

Figure 1

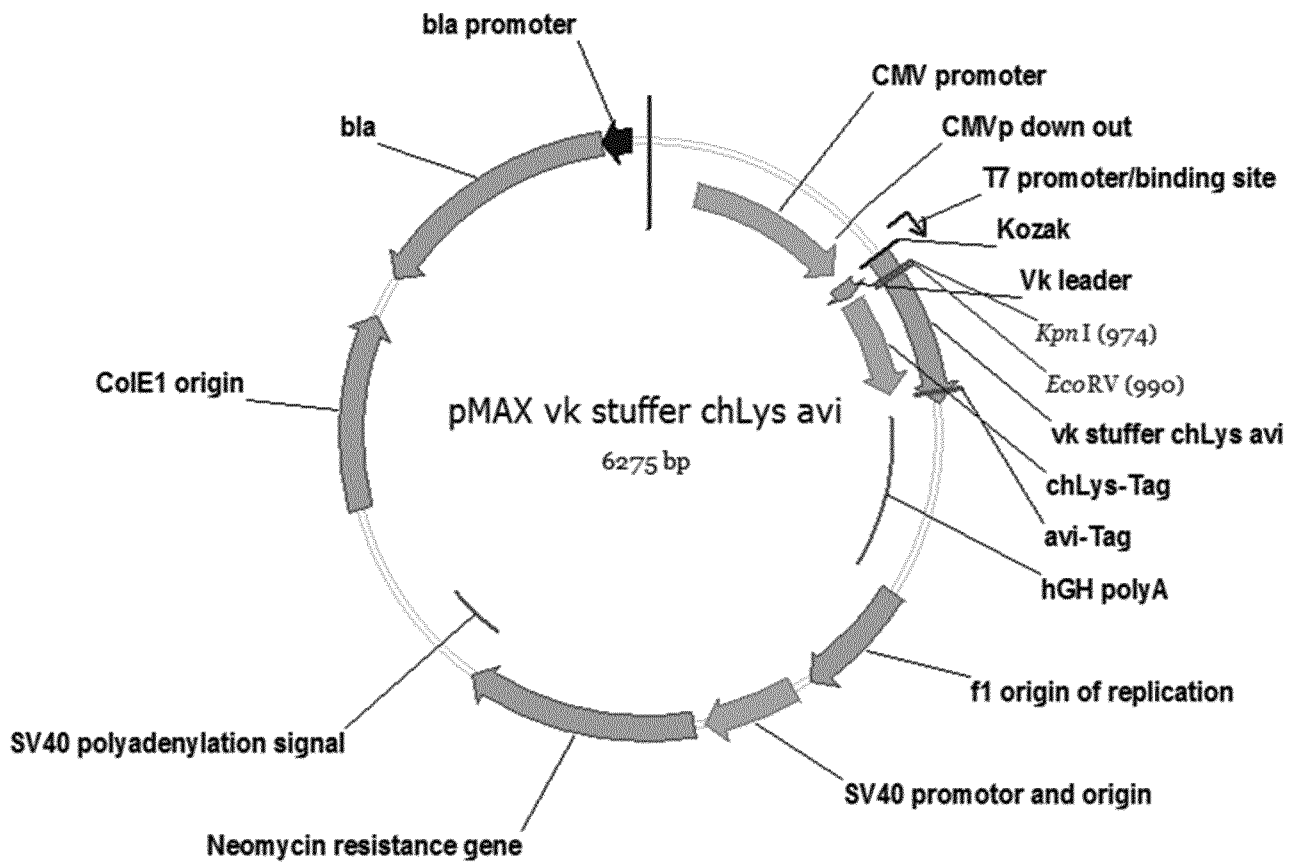


Figure 2A

|      |                   |                    |                    |                   |                   |
|------|-------------------|--------------------|--------------------|-------------------|-------------------|
| 1    | GATCTCCCGA        | TCCCCTATGG         | TGCACTCTCA         | GTACAATCTG        | CTCTGATGCC        |
| 51   | GCATAGTTAA        | GCCAGTATCT         | GCTCCCTGCT         | TGTGTGTTGG        | AGGTCGCTGA        |
| 101  | GTAGTGCGCG        | AGCAAAATTT         | AAGCTACAAC         | AAGGCAAGGC        | TTGACCGACA        |
| 151  | TTTGCA TGAA       | GAATCTGCTT         | AGGGTTAGGC         | GTTTTGCGCT        | GCTTCGCGAT        |
| 201  | GTACGGGCCA        | GATATACGCG         | TTGACATTGA         | TTATTGACTA        | GTTATTAATA        |
| 251  | GTAATCAATT        | ACGGGGTCAT         | TAGTTCATAG         | CCCATATATG        | GAGTTCGCG         |
| 301  | TTACATAACT        | TACGGTAAAT         | GGCCCGCCTG         | GCTGACCGCC        | CAACGACCCC        |
| 351  | CGCCCAT TGA       | CGTCAATAAT         | GACGTATGTT         | CCCATAGTAA        | CGCCAATAGG        |
| 401  | GACTTTCCAT        | TGACGTCAAT         | GGGTGGAGTA         | TTTACGGTAA        | ACTGCCC ACT       |
| 451  | TGGCAGTACA        | TCAAGTGTAT         | CATATGCCAA         | GTACGCCCCC        | TATTGACGTC        |
| 501  | AATGACGGTA        | AATGGCCCGC         | CTGGCATTAT         | GCCCAGTACA        | TGACCTTATG        |
| 551  | GGACTTTTCT        | ACTTGGCAGT         | ACATCTACGT         | ATTAGTCATC        | GCTATTACCA        |
| 601  | TGGTGATGCG        | GTTTTGGCAG         | TACATCAATG         | GGCGTGGATA        | GCGGTTTGAC        |
| 651  | TCACGGGGAT        | TTCCAAGTCT         | CCACCCCAT          | GACGTCAATG        | GGAGTTTGT         |
| 701  | TTGGCACCAA        | AATCAACGGG         | ACTTTCCAAA         | ATGTCGTAAC        | AACTCCGCCC        |
| 751  | CATTGACGCA        | AATGGGCGGT         | AGGCGTGTAC         | GGTGGGAGGT        | CTATATAAGC        |
| 801  | AGAGCTCTCT        | GGCTAACTAG         | AGAACCCACT         | GCTTACTGGC        | TTATCGAAAT        |
| 851  | TAATACGACT        | CACTATAGGG         | AGACCCAAGC         | TGGCTAGCAA        | GCTTCTAGCG        |
