAATATTTTGGCATTFTTTAATTTTCGTATTAGCTTACGACGCT	960
		*  N  L  V  P  N  Y  F  V  I  F  N  F  R  I  S  L  R  R  Y	
Seq_1	926	ACACCAGTTCCTATCTATTTTGGCACTCTTCCCTAAATAATCCTTAAAACTCCATTTTC	985
Seq_2	961	ACACCAGTTCCTATCTATTTTGGCACTCTTCCCTAAATAATCCTTAAAACTCCATTTTC	1020
		T  Q  F  P  S  I  L  S  L  F  P  K  *  S  L  K  T  P  F  P	
Seq_1	986	CACCCCTCCAGTTCCTAACTATTTGCCCCG-CCACAGCGGGGCATTTT-----	1033
Seq_2	1021	CACCCCTCCAGTTCCTAACTATTTGCCCCGCCCCACAGCGGGGCATTTTCTTCCTGTTA	1080
		P  L  P  V  P  N  Y  F  V  R  P  Q  R  G  I  F  L  P  V  M	

**Supplementary Figure 24. Alignment between theoretical coding nucleotide sequence and the sequencing result of pFastBac1-SKLGLucFI construct (clone #414) for crystallization *in vivo* using Sf21 insect cells**