

pUNO1-Spike-dfur

Expression vector containing the Wuhan-Hu-1 D614 SARS-CoV-2 Spike (delta furin) gene

Catalog code: p1-spike-df

<https://www.invivogen.com/sars2-spike-d614-expression-vectors>

For research use only

Version 21C15-ED

PRODUCT INFORMATION

Contents

- 20 µg of lyophilized pUNO1-Spike-dfur (plasmid DNA)
- 2 x 1 ml of Blasticidin (10 mg/ml)

Storage and Stability

- Product is shipped at room temperature.
- Lyophilized DNA should be stored at -20°C.
- Resuspended DNA should be stored at -20°C and is stable for at least 1 year.
- Store Blasticidin at 4°C or -20°C. The expiry date is specified on the product label.

Quality control

- Plasmid construct is confirmed by restriction analysis and full-length open reading frame (ORF) sequencing.
- After purification by ion exchange chromatography, predominant supercoiled conformation is verified by electrophoresis.

PLASMID FEATURES

Wuhan-Hu-1 SARS-CoV-2 Spike (D614) cassette

• **EF-1α/HTLV hybrid promoter** is a composite promoter comprised of the Elongation Factor-1α (EF-1α) core promoter¹ and the 5' untranslated region of the Human T-Cell Leukemia Virus (HTLV). EF-1α utilizes a type 2 promoter that encodes a "house-keeping" gene. It is expressed at high levels in all cell cycles and lower levels during the G0 phase. Additionally, since the promoter is not tissue-specific it is highly expressed in all cell types. The R segment and part of the U5 sequence (R-U5') of the HTLV Type 1 Long Terminal Repeat² has been coupled to the EF-1α promoter to enhance stability of DNA and RNA. This modification not only increases steady state transcription, but also significantly increases translation efficiency.

- **Codon-optimized Spike ORF**

pUNO1-Spike-dfur contains the Spike (S) coding sequence from the original Wuhan-Hu-1 SARS-CoV-2 isolate. Notably, the furin cleavage site has been inactivated (dfur) by the inclusion of two mutations (R683/5A). Furthermore; to improve expression of the S protein in cell lines, the gene is codon-optimized and the last 19 amino acids, which contain an endoplasmic reticulum (ER)-retention motif (KxHxx), have been removed^{3,4}.

pUNO1-Spike-dfur is characterized by the following features:

- **S1 domain:** D614
- **S1/S2 boundary:** R683A, R685A

Spike (S) is a structural glycoprotein expressed on the surface of SARS-CoV-2. It mediates membrane fusion and viral entry into target cells upon binding to the host receptor ACE2 and the proteolytic activity of host proteases such as furin and TMPRSS2⁵.

For more information visit: <https://www.invivogen.com/sars2-spike>

- **SV40 pAn** is the Simian Virus 40 late polyadenylation (pAn) signal and it enables efficient cleavage and polyadenylation reactions resulting in high levels of steady-state mRNA⁶.

Antibiotic selection cassette

- **hCMV (human cytomegalovirus) enhancer & promoter** drive the expression of the blasticidin resistance gene (*bsr*) in mammalian cells.
- **EM7** is a bacterial promoter that enables the constitutive expression of the blasticidin resistance gene (*bsr*) in *E. coli*.
- ***bsr* (blasticidin resistance gene)** encodes a deaminase from *Bacillus cereus* that confers resistance to the antibiotic blasticidin. The expression of the *bsr* gene is driven by the CMV promoter/enhancer and the bacterial EM7 promoter. Therefore, **Blasticidin** can be used to select stable clones in mammalian cells and *E. coli* transformants.
- **Human β-Globin pAn** is a strong polyadenylation (pAn) signal placed downstream of *bsr*. The use of β-globin pAn minimizes interference and possible recombination events with the SV40 pAn signal⁷.

General features of pUNO1-Spike-dfur

- **pMB1 ori** is a minimal *E. coli* origin of replication.

APPLICATIONS

Stable gene expression in mammalian cells.

pUNO1 plasmids are designed for both transient and stable transfection in mammalian cell lines by selection with **Blasticidin**. Furthermore, they facilitate high levels of expression of the gene of interest.

Antibody screening by flow cytometry

pUNO1-Spike-dfur has been specifically designed for mammalian cell expression of the SARS-CoV-2 S protein. Notably, due to the inactivated furin cleavage site, when this plasmid is expressed by a host cell (e.g. 293T cells) there is high surface expression of the full-length S protein^{3,8}. Ideal for SARS-CoV-2 S-specific antibody screening by flow cytometry (*in-house data*).