| 901  | CCACCATGGT        | GTTGCAGACC         | CAGGTCTTCA         | TTTCTCTGTT        | GCTCTGGATC        |
| 951  | TCTGGTGCCT        | ACGGGGATGG         | TACCGGGAAA         | ATGATGGATA        | TCATCGAGGG        |
| 1001 | CCGGATGGAC        | <u>AAGGTGTTTCG</u> | <u>GCAGATGCGA</u>  | <u>GCTGGCCGCT</u> | <u>GCCATGAAGC</u> |
| 1051 | <u>GGCACGGCCT</u> | <u>GGACAAC TAC</u> | <u>CGGGGCTACA</u>  | <u>GCCTGGGCAA</u> | <u>CTGGGTCTGC</u> |
| 1101 | <u>GCCGCCAAGT</u> | <u>TCGAGAGCAA</u>  | <u>CTTCAATACT</u>  | <u>CAGGCCACCA</u> | <u>ACCGGAACAC</u> |
| 1151 | <u>CGACGGCAGC</u> | <u>ACCGACTACG</u>  | <u>GCATCCTGCA</u>  | <u>GATCAACAGC</u> | <u>CGGTGGTGGT</u> |
| 1201 | <u>GCAACGACGG</u> | <u>CAGGACCCCC</u>  | <u>GGCAGCCGGA</u>  | <u>ACCTGTGCAA</u> | <u>CATCCCTTGC</u> |
| 1251 | <u>AGCGCCCTGC</u> | <u>TGTCCAGCGA</u>  | <u>CATCACC GCC</u> | <u>AGCGTGAACT</u> | <u>GCGCCAAGAA</u> |
| 1301 | <u>AATCGTGTCC</u> | <u>GACGGCAACG</u>  | <u>GCATGAACGC</u>  | <u>CTGGGTGGCC</u> | <u>TGGCGGAACC</u> |
| 1351 | <u>GGTGCAAGGG</u> | <u>CACAGACGTG</u>  | <u>CAGGCCTGGA</u>  | <u>TCAGAGGCTG</u> | <u>CAGACTGGTT</u> |
| 1401 | AACTCTAGAG        | GTCTGAACGA         | CATCTTCGAG         | GCTCAGAAAA        | TCGAATGGCA        |
| 1451 | CGAATAATGA        | GAATTCTCTA         | GATAATGAGT         | TTAAACGGGT        | GGCATCCCTG        |
| 1501 | TGACCCCTCC        | CCAGTGCCCTC        | TCCTGGCCCT         | GGAAGTTGCC        | ACTCCAGTGC        |
| 1551 | CCACCAGCCT        | TGTCCTAATA         | AAATTAAGTT         | GCATCATTTT        | GTCTGACTAG        |
| 1601 | GTGTCCTTCT        | ATAATATTAT         | GGGGTGGAGG         | GGGGTGGTAT        | GGAGCAAGGG        |
| 1651 | GCAAGTTGGG        | AAGACAACCT         | GTAGGGCCTG         | CGGGGTCTAT        | TGGGAACCAA        |
| 1701 | GCTGGAGTGC        | AGTGGCACAA         | TCTTGGCTCA         | CTGCAATCTC        | CGCCTCCTGG        |
| 1751 | GTTCAAGCGA        | TTCTCCTGCC         | TCAGCCTCCC         | GAGTTGTTGG        | GATTCCAGGC        |
| 1801 | ATGCATGACC        | AGGCTCACCT         | AATTTTTGTT         | TTTTTGGTAG        | AGACGGGGTT        |
| 1851 | TCACCATATT        | GGCCAGGCTG         | GTCTCCA ACT        | CCTAATCTCA        | GGTGATCTAC        |
| 1901 | CCACCTTGGC        | CTCCCAAATT         | GCTGGGATTA         | CAGGCGTGAA        | CCACTGCTCC        |
