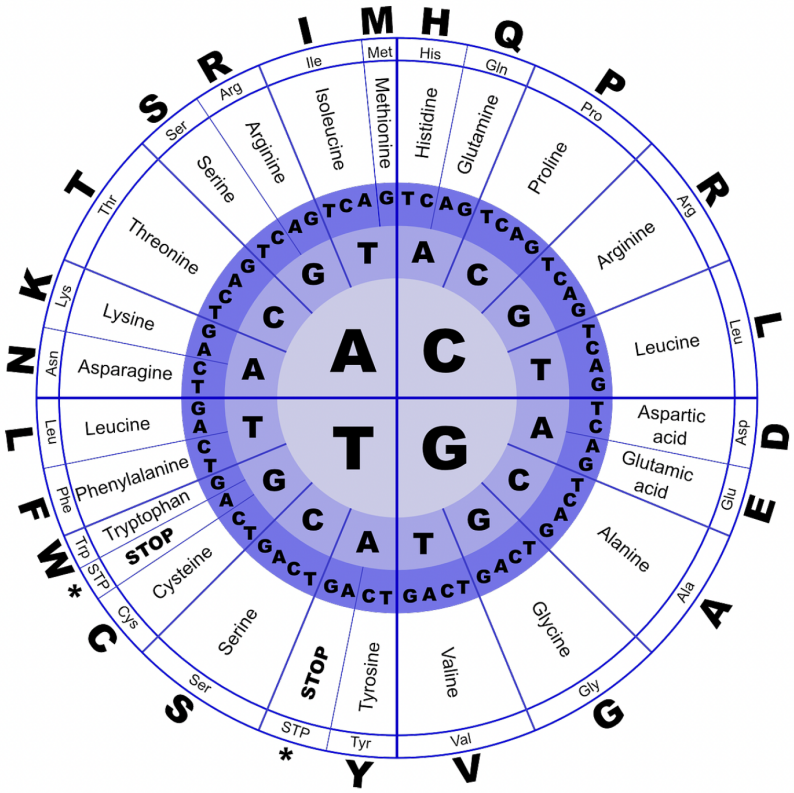
Valeurs de pK des 20 acides aminés courants

	pK_A	pK_B	pK_R
Gly	2,3	9,6	
Ala	2,3	9.7	
Val	2,3	9,6	
Leu	2,4	9,6	
Ile	2,4	9,7	
Pro	2	9,6	
Phe	1,8	9,1	
Trp	2,4	9,4	
Asn	2	8,8	
Gln	2,2	9,1	
Tyr	2,2	9,1	10,1
Ser	2,2	9,2	
Thr	2,6	10,4	
Cys	1,7	10,8	8,3
Met	2,3	9,2	
Lys	2,2	9,2	10,5
Arg	2,2	9,2	12,5
His	1,8	9,2	6,0
Asp	2,1	9,8	3,9
Glu	2,2	9,7	4,3

Code génétique



IMPORTANT : L'**<u>acide aminé</u>** correspondant à chacun des codons est indiqué **<u>sous la 1ère base</u>** de celui-ci. Pour faciliter le repérage, les numérotations du 1er et du dernier nucléotide ou acide aminé sont indiquées pour chacune des lignes. Dans la séquence 2, les exons sont en majuscules et les bordures introniques en minuscules. (…) indique que certaines bases ne sont pas indiquées.