METHODS

Plasmid resuspension

- Quickly spin the tube containing the lyophilized plasmid to pellet the DNA.

- To obtain a plasmid solution at 1 µg/µl, resuspend the DNA in 20 µl of sterile water.

- Store resuspended plasmid at -20°C.

Plasmid amplification and cloning

Plasmid amplification and cloning can be performed in *E. coli* GT116 or other commonly used laboratory *E. coli* strains, such as DH5α.

Blasticidin usage

Blasticidin should be used at 25-100 µg/ml in bacteria and 1-30 µg/ml in mammalian cells. Blasticidin is supplied as a 10 mg/ml colorless solution in HEPES buffer.

TECHNICAL SUPPORT

InvivoGen USA (Toll-Free): 888-457-5873

InvivoGen USA (International): +1 (858) 457-5873

InvivoGen Europe: +33 (0) 5-62-71-69-39

InvivoGen Hong Kong: +852 3622-3480

E-mail: info@invivogen.com

REFERENCES

1. Kim D. *et al.*, 1990. Use of the human elongation factor 1 α promoter as a versatile and efficient expression system. *Gene* 91(2):217-23

2. Takebe Y. *et al.*, 1988. SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol Cell Biol*, 8(1):466-72.

3. Johnson, M.C. *et al.* 2020. Optimized Pseudotyping Conditions for the SARS-CoV-2 Spike Glycoprotein. *J Virol* 94.

4. Ou, X. *et al.* 2020. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* 11, 1620.

5. Hoffmann M. *et al.*, 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 181:1-16.

6. Carswell S. & Alwine J., 1989. Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. *Mol Cell Biol*. 9(10):4248-58.

7. Yu J. & Russell J., 2001. Structural and functional analysis of an mRNP complex that mediates the high stability of human β -globin mRNA. *Mol Cell Biol*. 21(17):5879-88.

8. Walls, A.C. *et al.* 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*.

RELATED PRODUCTS

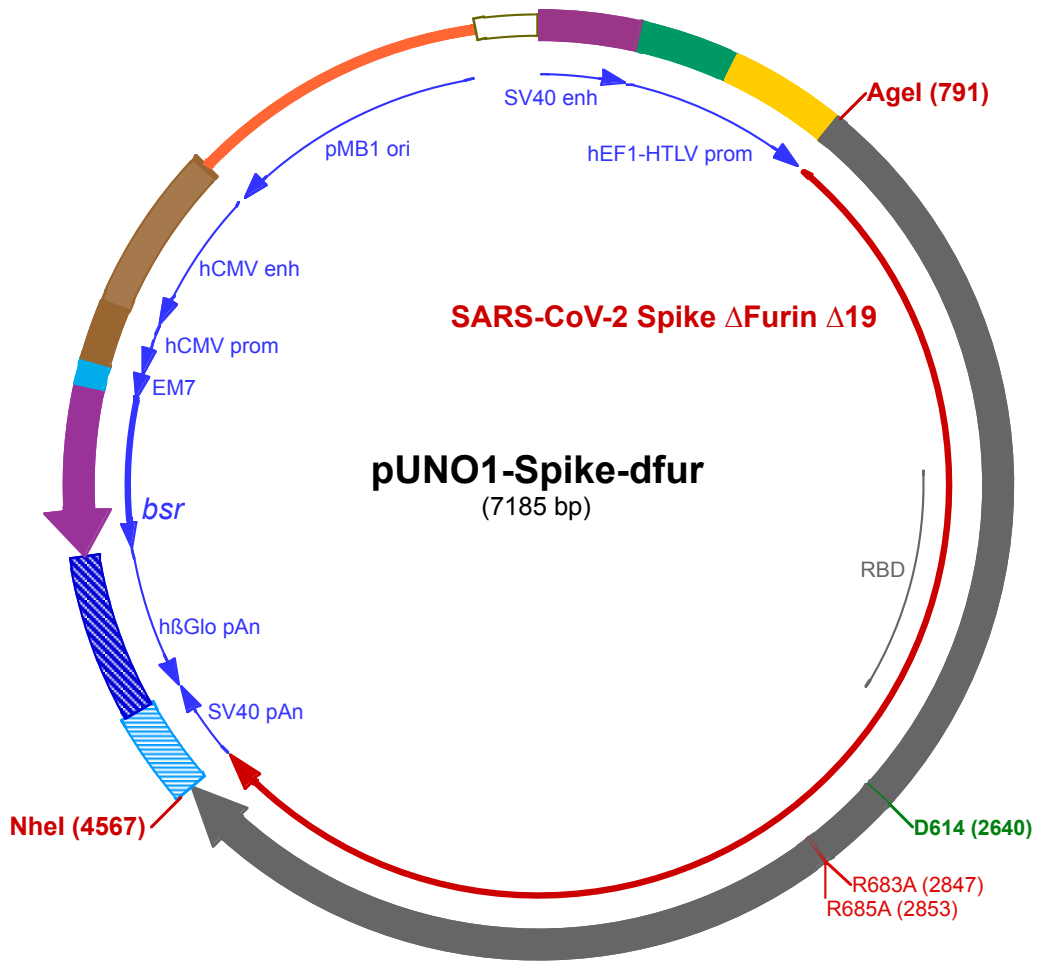
| Product | Description | Cat. Code |
|-------------------------------|--------------------------|-----------------|
| Blasticidin | Selection antibiotic | ant-bl-1 |
| ChemiComp GT116 | Competent <i>E. coli</i> | gt116-11 |
| COVID-19 Product Range | | |
| HEK-Blue™ hACE2 Cells | Cell line | hkb-hace2 |
| A549-hACE2-TMPRSS2 Cells | Cell Line | a549-hace2-tpsa |
| pUNO1-hACE2 | Expression vector | puno1-hace2 |
| pUNO1-hTMPRSS2a | Expression vector | puno1-htp2a |
| Anti-CoV2RBD-c1-hIgG1 | Recombinant Antibody | cov2rbdc1-mab1 |