| 1951 | CTTCCCTGTC        | CTTCTGATTT         | TAAAATAACT         | ATACCAGCAG        | GAGGACGTCC        |
| 2001 | AGACACAGCA        | TAGGCTACCT         | GGCCATGCCC         | AACCGGTGGG        | ACATTTGAGT        |
| 2051 | TGCTTGCTTG        | GCACTGTCCT         | CTCATGCGTT         | GGGTCCACTC        | AGTAGATGCC        |
| 2101 | TGTTGAATTG        | GGTACGCGGC         | ATCGATTCCA         | CGCGCCCTGT        | AGCGGCGCAT        |
| 2151 | TAAGCGCGGC        | GGGTGTGGTG         | GTTACGCGCA         | GCGTGACCGC        | TACACTTGCC        |
| 2201 | AGCGCCCTAG        | CGCCCGCTCC         | TTTCGCTTTC         | TTCCCTTCCT        | TTCTCGCCAC        |
| 2251 | GTTCGCCGGC        | TTTCCCCGTC         | AAGCTCTAAA         | TCGGGGGCTC        | CCTTTAGGGT        |
| 2301 | TCCGATTTAG        | TGCTTTACGG         | CACCTCGACC         | CCAAAAA ACT       | TGATTAGGGT        |
| 2351 | GATGGTTTAC        | GTAGTGGGCC         | ATCGCCCTGA         | TAGACGGTTT        | TTGCCCCTTT        |
| 2401 | GACGTTGGAG        | TCCACGTTCT         | TTAATAGTGG         | ACTCTTGTTT        | CAAACTGGAA        |
| 2451 | CAACACTCAA        | CCCTATCTCG         | GTCTATTCTT         | TTGATTTATA        | AGGGATTTTG        |
| 2501 | CCGATTTTCG        | CCTATTGGTT         | AAAAAATGAG         | CTGATTTAAC        | AAAAATTTAA        |

Figure 2B

|      |              |             |             |            |            |
|------|--------------|-------------|-------------|------------|------------|
| 2551 | CGCGAATTAA   | TTCTGTGGAA  | TGTGTGTCAG  | TTAGGGTGTG | GAAAGTCCCC |
| 2601 | AGGCTCCCCA   | GCAGGCAGAA  | GTATGCAAAG  | CATGCATCTC | AATTAGTCAG |
| 2651 | CAACCAGGTG   | TGGAAAGTCC  | CCAGGCTCCC  | CAGCAGGCAG | AAGTATGCAA |
| 2701 | AGCATGCATC   | TCAATTAGTC  | AGCAACCATA  | GTCCCGCCCC | TAACTCCGCC |
| 2751 | CATCCCGCCC   | CTAACTCCGC  | CCAGTTCCGC  | CCATTCTCCG | CCCCATGGCT |
| 2801 | GACTAATTTT   | TTTTATTTAT  | GCAGAGGCCG  | AGGCCGCCTC | TGCCTCTGAG |
| 2851 | CTATTCCAGA   | AGTAGTGAGG  | AGGCTTTTTT  | GGAGGCCTAG | GCTTTTGCAA |
| 2901 | AAAGCTCCCG   | GGAGCTTGTA  | TATCCATTTT  | CGGATCTGAT | CAAGAGACAG |
| 2951 | GATGAGGATC   | GTTTCGCATG  | ATTGAACAAG  | ATGGATTGCA | CGCAGGTTCT |
| 3001 | CCGGCCGCTT   | GGGTGGAGAG  | GCTATTCGGC  | TATGACTGGG | CACAACAGAC |
| 3051 | AATCGGCTGC   | TCTGATGCCG  | CCGTGTTCCG  | GCTGTCAGCG | CAGGGGCGCC |