Séquence 1

```
   1 AGTGCGCGGTGGCGGCGGCGTCGCGGGCAGCTGGCGCCGCGCGGTCCTGCTCTGCCGGTC 60
  61 GCACGGACGCACCGGCGGGCCGCCGGCCGGAGGGACGGGGCGGGAGCTGGGCCCGCGGAC 120
 121 AGCGAGCCGGAGCGGGAGCCGCGCGTAGCGAGCCGGGCTCCGGCGCTCGCCAGTCTCCCG 180
 181 AGCGGCGCCCGCCTCCCGCCGGTGCCCGCGCCGGGCCGTGGGGGGCAGCATGCCCGCGCG 240
 241 CGCTGCCTGAGGACGCCGCGGCCCCCGCCCCCGCCATGGGCGCCCCTGCCTGCGCCCTCG 300
   1                                   M  G  A  P  A  C  A  L  A 9
 301 CGCTCTGCGTGGCCGTGGCCATCGTGGCCGGCGCCTCCTCGGAGTCCTTGGGGACGGAGC 360
  10   L  C  V  A  V  A  I  V  A  G  A  S  S  E  S  L  G  T  E  Q 29
 361 AGCGCGTCGTGGGGCGAGCGGCAGAAGTCCCGGGCCCAGAGCCCGGCCAGCAGGAGCAGT 420
  30   R  V  V  G  R  A  A  E  V  P  G  P  E  P  G  Q  Q  E  Q  L 49
 421 TGGTCTTCGGCAGCGGGGATGCTGTGGAGCTGAGCTGTCCCCCGCCCGGGGGTGGTCCCA 480
  50   V  F  G  S  G  D  A  V  E  L  S  C  P  P  P  G  G  G  P  M 69
 481 TGGGGCCCACTGTCTGGGTCAAGGATGGCACAGGGCTGGTGCCCTCGGAGCGTGTCCTGG 540
  70   G  P  T  V  W  V  K  D  G  T  G  L  V  P  S  E  R  V  L  V 89
 541 TGGGGCCCCAGCGGCTGCAGGTGCTGAATGCCTCCCACGAGGACTCCGGGGCCTACAGCT 600
  90   G  P  Q  R  L  Q  V  L  N  A  S  H  E  D  S  G  A  Y  S  C 109
 601 GCCGGCAGCGGCTCACGCAGCGCGTACTGTGCCACTTCAGTGTGCGGGTGACAGACGCTC 660
 110   R  Q  R  L  T  Q  R  V  L  C  H  F  S  V  R  V  T  D  A  P 129
 661 CATCCTCGGGAGATGACGAAGACGGGGAGGACGAGGCTGAGGACACAGGTGTGGACACAG 720
 130   S  S  G  D  D  E  D  G  E  D  E  A  E  D  T  G  V  D  T  G 149
 721 GGGCCCCTTACTGGACACGGCCCGAGCGGATGGACAAGAAGCTGCTGGCCGTGCCGGCCG 780
 150   A  P  Y  W  T  R  P  E  R  M  D  K  K  L  L  A  V  P  A  A 169
 781 CCAACACCGTCCGCTTCCGCTGCCCAGCCGCTGGCAACCCCACTCCCTCCATCTCCTGGC 840
 170   N  T  V  R  F  R  C  P  A  A  G  N  P  T  P  S  I  S  W  L 189
 841 TGAAGAACGGCAGGGAGTTCCGCGGCGAGCACCGCATTGGAGGCATCAAGCTGCGGCATC 900
 190   K  N  G  R  E  F  R  G  E  H  R  I  G  G  I  K  L  R  H  Q 209
 901 AGCAGTGGAGCCTGGTCATGGAAAGCGTGGTGCCCTCGGACCGCGGCAACTACACCTGCG 960
 210   Q  W  S  L  V  M  E  S  V  V  P  S  D  R  G  N  Y  T  C  V 229
 961 TCGTGGAGAACAAGTTTGGCAGCATCCGGCAGACGTACACGCTGGACGTGCTGGAGCGCT 1020
 230   V  E  N  K  F  G  S  I  R  Q  T  Y  T  L  D  V  L  E  R  S 249
1021 CCCCGCACCGGCCCATCCTGCAGGCGGGGCTGCCGGCCAACCAGACGGCGGTGCTGGGCA 1080
 250   P  H  R  P  I  L  Q  A  G  L  P  A  N  Q  T  A  V  L  G  S 269
1081 GCGACGTGGAGTTCCACTGCAAGGTGTACAGTGACGCACAGCCCCACATCCAGTGGCTCA 1140
 270   D  V  E  F  H  C  K  V  Y  S  D  A  Q  P  H  I  Q  W  L  K 289
1141 AGCACGTGGAGGTGAATGGCAGCAAGGTGGGCCCGGACGGCACACCCTACGTTACCGTGC 1200
 290   H  V  E  V  N  G  S  K  V  G  P  D  G  T  P  Y  V  T  V  L 309
1201 TCAAGACGGCGGGCGCTAACACCACCGACAAGGAGCTAGAGGTTCTCTCCTTGCACAACG 1260
 310   K  T  A  G  A  N  T  T  D  K  E  L  E  V  L  S  L  H  N  V 329
1261 TCACCTTTGAGGACGCCGGGGAGTACACCTGCCTGGCGGGCAATTCTATTGGGTTTTCTC 1320
 330   T  F  E  D  A  G  E  Y  T  C  L  A  G  N  S  I  G  F  S  H 349
1321 ATCACTCTGCGTGGCTGGTGGTGCTGCCAGCCGAGGAGGAGCTGGTGGAGGCTGACGAGG 1380
 350   H  S  A  W  L  V  V  L  P  A  E  E  E  L  V  E  A  D  E  A 369
1381 CGGGCAGTGTGTATGCAGGCATCCTCAGCTACGGGGTGGGCTTCTTCCTGTTCATCCTGG 1440
 370   G  S  V  Y  A  G  I  L  S  Y  G  V  G  F  F  L  F  I  L  V 389
1441 TGGTGGCGGCTGTGACGCTCTGCCGCCTGCGCAGCCCCCCCAAGAAAGGCCTGGGCTCCC 1500
 390   V  A  A  V  T  L  C  R  L  R  S  P  P  K  K  G  L  G  S  P 409
1501 CCACCGTGCACAAGATCTCCCGCTTCCCGCTCAAGCGACAGGTGTCCCTGGAGTCCAACG 1560
 410   T  V  H  K  I  S  R  F  P  L  K  R  Q  V  S  L  E  S  N  A 429
1561 CGTCCATGAGCTCCAACACACCACTGGTGCGCATCGCAAGGCTGTCCTCAGGGGAGGGCC 1620
 430   S  M  S  S  N  T  P  L  V  R  I  A  R  L  S  S  G  E  G  P 449
1621 CCACGCTGGCCAATGTCTCCGAGCTCGAGCTGCCTGCCGACCCCAAATGGGAGCTGTCTC 1680
 450   T  L  A  N  V  S  E  L  E  L  P  A  D  P  K  W  E  L  S  R 469
1681 GGGCCCGGCTGACCCTGGGCAAGCCCCTTGGGGAGGGCTGCTTCGGCCAGGTGGTCATGG 1740
 470   A  R  L  T  L  G  K  P  L  G  E  G  C  F  G  Q  V  V  M  A 489
```

```
1741 CGGAGGCCATCGGCATTGACAAGGACCGGGCCGCCAAGCCTGTCACCGTAGCCGTGAAGA 1800
 490   E  A  I  G  I  D  K  D  R  A  A  K  P  V  T  V  A  V  K  M 509
1801 TGCTGAAAGACGATGCCACTGACAAGGACCTGTCGGACCTGGTGTCTGAGATGGAGATGA 1860
 510   L  K  D  D  A  T  D  K  D  L  S  D  L  V  S  E  M  E  M  M 529
1861 TGAAGATGATCGGGAAACACAAAAACATCATCAACCTGCTGGGCGCCTGCACGCAGGGCG 1920
 530   K  M  I  G  K  H  K  N  I  I  N  L  L  G  A  C  T  Q  G  G 549
1921 GGCCCCTGTACGTGCTGGTGGAGTACGCGGCCAAGGGTAACCTGCGGGAGTTTCTGCGGG 1980
 550   P  L  Y  V  L  V  E  Y  A  A  K  G  N  L  R  E  F  L  R  A 569
1981 CGCGGCGGCCCCCGGGCCTGGACTACTCCTTCGACACCTGCAAGCCGCCCGAGGAGCAGC 2040
 570   R  R  P  P  G  L  D  Y  S  F  D  T  C  K  P  P  E  E  Q  L 589
2041 TCACCTTCAAGGACCTGGTGTCCTGTGCCTACCAGGTGGCCCGGGGCATGGAGTACTTGG 2100
 590   T  F  K  D  L  V  S  C  A  Y  Q  V  A  R  G  M  E  Y  L  A 609
2101 CCTCCCAGAAGTGCATCCACAGGGACCTGGCTGCCCGCAATGTGCTGGTGACCGAGGACA 2160
 610   S  Q  K  C  I  H  R  D  L  A  A  R  N  V  L  V  T  E  D  N 629
2161 ACGTGATGAAGATCGCAGACTTCGGGCTGGCCCGGGACGTGCACAACCTCGACTACTACA 2220
 630   V  M  K  I  A  D  F  G  L  A  R  D  V  H  N  L  D  Y  Y  K 649
2221 AGAAGACGACCAACGGCCGGCTGCCCGTGAAGTGGATGGCGCCTGAGGCCTTGTTTGACC 2280
 650   K  T  T  N  G  R  L  P  V  K  W  M  A  P  E  A  L  F  D  R 669
2281 GAGTCTACACTCACCAGAGTGACGTCTGGTCCTTTGGGGTCCTGCTCTGGGAGATCTTCA 2340
 670   V  Y  T  H  Q  S  D  V  W  S  F  G  V  L  L  W  E  I  F  T 689
2341 CGCTGGGGGGCTCCCCGTACCCCGGCATCCCTGTGGAGGAGCTCTTCAAGCTGCTGAAGG 2400
 690   L  G  G  S  P  Y  P  G  I  P  V  E  E  L  F  K  L  L  K  E 709
2401 AGGGCCACCGCATGGACAAGCCCGCCAACTGCACACACGACCTGTACATGATCATGCGGG 2460
 710   G  H  R  M  D  K  P  A  N  C  T  H  D  L  Y  M  I  M  R  E 729
2461 AGTGCTGGCATGCCGCGCCCTCCCAGAGGCCCACCTTCAAGCAGCTGGTGGAGGACCTGG 2520
 730   C  W  H  A  A  P  S  Q  R  P  T  F  K  Q  L  V  E  D  L  D 749
2521 ACCGTGTCCTTACCGTGACGTCCACCGACGAGTACCTGGACCTGTCGGCGCCTTTCGAGC 2580
 750   R  V  L  T  V  T  S  T  D  E  Y  L  D  L  S  A  P  F  E  Q 769
2581 AGTACTCCCCGGGTGGCCAGGACACCCCCAGCTCCAGCTCCTCAGGGGACGACTCCGTGT 2640
 770   Y  S  P  G  G  Q  D  T  P  S  S  S  S  S  G  D  D  S  V  F 789
2641 TTGCCCACGACCTGCTGCCCCCGGCCCCACCCAGCAGTGGGGGCTCGCGGACGTGAAGGG 2700
 790   A  H  D  L  L  P  P  A  P  P  S  S  G  G  S  R  T  *       806
2701 CCACTGGTCCCCAAC(…)GACACCTGGTTGCTAA
```

Séquence 2

```
     AGTGCGCGGTGGCGGCGGCGTCGCGGGCAGCTGGCGCCGCGCGGTCCTGCTCTGCCGGTC
     GCACGGACGCACCGGCGGGCCGCCGGCCGGAGGGACGGGGCGGGAGCTGGGCCCGCGGAC
     AGCGAGCCGGAGCGGGAGCCGCGCGTAGCGAGCCGGGCTCCGGCGCTCGCCAGgtccgtg
     cttggggccgggacgtacgtacgtacgtacgtacgtacgtacgtacgtacgta(…)
     acgtacgagtacgtacgacgtacgtacacgtacctgccttcctcctcctgtagTCTCCCG
     AGCGGCGCCCGCCTCCCGCCGGTGCCCGCGCCGGGCCGTGGGGGGCAGCATGCCCGCGCG
     CGCTGCCTGAGGACGCCGCGGCCCCCGCCCCCGCCATGGGCGCCCCTGCCTGCGCCCTCG
   1                                  M  G  A  P  A  C  A  L  A 9
     CGCTCTGCGTGGCCGTGGCCATCGTGGCCGGCGCCTCCTCGGAGTCCTTGGGGACGGAGC
  10   L  C  V  A  V  A  I  V  A  G  A  S  S  E  S  L  G  T  E  Q 29
     AGCGCGTCGTGGGGCGAGCGGCAGgtaagaagggacccactagacgtacgtacgtac(…)
  30   R  V  V  G  R  A  A  E                                    37
     acgtgttgcccatcttccccacagAAGTCCCGGGCCCAGAGCCCGGCCAGCAGGAGCAGT
  38                          V  P  G  P  E  P  G  Q  Q  E  Q  L 49
     TGGTCTTCGGCAGCGGGGATGCTGTGGAGCTGAGCTGTCCCCCGCCCGGGGGTGGTCCCA
  50   V  F  G  S  G  D  A  V  E  L  S  C  P  P  P  G  G  G  P  M 69
     TGGGGCCCACTGTCTGGGTCAAGGATGGCACAGGGCTGGTGCCCTCGGAGCGTGTCCTGG
  70   G  P  T  V  W  V  K  D  G  T  G  L  V  P  S  E  R  V  L  V 89
     TGGGGCCCCAGCGGCTGCAGGTGCTGAATGCCTCCCACGAGGACTCCGGGGCCTACAGCT
  90   G  P  Q  R  L  Q  V  L  N  A  S  H  E  D  S  G  A  Y  S  C 109
     GCCGGCAGCGGCTCACGCAGCGCGTACTGTGCCACTTCAGTGTGCGGGTGACAGgtgagc
 110   R  Q  R  L  T  Q  R  V  L  C  H  F  S  V  R  V  T  D       127
     tctggggccacgcacgtacgtacgt(…)acgtaccggccatctctgccttgcagACGCTC
 128                                                       A  P 129
     CATCCTCGGGAGATGACGAAGACGGGGAGGACGAGGCTGAGGACACAGGTGTGGACACAG
 130   S  S  G  D  D  E  D  G  E  D  E  A  E  D  T  G  V  D  T  G 149
     gtaggagcagggtccagggacgtacgtacgt(…)acgtacacacgcacctcggcccgcag
     GGGCCCCTTACTGGACACGGCCCGAGCGGATGGACAAGAAGCTGCTGGCCGTGCCGGCCG
 150   A  P  Y  W  T  R  P  E  R  M  D  K  K  L  L  A  V  P  A  A 169
     CCAACACCGTCCGCTTCCGCTGCCCAGCCGCTGGCAACCCCACTCCCTCCATCTCCTGGC
 170   N  T  V  R  F  R  C  P  A  A  G  N  P  T  P  S  I  S  W  L 189
     TGAAGAACGGCAGGGAGTTCCGCGGCGAGCACCGCATTGGAGGCATCAAGgtgggcgcgg
 190   K  N  G  R  E  F  R  G  E  H  R  I  G  G  I  K            205
     cggggtggcacgtacgtacgt(…)acgtacccgcccgcgtcccggtgcagCTGCGGCATC
 206                                                   L  R  H  Q 209
     AGCAGTGGAGCCTGGTCATGGAAAGCGTGGTGCCCTCGGACCGCGGCAACTACACCTGCG
 210   Q  W  S  L  V  M  E  S  V  V  P  S  D  R  G  N  Y  T  C  V 229
     TCGTGGAGAACAAGTTTGGCAGCATCCGGCAGACGTACACGCTGGACGTGCTGGgtgagg
 230   V  E  N  K  F  G  S  I  R  Q  T  Y  T  L  D  V  L  E       247
     gccctggggcggcacgtacgtacgt(…)acgtacagcgtcatctgcccccacagAGCGCT
 248                                                       R  S 249
     CCCCGCACCGGCCCATCCTGCAGGCGGGGCTGCCGGCCAACCAGACGGCGGTGCTGGGCA
 250   P  H  R  P  I  L  Q  A  G  L  P  A  N  Q  T  A  V  L  G  S 269
     GCGACGTGGAGTTCCACTGCAAGGTGTACAGTGACGCACAGCCCCACATCCAGTGGCTCA
 270   D  V  E  F  H  C  K  V  Y  S  D  A  Q  P  H  I  Q  W  L  K 289
     AGCACGTGGAGGTGAATGGCAGCAAGGTGGGCCCGGACGGCACACCCTACGTTACCGTGC
 290   H  V  E  V  N  G  S  K  V  G  P  D  G  T  P  Y  V  T  V  L 309
     TCAAGgtgggccaccgtgtgcacgacgtacgat(…)acgtacgtagctctgctctctctt
 310   K                                                          310
     tgtagACGGCGGGCGCTAACACCACCGACAAGGAGCTAGAGGTTCTCTCCTTGCACAACG
 311      T  A  G  A  N  T  T  D  K  E  L  E  V  L  S  L  H  N  V 329
     TCACCTTTGAGGACGCCGGGGAGTACACCTGCCTGGCGGGCAATTCTATTGGGTTTTCTC
 330   T  F  E  D  A  G  E  Y  T  C  L  A  G  N  S  I  G  F  S  H 349
     ATCACTCTGCGTGGCTGGTGGTGCTGCCAGgtaccggcttctgctgctgacgtacgt(…)
 350   H  S  A  W  L  V  V  L  P  A                              359
     acgtacgtaccaacgcccatgtctttgcagCCGAGGAGGAGCTGGTGGAGGCTGACGAGG
 360                                 E  E  E  L  V  E  A  D  E  A 369
     CGGGCAGTGTGTATGCAGGCATCCTCAGCTACGGGGTGGGCTTCTTCCTGTTCATCCTGG
 370   G  S  V  Y  A  G  I  L  S  Y  G  V  G  F  F  L  F  I  L  V 389
     TGGTGGCGGCTGTGACGCTCTGCCGCCTGCGCAGCCCCCCCAAGAAAGGCCTGGGCTCCC
```

390 V A A V T L C R L R S P P K K G L G S P 409
CCACCGTGCACAAGATCTCCCGCTTCCCGCTCAAGCGACAGgtaacagaaagtagata(…)
410 T V H K I S R F P L K R Q 422
acgtacgtacgtacgtacgtagcccctgctgacccaagcagGTGTCCCTGGAGTCCAACG
423 V S L E S N A 429
CGTCCATGAGCTCCAACACACCACTGGTGCGCATCGCAAGGCTGTCCTCAGGGGAGGGCC
430 S M S S N T P L V R I A R L S S G E G P 449
CCACGCTGGCCAATGTCTCCGAGCTCGAGCTGCCTGCCGACCCCAAATGGGAGCTGTCTC
450 T L A N V S E L E L P A D P K W E L S R 469
GGGCCCGgtcagtggtgctgagggccacgtacgtacgt(…)acgtacccctctgcctcca
470 A R 471
ctgccagGCTGACCCTGGGCAAGCCCCTTGGGGAGGGCTGCTTCGGCCAGGTGGTCATGG
472 L T L G K P L G E G C F G Q V V M A 489
CGGAGGCCATCGGCATTGACAAGGACCGGGCCGCCAAGCCTGTCACCGTAGCCGTGAAGA
490 E A I G I D K D R A A K P V T V A V K M 509
TGCTGAAAGgtgaggaggggcggccagacgtacgt(…)acgtacgtacgcccccgctc
510 L K D 512
cgtgcacagACGATGCCACTGACAAGGACCTGTCGGACCTGGTGTCTGAGATGGAGATGA
513 D A T D K D L S D L V S E M E M M 529
TGAAGATGATCGGGAAACACAAAAACATCATCAACCTGCTGGGCGCCTGCACGCAGGGCG
530 K M I G K H K N I I N L L G A C T Q G G 549
gtaggtgcggtagcggcggacgtcagtcagt(…)acgtactgaggagcccgtgtccccag
GGCCCCTGTACGTGCTGGTGGAGTACGCGGCCAAGGGTAACCTGCGGGAGTTTCTGCGGG
550 P L Y V L V E Y A A K G N L R E F L R A 569
CGCGGCGGCCCCCGGGCCTGGACTACTCCTTCGACACCTGCAAGCCGCCCGAGGAGCAGC
570 R R P P G L D Y S F D T C K P P E E Q L 589
TCACCTTCAAGGACCTGGTGTCCTGTGCCTACCAGGTGGCCCGGGGCATGGAGTACTTGG
590 T F K D L V S C A Y Q V A R G M E Y L A 609
CCTCCCAGAAGgtgggcagggcggcaggtgagctacgtcagt(…)gcatacgcctcccac
610 S Q K 612
cccttccccagTGCATCCACAGGGACCTGGCTGCCCGCAATGTGCTGGTGACCGAGGACA
613 C I H R D L A A R N V L V T E D N 629
ACGTGATGAAGATCGCAGACTTCGGGCTGGCCCGGGACGTGCACAACCTCGACTACTACA
630 V M K I A D F G L A R D V H N L D Y Y K 649
AGAAGACGACCAACgtgagcccggccctggggtacgtacga(…)acgtacgtacgccttc
650 K T T N 653
ccacaccctcccagGGCCGGCTGCCCGTGAAGTGGATGGCGCCTGAGGCCTTGTTTGACC
654 G R L P V K W M A P E A L F D R 669
GAGTCTACACTCACCAGAGTGACGTgtacgtgtcctgcagagctacgttgcacgatg(…)
670 V Y T H Q S D V 677
acgtagacctcaccttcccctgcagCTGGTCCTTTGGGGTCCTGCTCTGGGAGATCTTCA
678 W S F G V L L W E I F T 689
CGCTGGGGGGCTCCCCGTACCCCGGCATCCCTGTGGAGGAGCTCTTCAAGCTGCTGAAGG
690 L G G S P Y P G I P V E E L F K L L K E 709
AGGGCCACCGCATGGACAAGCCCGCCAACTGCACACACGACCT gtgagtggcatccc(…)
710 G H R M D K P A N C T H D L 723
acgtcgtacagtacgtacagtggcctgagcgccctgcccgcagGTACATGATCATGCGGG
724 Y M I M R E 729
AGTGCTGGCATGCCGCGCCCTCCCAGAGGCCCACCTTCAAGCAGCTGGTGGAGGACCTGG
730 C W H A A P S Q R P T F K Q L V E D L D 749
ACCGTGTCCTTACCGTGACGTCCACCGACgtgagtgctggctctggccacgtcgacg(…)
750 R V L T V T S T D 758
acgttgcacaccccgcctcccgccagcagGAGTACCTGGACCTGTCGGCGCCTTTCGAGC
759 E Y L D L S A P F E Q 769
AGTACTCCCCGGGTGGCCAGGACACCCCCAGCTCCAGCTCCTCAGGGGACGACTCCGTGT
770 Y S P G G Q D T P S S S S S G D D S V F 789
TTGCCCACGACCTGCTGCCCCCGGCCCCACCCAGCAGTGGGGGCTCGCGGACGTGAAGGG
790 A H D L L P P A P P S S G G S R T * 806
CCACTGGTCCCCAAC(…)GACACCTGGTTGCTAA