For a complete list of InvivoGen's COVID-19 related products visit: <https://www.invivogen.com/covid-19>

TECHNICAL SUPPORT

InvivoGen USA (Toll-Free): 888-457-5873
InvivoGen USA (International): +1 (858) 457-5873
InvivoGen Europe: +33 (0) 5-62-71-69-39
InvivoGen Hong Kong: +852 3622-3480
E-mail: info@invivogen.com





150

1 GGACCTGCAGGGCTGAATAACCTCTGAAAGAGGAACCTGGTTAGGTACCTCTGAGCGGAAAGAACAGCTGTGGAATGTGTGTCAGTTAGGGTGTG
101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCAGGCTCCCAGCAGGCAG
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTGGCCGTGAGTGGGAGAGCGCACATCGCCACAGTCCCCGA
301 GAAGTTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCCC
401 GAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTC
501 GCATCTCTCTTACGCGCCCGCCCTACCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGC
601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGC

701 CTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCTGTTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTCAA
801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAGTCCCAATGCGTAAATCTACCACCCGAACCTCAACTCCCACCCGCATATACAAATCCTTACC
101 AGAGGAGTGTACTATCCTGACAAAGTGTTCGGTCAAGTGCCTCCACTCTACTCAGGACCTTTTCTGCCTTTCTTTTCTAACGTTACATGGTTTCATG
1201 CAATCCATGTGTCTGGGACAAACCGCACAAACGCTTCGACAAACCCTGTATTGCCATTCAATGATGGGGTGTACTTTGCTCCACAGAGAAATCCAACAT
1401 CATTGAGGATGGATTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGTGATCGTTAAACAACGCCACAAACGTTGTCATCAAAGTGTGCGAATTC
1601 CCTTCGAGTACGTGAGTCAACCTTTCTGATGGACCTGGAAGGAAACAGGGAAACTTCAAGAACCTGAGAGAGTTTGTCTTTAAGAACATCGAGCGCTA
1801 TTTAAGATCTATAGTAAGCATAACGCTATCAACCTGGTAAGGGATCTTCCCAGGGCTTTTCCAGCCTGGAACCTTTGGTTGACTTGCCTATTGGTATC
2001 AATATCACCAGATTTCCAGACCCTTCTGGCATTGCATCGGTCTTATCTTACTCCAGGTATTCTCCTCCGGGTGGACTGCCGGCGCCGCTGCCTACTATG
2201 TCGGCTATCTGCAACCAAGAAGTTCCTGCTCAAGTACAACGAAAACGGCACTATTACGGATGCTGTTGATTGTGCCCTGGACCTCTGTCTGAGACTAA
2401 ATGCACCTCAAGAGCTTTACCGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGGTCCAACCCACCGAAAGCATTGTGCGGTTCCCAAATATCACC
2601 AATCTGTGTCCTTTGGCGAAGTGTTCATGCTACAAGTTTGTCTTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTCGCTGATTACTCCG
2801 TGCTGTACAATTCGCTCTTTTCTCAACCTTCAAGTGTATGGCGTTTACCTACCAAACCTAACGACCTGTGCTTACTAATGTGTATGCCGACTCTTT
3001 TGTGATACGAGGGCATGAAGTGAAGCAGATTGCACCAGGGCAGACCGGCAAAATTGCCGACTACAACATAAGCTTCCAGATGACTTTACCGGATGTGTT
3201 ATTGATGGAACCTAAACAATCTGGATTCCAAGTGGGTGGCAACTATAACTACCTGTATAGACTGTTCCAGGAAATCCAACCTGAAACCATTCCAGCGAG
3401 ATATAAGCACAGAAATCTACCAGGCTGGAAGTACGCCCTGCAACGGCGTGGAAAGGTTCAACTGCTACTTCCCATTGCAGAGTTACGGATTCCAGCCTAC
3601 AAACGGGGTGGGTTACCAACCCTATCGTGTCTGAGTCTGAGTCTTTCGAGTCTCCATGCCCCAGCCACAGTCTGTGGCCCCAAGAAAAGCACCAATCTG
3801 GTGAAGAACAATGCGTGAACCTTAACTTAAACGGACTCACAGGAACCGCGTATTGACGGAGAGTAACAAGAAGTCTGCTCCATTCCAGCAGTTCGGTC
4001 GCGATATTGCCGACACTACCGACGCTGTCCGAGATCCCAGACATTGGAGATTCTTGATATCACACCCTGTAGTTTCCGGCGGAGTGAGCGTGATTACGCC
4201 CGGAACCAATACCAGCAATCAGGTTGCCGCTCTGTATCAGGACGTGAATTGCACCGAGGTACCTGTGCCATCCACGCTGACCAACTTACACCACATGG
4401 CGAGTATATTCCACCGGCTCCAACGCTTTTCCAGACAGTGTGGATGTCTGATCGGTGCAAGACAGTAAATAATAGTCTACGAGTGTGATATCCCATCG
4601 R V Y S T G S N V F Q T R A G C L I G A E H V N N S Y E C D I P I