| 3101 | CGGTTCTTTT   | TGTCAAGACC  | GACCTGTCCG  | GTGCCCTGAA | TGAACTGCAG |
| 3151 | GACGAGGCAG   | CGCGGCTATC  | GTGGCTGGCC  | ACGACGGGCG | TTCCTTGCGC |
| 3201 | AGCTGTGCTC   | GACGTTGTCA  | CTGAAGCGGG  | AAGGGACTGG | CTGCTATTGG |
| 3251 | GCGAAGTGCC   | GGGGCAGGAT  | CTCCTGTCAT  | CTCACCTTGC | TCCTGCCGAG |
| 3301 | AAAGTATCCA   | TCATGGCTGA  | TGCAATGCGG  | CGGCTGCATA | CGCTTGATCC |
| 3351 | GGCTACCTGC   | CCATTTCGACC | ACCAAGCGAA  | ACATCGCATC | GAGCGAGCAC |
| 3401 | GTA CT CGGAT | GGAAGCCGGT  | CTTGTCGATC  | AGGATGATCT | GGACGAAGAG |
| 3451 | CATCAGGGGC   | TCGCGCCAGC  | CGAACTGTTC  | GCCAGGCTCA | AGGCGCGCAT |
| 3501 | GCCCCACGGC   | GAGGATCTCG  | TCGTGACCCA  | TGGCGATGCC | TGCTTGCCGA |
| 3551 | ATATCATGGT   | GGAAAATGGC  | CGTTTTTCTG  | GATTCATCGA | CTGTGGCCGG |
| 3601 | CTGGGTGTGG   | CGGACCGCTA  | TCAGGACATA  | GCGTTGGCTA | CCCGTGATAT |
| 3651 | TGCTGAAGAG   | CTTGCGGGCG  | AATGGGCTGA  | CCGCTTCCTC | GTGCTTTACG |
| 3701 | GTATCGCCGC   | TCCCGATTTCG | CAGCGCATCG  | CCTTCTATCG | CCTTCTTGAC |
| 3751 | GAGTTCTTCT   | GAGCGGGACT  | CTGGGGTTTCG | AAATGACCGA | CCAAGCGACG |
| 3801 | CCCAACCTGC   | CATCACGAGA  | TTTCGATTCC  | ACCGCCGCCT | TCTATGAAAG |
| 3851 | GTTGGGCTTC   | GGAATCGTTT  | TCCGGGACGC  | CGGCTGGATG | ATCCTCCAGC |
| 3901 | GCGGGGATCT   | CATGCTGGAG  | TTCTTCGCCC  | ACCCCAACTT | GTTTATTGCA |
| 3951 | GCTTATAATG   | GTTACAAATA  | AAGCAATAGC  | ATCACAAATT | TCACAAATAA |
| 4001 | AGCATTTTTT   | TCACTGCATT  | CTAGTTGTGG  | TTTGTCCAAA | CTCATCAATG |
| 4051 | TATCTTATCA   | TGTCTGTATA  | CCGTCGACCT  | CTAGCTAGAG | CTTGGCGTAA |
| 4101 | TCATGGTCAT   | AGCTGTTTCC  | TGTGTGAAAT  | TGTTATCCGC | TCACAATTCC |
| 4151 | ACACAACATA   | CGAGCCGGAA  | GCATAAAGTG  | TAAAGCCTGG | GGTGCCTAAT |
| 4201 | GAGTGAGCTA   | ACTCACATTA  | ATTGCGTTGC  | GCTCACTGCC | CGCTTTCAG  |
| 4251 | TCGGGAAACC   | TGTCGTGCCA  | GCTGCATTAA  | TGAATCGGCC | AACGCGCGGG |
| 4301 | GAGAGGCGGT   | TTGCGTATTG  | GGCGCTCTTC  | CGCTTCCTCG | CTCACTGACT |
| 4351 | CGCTGCGCTC   | GGTCGTTCCG  | CTGCGGCGAG  | CGGTATCAGC | TCACTCAAAG |
| 4401 | GCGGTAATAC   | GGTTATCCAC  | AGAATCAGGG  | GATAACGCAG | GAAAGAACAT |
| 4451 | GTGAGCAAAA   | GGCCAGCAAA  | AGGCCAGGAA  | CCGTAAAAAG | GCCGCGTTGC |
| 4501 | TGGCGTTTTT   | CCATAGGCTC  | CGCCCCCTG   | ACGAGCATCA | CAAAAATCGA |
| 4551 | CGCTCAAGTC   | AGAGGTGGCG  | AAACCCGACA  | GGACTATAAA | GATACCAGGC |
| 4601 | GTTTCCCCCT   | GGAAGCTCCC  | TCGTGCGCTC  | TCCTGTTCCG | ACCCTGCCGC |
| 4651 | TTACCGGATA   | CCTGTCCGCC  | TTTCTCCCTT  | CGGGAAGCGT | GGCGCTTTCT |
| 4701 | CATAGCTCAC   | GCTGTAGGTA  | TCTCAGTTTCG | GTGTAGGTCG | TTGCTCCAA  |
| 4751 | GCTGGGCTGT   | GTGCACGAAC  | CCCCCGTTCA  | GCCCGACCGC | TGCGCCTTAT |
| 4801 | CCGGTAACTA   | TCGTCTTGAG  | TCCAACCCGG  | TAAGACACGA | CTTATCGCCA |
| 4851 | CTGGCAGCAG   | CCACTGGTAA  | CAGGATTAGC  | AGAGCGAGGT | ATGTAGGCGG |
| 4901 | TGCTACAGAG   | TTCTTGAAGT  | GGTGGCCTAA  | CTACGGCTAC | ACTAGAAGAA |
| 4951 | CAGTATTTGG   | TATCTGCGCT  | CTGCTGAAGC  | CAGTTACCTT | CGGAAAAAGA |
| 5001 | GTTGGTAGCT   | CTTGATCCGG  | CAAACAAACC  | ACCGCTGGTA | GCGGTGGTTT |
| 5051 | TTTTGTTTGC   | AAGCAGCAGA  | TTACGCGCAG  | AAAAAAAGGA | TCTCAAGAAG |

## Figure 2C

```
5101 ATCCTTTGAT CTTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA
5151 CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT
5201 CCTTTTAAAT TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT
5251 AAAC TTGGTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA
5301 GCGATCTGTC TATTTTCGTT ATCCATAGTT GCCTGACTCC CCGTCGTGTA
5351 GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT GCTGCAATGA
5401 TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCCAG
5451 CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC
5501 CATCCAGTCT ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG
5551 TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA
5601 CGCTCGTCGT TTGGTATGGC TTCATT CAGC TCCGGTTCCC AACGATCAAG
5651 GCGAGTTACA TGATCCCCCA TGTGTGCAA AAAAGCGGTT AGCTCCTTCG
5701 GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG
5751 GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG
5801 CTTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA
5851 TCGGCGGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA TAATACCGCG
5901 CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG
5951 GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC
6001 CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTTACTTT CACCAGCGTT
6051 TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAA AGGGAATAAG
6101 GGCGACACGG AAATGTTGAA TACTCATACT CTTCTTTTT CAATATTATT
6151 GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT
6201 ATTTAGAAAA ATAAACAAAT AGGGGTTCCG CGCACATTTT CCCGAAAAGT
6251 GCCACCTGAC GTCGACGGAT CGGGA
```