Agel (791)

D614 (2640)

R685A (2853)

R683A (2847)

2801 GTGCTGGAATATGCGCCTTTATCAAACCTCAAACCACTCTCCTAGGCGGGCAGCTAGTGTAGCATCCCAAAGTATCATTGCCTACACAATGAGCCTCGG
667▶ G A G I C A S Y Q T Q T N S P R A A A S V A S Q S I I A Y T M S L G
2901 TGCTGAGAATTCTGTCGCTACAGCAACAACTCCATTGCTATCCCTACTAACTTCACAATCAGTGTGACAACCTGAAATTCTGCCCGTATCTATGACCAAA
700▶ A E N S V A Y S N N S I A I P T N F T I S V T T E I L P V S M T K
3001 ACAAGCGTTGACTGCACCATGTACATCTGTGGCGATTCTACCGAATGTAGCAATCTCCTCCTGCAATACGGATCATTCTGCACCTCAGTGAATCGTGCCC
734▶ T S V D C T M Y I C G D S T E C S N L L L Q Y G S F C T Q L N R A
3101 TCACAGGTATTGAGTGTGAGCAGGACAAGAATACGCGAGGAAGTGTTCGCCAGGTGAAGCAAATCTACAAAACCTCCACCCATAAAAAGACTTTGGCGGATT
767▶ L T G I A V E Q D K N T Q E V F A Q V K Q I Y K T P P I K D F G G F
3201 CAATTTCTCACAGATCTGCCGATCCCTCAAACCCCTCAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAAGGTAACCTCGCAGATGCCGGTTTC
800▶ N F S Q I L P D P S K P S K R S F I E D L L F N K V T L A D A G F
3301 ATCAAGCAGTATGGCGATTGTCTGGGAGACATCGCCGCTCGGGACCTGATCTGTGCACAGAAGTTCATGGACTGACCGTGTGCCTCCCTTGTGACCG
834▶ I K Q Y G D C L G D I A A R D L I C A Q K F N G L T V L P P L L T
3401 ACGAGATGATAGCCCAATACACTAGCGCCTGTGGCGGCACCATCACTTCTGGGTGGACATTCGGAGCTGGCGTGCCTTTCAGATTCTTTTGTAT
867▶ D E M I A Q Y T S A L L A G T I T S G W T F G A G A A L Q I P F A M
3501 GCAGATGGCCTACCGCTTAAACGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAAACTCATCGCAACAGGTTCAACAGTGTATCGGTAAG
900▶ Q M A Y R F N G I G V T Q N V L Y E N Q K L I A N Q A C F N S A I G K
3601 ATACAGGATAGCTGTATCCACTGCCAGCGATTGGGAAAGTTGCAGGATGTAGTGAACAGAATGCCAGGCATTAACACCCCTGGTGAACAGCTCT
934▶ I Q D S L S S T A S A L G K L Q D V V N Q N A Q A L N T L V K Q L
3701 CTTCAAATTTTGGTCCATTTCTAGCGTGTGATGACATACTGAGCCGGTTGGACAAGTGGAGGCTGAAGTGCAGATTGATAGGCTGATAACTGGGCG
967▶ S S N F G A I S S V L N D I L S R L D K V E A E V Q I D R L I T G R
3801 CCTCAGTCTCTCAGACCTATGTGACCAGCAGCTCATCCGCGCTGTGAAATTCGCGCATCCGCTAACCTGGCAGCAACAAAATGTCCGAGTGTGTG
1000▶ L Q S L Q T Y V T Q Q L I R A A E I R A S A N L A A T K M S E C V
3901 CTGGTCACTAAGAGAGTGGACTTTTGGGGAGGGGATCACCTGATGTCTTTCCTCAGTCTGCACCCATGGTGTGGTCTTTCGACGCTGACTT
1034▶ L G Q S K R V D F T C G K G K G Y H L M S F P Q S A P H G V V F L H V T
4001 ATGTCCAGCTCAGGAAAAGAACTTCACTACAGKCCCAAGCCATCTGCCACGATGGGAAAAGCCCACTTTCCAGGGAAGGCGTATTTCGTCCCAATGGTAC
1067▶ Y V P A Q E K N F T T A P A I C H D G K A H F P R E G V F V S N G T
4101 TCATTGGTTCGCTACTCAGAGAAATTTCTACGAGCCCCAGATTATAACCACTGACAATACATTTGTATCCGGCAATTTGTATGGTTATCGGGATTGTG
1100▶ H W F V T Q R N F Y E P Q I I T T D N T F V S G N C D V V I G I V
4201 AATAACTGTTTACGATCCTTTGACGACGAGCTGGACTCCTTCAAGGAGGAGCTTACAAAATTTTAAAGAAATCACACATCACCTGACGTGCACCTCG
1134▶ N N T V Y D P L Q P E L D S F K E E L D K Y F K N H T S P D V D L
4301 GAGATATTTAGGAATCAATGCTCCGTGGTCAATATTCAGAAGGAGATAGACAGGCTGAATGAGGTTGCCAAGAACTCAACGAGTCTCTGATCGATCT
1167▶ G D I S G I N A S V V N I Q K E I D R L N E V A K N L N E S L I D L
4401 GCAGGAGTTGGGCAAGTACGAACAGTATATCAAATGGCCATGGTACATTTGGCTTGGGTTTATTGCTGGGCTGATAGCTATCGTCATGGTGAACAATTATG
1200▶ Q E L G K Y E Q Y I K W P W Y I W L G F I A G L I A I V M V T I M

NheI (4567)

4501 TTGTGTTGCATGACATCTGCTGTAGTTGTCTGAAGGGCTGCTGCTCATGCGGCAGCTGTTGCTAAAAGCTAGCTGGCCAGACATGATAAGATACATTGAT
1234▶ L C C M T S C C S C L K G C C S C G S C C •
4601 GAGTTTGGACAAACCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA
4701 AACAAGTTAACAACAACAATTGCATTATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAT
4801 GGAATTCTAAAATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTG
4901 TTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAACCTGCTTCTATTCTTTAT
5001 GTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCA
5101 GAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGA
5201 GCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGG
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I L V F C D P
5301 AGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATG
111▶ A Y D S I L E R C M G C P S V V R I S R D V E D S Y P H R V A V I
5401 GTGTCAAAGCTCTTCTGCCGTTGTCTCACAGCAGACCAATGGCAATGGCTTACGACAGCAGTACCTGCAATGTAGGCTCAATGTGGACAGCAG
77▶ T D F D K Q G N S V A S G I A I A E A C V T V R G I Y A E I H V A S
5501 AGATGATCTCCCACTTGGTCTGATGGCCGCCGACATGGTGTGTTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACCAGCTC
44▶ I I E G T K T R I A A G V H H K N D E Y L M T I K E T A V E V L E
5601 CAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAA
11▶ L D Q Q S I N F T K M
5701 ACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAAT
5801 GGGGCGGAGTTGTTACGACATTTTGAAGTCCCCTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGTGGAGACTTGGAAATCCCCGTG
5901 AGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTC
6001 CCATAAGTTCATGTAAGTGGGCATAATGCCAGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTC

6101 AAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGC
6201 GGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC
6301 AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
6401 CGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCT
6501 CCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
6601 GTTCAGCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
6701 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT
6801 GAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGAAGCAGCAGATTACG
6901 CGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTA
7001 GTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTC
7101 CATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA