Peptides were designed from Sabaeus TAT, ENV and GAG proteins, and from **Vervet GAG** protein. **The Sabaeus alignment is conserved with the 92018 sequence and is indicated in purple (N/A for the Vervet).** When the sequence was divergent from 92018, another proposed sequence was designed (in grey), and **the amino acids that are different are in green** (for more information, contact **Béatrice JACQUELIN**, Ph.D., Institut Pasteur, Unité de Biologie des Rétrovirus et Plateforme 2 - Puces à ADN, 25-28 rue du Dr. ROUX, 75724 PARIS CEDEX 15, Tél: 33-1-45-68-87-92, Fax: 33-1-45-68-89-57). No restrictions were applied in designing these peptides. The number of peptides produced is indicated for each protein in **black** for the ones conserved with Sabaeus 92018, in **green** for the supplemental ones from the "consensus" Sabaeus, and in **red** for the total. **The green peptides have been made from the grey sequence which is a consensus for all sabaeus sequences. Note that a single sequence is identical in both the Sabaeus and Vervet GAG alignments (**indicated in yellow**).** The predictions for proteasome clevage sites of the different sequences are indicated in **bold** text (threshold used: 0.5).

TAT.SABAEUS

Word length: 15
Overlap consecutive peptides by: 11
Shorten by: 3
Lengthen by: 2
Number of peptides generated: **16**

1 21 41 61

MDQEQEARPQ**VW**EE**L**QEELHRPLQACDNTC**YC**KVCCFHCI**L**CFHRKA**L**GIC**YY**VPRPRRASKK**I**QNNKVPVHN

```
MDQEQEARPQVWEEL (15) [-1.55]
    QEARPQVWEELQEEL (15) [-1.43]
        PQVWEELQEELHRPL (15) [-1.15]
            EELQEELHRPLQACD (15) [-1.21]
                EELHRPLQACDNTCY (15) [-0.96]
                    RPLQACDNTCYCKVC (15) [-0.18]
                        ACDNTCYCKVCCFHC (15) [0.51]
                            TCYCKVCCFHCILCF (15) [1.60]
                                KVCCFHCILCFHRKA (15) [0.75]
                                    FHCILCFHRKALGIC (15) [1.09]
                                        LCFHRKALGICYYVP (15) [0.65]
                                            RKALGICYYVPRPRR (15) [-0.75]
                                                GICYYVPRPRRASKK (15) [-1.02]
                                                    YVPRPRRASKKIQNN (15) [-1.77]
                                                        PRRASKKIQNNKVPV (15) [-1.37]
                                                            SKKIQNNKVPVHN(14) [-1.31]
```

ENV.SABAEUS

```
                  Word length: 15
                  Overlap consecutive peptides by: 11
                  Shorten by: 3
                  Lengthen by: 2
                  Number of peptides generated: 221 + 52 = 273

1                   21                  41                  61                  81
MKLLTVLLWLSGCWSIVWLVQYVTVFYGIPVWKNSSVQAFCKTPNTNLWASTNCIPDDEPEGTIAEVPIPNITEKFDAWKNRNPLVG
MKLLTVLLWLSGCWS (15) [1.28]
    TVLLWLSGCWSIVWL (15) [1.68]
        WLSGCWSIVWLVQYV (15) [1.18]
            CWSIVWLVQYVTVFY (15) [1.40]
                VWLVQYVTVFYGIPV (15) [1.49]
                    QYVTVFYGIPVWKNS (15) [0.13]
                        VFYGIPVWKNSSVQA (15) [0.33]
                            IPVWKNSSVQAFCKT (15) [0.03]
                                KNSSVQAFCKTPNTN (15) [-1.01]
                                    VQAFCKTPNTNLWAS (15) [-0.15]
                                        CKTPNTNLWASTNCI (15) [-0.31]
                                            NTNLWASTNCIPDDE (15) [-0.87]
                                                WASTNCIPDDEPEGT (15) [-1.03]
                                                    NCIPDDEPEGTIAEV (15) [-0.52]
                                                        DDEPEGTIAEVPIPN (15) [-0.79]
                                                            EGTIAEVPIPNITEK (15) [-0.23]
                                                                AEVPIPNITEKFDAW (15) [-0.21]
                                                                    IPNITEKFDAWKNRN (15) [-1.29]
                                                                        TEKFDAWKNRNPLVG (15) [-1.15]

    81                  101                 121                 141                 161
DAWKNRNPLVGQAESNIHLLFESTLKPCVKLTPMCIKMNCYRLEGGAATTKTPSTTTARPEVVSVGYNDSVIEKEMEKEQAMNCSFAMAGY
DAWKNRNPLVGQAES (15) [-1.20]
    NRNPLVGQAESNIHL (15) [-0.66]
        LVGQAESNIHLLFES (15) [0.37]
            AESNIHLLFESTLKP (15) [-0.07]
                IHLLFESTLKPCVKL (15) [0.77]
                    FESTLKPCVKLTPMC (15) [0.32]
                        LKPCVKLTPMCIKMN (15) [0.40]
                            VKLTPMCIKMNCYRL (15) [0.38]
                                PMCIKMNCYRLEGGA (15) [-0.01]
                                    KMNCYRLEGGAATTK (15) [-0.73]
                                        YRLEGGAATTKTPST (15) [-0.79]
                                            GGAATTKTPSTTTAR (15) [-0.69]
                                                TTKTPSTTTARPEVV (15) [-0.66]
                                                    PSTTTARPEVVSVGY (15) [-0.15]
                                                        TARPEVVSVGYNDSV (15) [-0.13]
                                                            EVVSVGYNDSVIEKE (15) [-0.23]
                                                                VGYNDSVIEKEMEKE (15) [-1.10]
                                                                    DSVIEKEMEKEQAMN (15) [-1.25]
                                                                        EKEMEKEQAMNCSFA (15) [-1.13]
```

```
                                                                            EKEQAMNCSFAMAGY (15) [-0.39]
```

```
    161                 181                 201                 221                 241
AMNCSFAMAGYRRDVKKNYSTVWDDQEVVCEEGREKSNATNTVGCYMIHCNDSVIKEACDKTYWDTFRLRYCAPAGYALLKCNDHDYRGHK
AMNCSFAMAGYRRDVKKNYSTVWDDQEVVCEEGREKSNATNTVGCYMIHCNDSVIKEACDKTYWDTFRLRYCAPAGYALLRCADTDYSGHK
AMNCSFAMAGYRRDV (15) [0.01]
    SFAMAGYRRDVKKNY (15) [-1.01]
        AGYRRDVKKNYSTVW (15) [-1.27]
            RDVKKNYSTVWDDQE (15) [-1.91]
                KNYSTVWDDQEVVCE (15) [-0.90]
                    TVWDDQEVVCEEGRE (15) [-1.06]
                        DQEVVCEEGREKSNA (15) [-1.43]
                            VCEEGREKSNATNTV (15) [-1.05]
                                GREKSNATNTVGCYM (15) [-0.85]
                                    SNATNTVGCYMIHCN (15) [-0.01]
                                        NTVGCYMIHCNDSVI (15) [0.49]
                                            CYMIHCNDSVIKEAC (15) [0.31]
                                                HCNDSVIKEACDKTY (15) [-0.82]
                                                    SVIKEACDKTYWDTF (15) [-0.46]
                                                        EACDKTYWDTFRLRY (15) [-1.16]
                                                            KTYWDTFRLRYCAPA (15) [-0.68]
                                                                DTFRLRYCAPAGYAL (15) [0.03]
                                                                    LRYCAPAGYALLKCN (15) [0.35]
                                                                    LRYCAPAGYALLRCA (15) [0.67]
                                                                        APAGYALLKCNDHDY (15) [-0.45]
                                                                        APAGYALLRCADTDY (15) [0.03]
                                                                            YALLKCNDHDYRGHK (15) [-1.35]
                                                                            YALLRCADTDYSGHK (15) [-0.63]
```

```
    241                 261                 281                 301                 321
KCNDHDYRGHKNCQNVSVAACTGLMNTTISTGIGINGSRAENRTEIWQEDGDSNDTVIIKLNKYYNLTIRCRRPGNKTVLPVTIMAGLVFH
RCADTDYSGHKACQNVTVSACTRLINTTVSTGIGINGSYVANRTEIWQENGNSNDTVIIRLNRYYNLTIRCRRPGNKTVLPVTIMAGLVFH
KCNDHDYRGHKNCQN (15) [-2.43]
RCADTDYSGHKACQN (15) [-1.35]
    HDYRGHKNCQNVSVA (15) [-1.24]
    TDYSGHKACQNVTVS (15) [-0.64]
        GHKNCQNVSVAACTG (15) [-0.19]
        GHKACQNVTVSACTR (15) [-0.28]
            CQNVSVAACTGLMNT (15) [0.64]
            CQNVTVSACTRLINT (15) [0.37]
                SVAACTGLMNTTIST (15) [0.81]
                TVSACTRLINTTVST (15) [0.53]
                    CTGLMNTTISTGIGI (15) [0.89]
                    CTRLINTTVSTGIGI (15) [0.77]
                        MNTTISTGIGINGSR (15) [-0.07]
                        INTTVSTGIGINGSY (15) [0.30]
                            ISTGIGINGSRAENR (15) [-0.51]
                            VSTGIGINGSYVANR (15) [0.19]
                                IGINGSRAENRTEIW (15) [-0.73]
                                IGINGSYVANRTEIW (15) [-0.00]
                                    GSRAENRTEIWQEDG (15) [-1.79]
                                    GSYVANRTEIWQENG (15) [-1.07]
                                        ENRTEIWQEDGDSND (15) [-2.29]
                                        ANRTEIWQENGNSND (15) [-1.93]
                                            EIWQEDGDSNDTVII (15) [-0.64]
                                            EIWQENGNSNDTVII (15) [-0.64]
                                                EDGDSNDTVIIKLNK (15) [-0.91]
                                                ENGNSNDTVIIRLNR (15) [-0.99]
                                                    SNDTVIIKLNKYYNL (15) [-0.34]
                                                    SNDTVIIRLNRYYNL (15) [-0.42]
                                                        VIIKLNKYYNLTIRC (15) [0.35]
                                                        VIIRLNRYYNLTIRC (15) [0.27]
                                                            LNKYYNLTIRCRRPG (15) [-1.01]
                                                            LNRYYNLTIRCRRPG (15) [-1.05]
                                                                YNLTIRCRRPGNKTV (15) [-0.94]
                                                                    IRCRRPGNKTVLPVT (15) [-0.45]
                                                                        RPGNKTVLPVTIMAG (15) [0.21]
                                                                            KTVLPVTIMAGLVFH (15) [1.38]
```

```
    321                 341                 361                 381                 401
PVTIMAGLVFHSQKYNTRLKQAWCWFGGNWTQAWKEVRETIVKLPNTTYKGTRNISQIFLQRQWGDPEAENFFFNCQGEFFYCKMDWFLNY
PVTIMAGLVFHSQKYNTRLKQAWCWFGGNWRGAWQEVKETIVKLPPTRYSGTNDTSKIFLQRQWGDPEAEFFFFNCQGEFFYCKMDWFLNY
PVTIMAGLVFHSQKY (15) [0.52]
    MAGLVFHSQKYNTRL (15) [-0.23]
        VFHSQKYNTRLKQAW (15) [-1.14]
            QKYNTRLKQAWCWFG (15) [-1.07]
                TRLKQAWCWFGGNWT (15) [-0.63]
                TRLKQAWCWFGGNWR (15) [-0.88]
                    QAWCWFGGNWTQAWK (15) [-0.71]
                    QAWCWFGGNWRGAWQ (15) [-0.73]
                        WFGGNWTQAWKEVRE (15) [-1.19]
                        WFGGNWRGAWQEVKE (15) [-1.17]
                            NWTQAWKEVRETIVK (15) [-0.99]
                            NWRGAWQEVKETIVK (15) [-0.97]
                                AWKEVRETIVKLPNT (15) [-0.55]
                                AWQEVKETIVKLPPT (15) [-0.35]
                                    VRETIVKLPNTTYKG (15) [-0.53]
                                    VKETIVKLPPTRYSG (15) [-0.41]
                                        IVKLPNTTYKGTRNI (15) [-0.51]
                                        IVKLPPTRYSGTNDT (15) [-0.71]
                                            PNTTYKGTRNISQIF (15) [-0.89]
                                            PPTRYSGTNDTSKIF (15) [-1.11]
                                                YKGTRNISQIFLQRQ (15) [-0.97]
                                                YSGTNDTSKIFLQRQ (15) [-1.07]
                                                    RNISQIFLQRQWGDP (15) [-0.97]
                                                    NDTSKIFLQRQWGDP (15) [-1.28]
                                                        QIFLQRQWGDPEAEN (15) [-1.27]
                                                        KIFLQRQWGDPEAEF (15) [-0.87]
                                                            QRQWGDPEAENFFFN (15) [-1.45]
                                                            QRQWGDPEAEFFFFN (15) [-1.03]
                                                                GDPEAENFFFNCQGE (15) [-0.95]
                                                                GDPEAEFFFFNCQGE (15) [-0.53]
                                                                    AENFFFNCQGEFFYC (15) [0.11]
                                                                    AEFFFFNCQGEFFYC (15) [0.53]
                                                                        FFNCQGEFFYCKMDW (15) [-0.16]
                                                                            QGEFFYCKMDWFLNY (15) [-0.35]
```

```
    401                 421                 441                 461                 481
FYCKMDWFLNYLNNQSVDPDHNPCNKTGDKKCWQRTYVPCHIRQVVNVWYTLAKKTYAPPREGHLECTSTATALYVELNFDSNSRSNVTLS
FYCKMDWFLNYLNNQSVDPDHNPCNKNNDKKCWQRTYVPCHIRQVVNDWYTLSKKTYAPPREGHLECNSTATALYVELNYNSKNRTNVTLS
FYCKMDWFLNYLNNQ (15) [-0.49]
    MDWFLNYLNNQSVDP (15) [-0.61]
        LNYLNNQSVDPDHNP (15) [-1.41]
            NNQSVDPDHNPCNKT (15) [-1.97]
            NNQSVDPDHNPCNKN (15) [-2.16]
                VDPDHNPCNKTGDKK (15) [-2.00]
                VDPDHNPCNKNNDKK (15) [-2.39]
                    HNPCNKTGDKKCWQR (15) [-2.13]
                    HNPCNKNNDKKCWQR (15) [-2.53]
                        NKTGDKKCWQRTYVP (15) [-1.71]
                        NKNNDKKCWQRTYVP (15) [-2.10]
                            DKKCWQRTYVPCHIR (15) [-1.19]
                                WQRTYVPCHIRQVVN (15) [-0.51]
                                    YVPCHIRQVVNDWYT (15) [-0.29]
                                        HIRQVVNDWYTLAKK (15) [-0.69]
                                        HIRQVVNDWYTLSKK (15) [-0.87]
                                            VVNDWYTLAKKTYAP (15) [-0.37]
                                            VVNDWYTLSKKTYAP (15) [-0.54]
                                                WYTLAKKTYAPPREG (15) [-1.13]
                                                WYTLSKKTYAPPREG (15) [-1.30]
                                                    AKKTYAPPREGHLEC (15) [-1.21]
                                                    SKKTYAPPREGHLEC (15) [-1.39]
                                                        YAPPREGHLECTSTA (15) [-0.79]
                                                        YAPPREGHLECNSTA (15) [-0.98]
                                                            REGHLECTSTATALY (15) [-0.37]
                                                            REGHLECNSTATALY (15) [-0.56]
                                                                LECTSTATALYVELN (15) [0.47]
                                                                LECNSTATALYVELN (15) [0.28]
                                                                    STATALYVELNFDSN (15) [-0.01]
                                                                    STATALYVELNYNSK (15) [-0.31]
                                                                        ALYVELNFDSNSRSN (15) [-0.62]
                                                                        ALYVELNYNSKNRTN (15) [-1.09]
                                                                            ELNFDSNSRSNVTLS (15) [-0.75]
                                                                            ELNYNSKNRTNVTLS (15) [-1.23]
```

```
    481                 501                 521                 541                 561
DSNSRSNVTLSPQIRSIWSYELGDYKLVEIKPIGFAPTEVRRYTGPERQKRVPFVLGFLGFLGAAGAAMGAAATALTVQSQQLLAGILQQQ
NSKNRTNVTLSPQIRSIWANELGDYKLVEIKPIGFAPTEVRRYTGPERQKRVPFVLGFLGFLGAAGAAMGAAATALTVQSQQLLAGILQQQ
DSNSRSNVTLSPQIR (15) [-1.07]
NSKNRTNVTLSPQIR (15) [-1.27]
    RSNVTLSPQIRSIWS (15) [-0.36]
    RTNVTLSPQIRSIWA (15) [-0.18]
        TLSPQIRSIWSYELG (15) [-0.15]
        TLSPQIRSIWANELG (15) [-0.12]
            QIRSIWSYELGDYKL (15) [-0.52]
            QIRSIWANELGDYKL (15) [-0.49]
                IWSYELGDYKLVEIK (15) [-0.15]
                IWANELGDYKLVEIK (15) [-0.12]
                    ELGDYKLVEIKPIGF (15) [0.11]
                        YKLVEIKPIGFAPTE (15) [0.08]
                            EIKPIGFAPTEVRRY (15) [-0.51]
                                IGFAPTEVRRYTGPE (15) [-0.63]
                                    PTEVRRYTGPERQKR (15) [-2.30]
                                        RRYTGPERQKRVPFV (15) [-1.55]
                                            GPERQKRVPFVLGFL (15) [-0.15]
                                                QKRVPFVLGFLGFLG (15) [0.90]
                                                    PFVLGFLGFLGAAGA (15) [1.75]
                                                        GFLGFLGAAGAAMGA (15) [1.47]
                                                            FLGAAGAAMGAAATA (15) [1.40]
                                                                AGAAMGAAATALTVQ (15) [1.12]
                                                                    MGAAATALTVQSQQL (15) [0.52]
                                                                        ATALTVQSQQLLAGI (15) [0.83]
                                                                            TVQSQQLLAGILQQQ (15) [-0.07]
```

```
    561                 581                 601                 621                 641
QQLLAGILQQQKNLLAAVEQQQQMLKLTIWGVKNLNARVTALEKYLEDQARLNIWGCAFRQVCHTTVPWKFNNTPDWDNMTWQEWERQIAK
QQLLAGILQQQKNLL (15) [0.00]
    AGILQQQKNLLAAVE (15) [0.25]
        QQQKNLLAAVEQQQQ (15) [-1.33]
            NLLAAVEQQQQMLKL (15) [-0.00]
                AVEQQQQMLKLTIWG (15) [-0.23]
                    QQQMLKLTIWGVKNL (15) [-0.12]
                        LKLTIWGVKNLNARV (15) [0.32]
                            IWGVKNLNARVTALE (15) [0.21]
                                KNLNARVTALEKYLE (15) [-0.61]
                                    ARVTALEKYLEDQAR (15) [-0.78]
                                        ALEKYLEDQARLNIW (15) [-0.57]
                                            YLEDQARLNIWGCAF (15) [-0.01]
                                                QARLNIWGCAFRQVC (15) [0.21]
                                                    NIWGCAFRQVCHTTV (15) [0.34]
                                                        CAFRQVCHTTVPWKF (15) [0.12]
                                                            QVCHTTVPWKFNNTP (15) [-0.67]
                                                                TTVPWKFNNTPDWDN (15) [-1.43]
                                                                    WKFNNTPDWDNMTWQ (15) [-1.73]
                                                                        NTPDWDNMTWQEWER (15) [-2.19]
                                                                            WDNMTWQEWERQIAK (15) [-1.64]
```

```
    641                 661                 681                 701                 721
TWQEWERQIAKLEGNISRTLEQAHEQEQKNLDSYQKLVSWSDFWSWFDLTKWFGWMKIAIMVIAGIIIARILFVIISLLRKFRAGYAPLSS
TWQEWERQIAKLEGN (15) [-1.48]
    WERQIAKLEGNISRT (15) [-1.01]
        IAKLEGNISRTLEQA (15) [-0.27]
            EGNISRTLEQAHEQE (15) [-1.60]
                SRTLEQAHEQEQKNL (15) [-1.88]
                    EQAHEQEQKNLDSYQ (15) [-2.34]
                        EQEQKNLDSYQKLVS (15) [-1.56]
                            KNLDSYQKLVSWSDF (15) [-0.79]
                                SYQKLVSWSDFWSWF (15) [-0.30]
                                    LVSWSDFWSWFDLTK (15) [0.05]
                                        SDFWSWFDLTKWFGW (15) [-0.33]
                                            SWFDLTKWFGWMKIA (15) [0.11]
                                                LTKWFGWMKIAIMVI (15) [1.28]
                                                    FGWMKIAIMVIAGII (15) [2.09]
                                                        KIAIMVIAGIIIARI (15) [2.28]
                                                            MVIAGIIIARILFVI (15) [2.84]
                                                                GIIIARILFVIISLL (15) [2.77]
                                                                    ARILFVIISLLRKFR (15) [1.22]
                                                                        FVIISLLRKFRAGYA (15) [0.97]
```

```
                                                                            SLLRKFRAGYAPLSS (15) [-0.05]
    721                 741                 761                 781                 801
KFRAGYAPLSSLPSYNQQIHIHLDKGQPDKEEEQGSGGLNSDSRSYTWQREFLRHLCHRLITWLRNLTNWFLTIFSNLHRCLQVILQRIRQ
KFRAGYAPLSSLPSY (15) [-0.20]
    GYAPLSSLPSYNQQI (15) [-0.35]
        LSSLPSYNQQIHIHL (15) [-0.12]
            PSYNQQIHIHLDKGQ (15) [-1.27]
                QQIHIHLDKGQPDKE (15) [-1.63]
                    IHLDKGQPDKEEEQG (15) [-1.97]
                        KGQPDKEEEQGSGGL (15) [-1.93]
                            DKEEEQGSGGLNSDS (15) [-1.88]
                                EQGSGGLNSDSRSYT (15) [-1.41]
                                    GGLNSDSRSYTWQRE (15) [-1.69]
                                        SDSRSYTWQREFLRH (15) [-1.73]
                                            SYTWQREFLRHLCHR (15) [-1.18]
                                                QREFLRHLCHRLITW (15) [-0.49]
                                                    LRHLCHRLITWLRNL (15) [0.07]
                                                        CHRLITWLRNLTNWF (15) [-0.08]
                                                            ITWLRNLTNWFLTIF (15) [0.71]
                                                                RNLTNWFLTIFSNLH (15) [0.01]
                                                                    NWFLTIFSNLHRCLQ (15) [0.23]
                                                                        TIFSNLHRCLQVILQ (15) [0.68]
                                                                            NLHRCLQVILQRIRQ (15) [-0.24]
    801                 821                 841                 861                 881
CLQVILQRIRQLTAHLEYGWQEFKAAATSLAVAVVQAATSASDLIRHACRSIVRAVLAHPRRMRQELERWFNARRLINRRRYKNETSRKSG
CLQVILQRIRQLTAH (15) [0.37]
    ILQRIRQLTAHLEYG (15) [-0.19]
        IRQLTAHLEYGWQEF (15) [-0.55]
            TAHLEYGWQEFKAAA (15) [-0.47]
                EYGWQEFKAAATSLA (15) [-0.31]
                    QEFKAAATSLAVAVV (15) [1.05]
                        AAATSLAVAVVQAAT (15) [1.55]
                            SLAVAVVQAATSASD (15) [1.02]
                                AVVQAATSASDLIRH (15) [0.46]
                                    AATSASDLIRHACRS (15) [-0.05]
                                        ASDLIRHACRSIVRA (15) [0.21]
                                            IRHACRSIVRAVLAH (15) [0.56]
                                                CRSIVRAVLAHPRRM (15) [0.07]
                                                    VRAVLAHPRRMRQEL (15) [-0.55]
                                                        LAHPRRMRQELERWF (15) [-1.34]
                                                            RRMRQELERWFNARR (15) [-2.11]
                                                                QELERWFNARRLINR (15) [-1.31]
                                                                    RWFNARRLINRRRYK (15) [-1.81]
                                                                        ARRLINRRRYKNETS (15) [-1.97]
                                                                            INRRRYKNETSRKSG (15) [-2.39]

    881
RYKNETSRKSGRVVGR
RYKNETSRKSGRVVG (15) [-1.62]
```

```
    ETSRKSGRVVGR (13) [-1.20]

GAG.SABAEUS
                  Word length: 15
                  Overlap consecutive peptides by: 11
                  Shorten by: 3
                  Lengthen by: 2
                  Number of peptides generated: 136 + 39 = 175


1                   21                  41                  61                  81
MGASNSVLSGRKLDAFELVRLRPNGKKKYKLRHLVWASKELDRFGLSANLLETKEGVVKILSVLLPLVPTGSENLIALFNLCCVLAC
MGASNSVLSGRKLDAFESVRLRPNGKKKYKLRHLVWASKELDRFGLSANLLETKEGVVKILSVLLPLVPTGSENLIALFNLCCVLAC
MGASNSVLSGRKLDA (15) [-0.09]
    NSVLSGRKLDAFELV (15) [0.23]
    NSVLSGRKLDAFESV (15) [-0.07]
        SGRKLDAFELVRLRP (15) [-0.47]
        SGRKLDAFESVRLRP (15) [-0.77]
            LDAFELVRLRPNGKK (15) [-0.61]
            LDAFESVRLRPNGKK (15) [-0.91]
                ELVRLRPNGKKKYKL (15) [-1.29]
                ESVRLRPNGKKKYKL (15) [-1.59]
                    LRPNGKKKYKLRHLV (15) [-1.27]
                        GKKKYKLRHLVWASK (15) [-1.13]
                            YKLRHLVWASKELDR (15) [-0.84]
                                HLVWASKELDRFGLS (15) [-0.09]
                                    ASKELDRFGLSANLL (15) [0.05]
                                        LDRFGLSANLLETKE (15) [-0.30]
                                            GLSANLLETKEGVVK (15) [0.07]
                                                NLLETKEGVVKILSV (15) [0.55]
                                                    TKEGVVKILSVLLPL (15) [1.17]
                                                        VVKILSVLLPLVPTG (15) [1.83]
                                                            LSVLLPLVPTGSENL (15) [0.97]
                                                                LPLVPTGSENLIALF (15) [1.09]
                                                                    PTGSENLIALFNLCC (15) [0.77]
                                                                        ENLIALFNLCCVLAC (15) [1.82]
```

```
    81                  101                 121                 141                 161
ALFNLCCVLACIHAEIKVKDTEEAKAKIKQEVPLEMTESATVTSSGQKQELQQGKKNEAIAPSGGGSQNYPIVSVNNQWVHQPLSPRTLNA
ALFNLCCVLACVHAEIKVKDTEEAKAKVKQEVPAEMTESATATSSGQKQELQAKKKNEPTATSSGGSQNYPIVSVNNQWVHQPLSPRTLNA
ALFNLCCVLACIHAE (15) [1.71]
ALFNLCCVLACVHAE (15) [1.69]
    LCCVLACIHAEIKVK (15) [1.44]
    LCCVLACVHAEIKVK (15) [1.42]
        LACIHAEIKVKDTEE (15) [-0.17]
        LACVHAEIKVKDTEE (15) [-0.19]
            HAEIKVKDTEEAKAK (15) [-1.29]
                KVKDTEEAKAKIKQE (15) [-1.69]
                KVKDTEEAKAKVKQE (15) [-1.71]
                    TEEAKAKIKQEVPLE (15) [-1.03]
                    TEEAKAKVKQEVPAE (15) [-1.18]
                        KAKIKQEVPLEMTES (15) [-0.84]
                        KAKVKQEVPAEMTES (15) [-0.99]
                            KQEVPLEMTESATVT (15) [-0.43]
                            KQEVPAEMTESATAT (15) [-0.73]
                                PLEMTESATVTSSGQ (15) [-0.35]
                                PAEMTESATATSSGQ (15) [-0.65]
                                    TESATVTSSGQKQEL (15) [-0.87]
                                    TESATATSSGQKQEL (15) [-1.03]
                                        TVTSSGQKQELQQGK (15) [-1.41]
                                        TATSSGQKQELQAKK (15) [-1.45]
                                            SGQKQELQQGKKNEA (15) [-2.15]
                                            SGQKQELQAKKKNEP (15) [-2.25]
                                                QELQQGKKNEAIAPS (15) [-1.31]
                                                QELQAKKKNEPTATS (15) [-1.71]
                                                    QGKKNEAIAPSGGGS (15) [-1.00]
                                                    AKKKNEPTATSSGGS (15) [-1.42]
                                                        NEAIAPSGGGSQNYP (15) [-0.88]
                                                        NEPTATSSGGSQNYP (15) [-1.42]
                                                            APSGGGSQNYPIVSV (15) [-0.03]
                                                            ATSSGGSQNYPIVSV (15) [0.01]
                                                                GGSQNYPIVSVNNQW (15) [-0.72]
                                                                    NYPIVSVNNQWVHQP (15) [-0.65]
                                                                        VSVNNQWVHQPLSPR (15) [-0.73]
                                                                            NQWVHQPLSPRTLNA (15) [-0.91]
```

```
    161                 181                 201                 221                 241
HQPLSPRTLNAWVKVIEEKKFSAEVVPMFSALAEGAIPYDINQMLNAVGDHQGALQIVKDVINEEAADWDLRHPPPQQPPAQGVLREPQGS
HQPLSPRTLNAWVKVIEEKKFSAEVVPMFSALAEGAIPYDINQMLNAVGEHQGALQIVKDVINEEAADWDLRHPPPQQPPAQGVLRDPQGS
HQPLSPRTLNAWVKV (15) [-0.43]
    SPRTLNAWVKVIEEK (15) [-0.55]
        LNAWVKVIEEKKFSA (15) [-0.05]
            VKVIEEKKFSAEVVP (15) [0.09]
                EEKKFSAEVVPMFSA (15) [-0.13]
                    FSAEVVPMFSALAEG (15) [0.97]
                        VVPMFSALAEGAIPY (15) [1.17]
                            FSALAEGAIPYDINQ (15) [0.19]
                                AEGAIPYDINQMLNA (15) [-0.05]
                                    IPYDINQMLNAVGDH (15) [-0.22]
                                    IPYDINQMLNAVGEH (15) [-0.22]
                                        INQMLNAVGDHQGAL (15) [0.02]
                                        INQMLNAVGEHQGAL (15) [0.02]
                                            LNAVGDHQGALQIVK (15) [0.15]
                                            LNAVGEHQGALQIVK (15) [0.15]
                                                GDHQGALQIVKDVIN (15) [-0.16]
                                                GEHQGALQIVKDVIN (15) [-0.16]
                                                    GALQIVKDVINEEAA (15) [0.32]
                                                        IVKDVINEEAADWDL (15) [-0.07]
                                                            VINEEAADWDLRHPP (15) [-0.88]
                                                                EAADWDLRHPPPQQP (15) [-1.67]
                                                                    WDLRHPPPQQPPAQG (15) [-1.69]
                                                                        HPPPQQPPAQGVLRE (15) [-1.35]
                                                                        HPPPQQPPAQGVLRD (15) [-1.35]
                                                                            QQPPAQGVLREPQGS (15) [-1.24]
                                                                            QQPPAQGVLRDPQGS (15) [-1.24]
```

```
  241       261       281       301       321
AQGVLREPQGSDIAGTTSTIPEQIEWTTRAQNAINVGNIYKGWIILGLQKCVKMYNPVNILDIKQGPKEPFKDYVDRFYKALRAEQTDPAV
AQGVLRDPQGSDIAGTTSTIAEQIEWTTRAQNAINVGNIYKGWIILGLQKCVKMYNPVNILDIKQGPKEPFKDYVDRFYKALRAEQTDPAV
AQGVLREPQGSDIAG (15) [-0.40]
AQGVLRDPQGSDIAG (15) [-0.40]
  LREPQGSDIAGTTST (15) [-0.73]
  LRDPQGSDIAGTTST (15) [-0.73]
    QGSDIAGTTSTIPEQ (15) [-0.62]
    QGSDIAGTTSTIAEQ (15) [-0.39]
      IAGTTSTIPEQIEWT (15) [-0.11]
      IAGTTSTIAEQIEWT (15) [0.11]
        TSTIPEQIEWTTRAQ (15) [-0.92]
        TSTIAEQIEWTTRAQ (15) [-0.69]
          PEQIEWTTRAQNAIN (15) [-1.12]
          AEQIEWTTRAQNAIN (15) [-0.89]
            EWTTRAQNAINVGNI (15) [-0.53]
              RAQNAINVGNIYKGW (15) [-0.57]
                AINVGNIYKGWIILG (15) [0.90]
                  GNIYKGWIILGLQKC (15) [0.36]
                    KGWIILGLQKCVKMY (15) [0.47]
                      ILGLQKCVKMYNPVN (15) [0.22]
                        QKCVKMYNPVNILDI (15) [0.06]
                          KMYNPVNILDIKQGP (15) [-0.52]
                            PVNILDIKQGPKEPF (15) [-0.48]
                              LDIKQGPKEPFKDYV (15) [-1.02]
                                GPKEPFKDYVDRFYK (15) [-1.54]
                                  EPFKDYVDRFYKALR (15) [-1.07]
                                    DYVDRFYKALRAEQT (15) [-1.05]
                                      RFYKALRAEQTDPAV (15) [-0.72]
```

```
    321                 341                 361                 381                 401
ALRAEQTDPAVKNWMTQSLLIQNANPDCKTVLRGLGMNPTLEEMLTACQGIGGAQHKARLMAKAMSAAFQQQTVGNIFVQQGARPRGPPGG
ALRAEQTDPAVKNWMTQSLLIQNANPDCKTVLKGLGMNPTLEEMLTACQGIGGAQHKARLMAEAMSAAFQQQTVGNIFVQQGARPRGPPGG
ALRAEQTDPAVKNWM (15) [-0.69]
    EQTDPAVKNWMTQSL (15) [-0.96]
        PAVKNWMTQSLLIQN (15) [-0.13]
            NWMTQSLLIQNANPD (15) [-0.61]
                QSLLIQNANPDCKTV (15) [-0.26]
                    IQNANPDCKTVLRGL (15) [-0.30]
                    IQNANPDCKTVLKGL (15) [-0.26]
                        NPDCKTVLRGLGMNP (15) [-0.49]
                        NPDCKTVLKGLGMNP (15) [-0.45]
                            KTVLRGLGMNPTLEE (15) [-0.35]
                            KTVLKGLGMNPTLEE (15) [-0.31]
                                RGLGMNPTLEEMLTA (15) [-0.12]
                                KGLGMNPTLEEMLTA (15) [-0.08]
                                    MNPTLEEMLTACQGI (15) [0.19]
                                        LEEMLTACQGIGGAQ (15) [0.28]
                                            LTACQGIGGAQHKAR (15) [-0.29]
                                                QGIGGAQHKARLMAK (15) [-0.54]
                                                QGIGGAQHKARLMAE (15) [-0.51]
                                                    GAQHKARLMAKAMSA (15) [-0.24]
                                                    GAQHKARLMAEAMSA (15) [-0.21]
                                                        KARLMAKAMSAAFQQ (15) [-0.05]
                                                        KARLMAEAMSAAFQQ (15) [-0.02]
                                                            MAKAMSAAFQQQTVG (15) [0.11]
                                                            MAEAMSAAFQQQTVG (15) [0.14]
                                                                MSAAFQQQTVGNIFV (15) [0.54]
                                                                    FQQQTVGNIFVQQGA (15) [-0.15]
                                                                        TVGNIFVQQGARPRG (15) [-0.37]
                                                                            IFVQQGARPRGPPGG (15) [-0.61]
```

```
    401                 421                 441                 461                 481
QGARPRGPPGGRGRPMNPNIKCYNCGKPGHLARFCKAPRRQGCWKCGSPDHQMKDCQKQVNFLGFGPWGRGKPRNFPLTSIRPTAPPMERD
QGARPRGPPGGRGRP (15) [-1.87]
    PRGPPGGRGRPMNPN (15) [-1.88]
        PGGRGRPMNPNIKCY (15) [-1.22]
            GRPMNPNIKCYNCGK (15) [-1.11]
                NPNIKCYNCGKPGHL (15) [-0.90]
                    KCYNCGKPGHLARFC (15) [-0.45]
                        CGKPGHLARFCKAPR (15) [-0.59]
                            GHLARFCKAPRRQGC (15) [-0.75]
                                RFCKAPRRQGCWKCG (15) [-1.07]
                                    APRRQGCWKCGSPDH (15) [-1.47]
                                        QGCWKCGSPDHQMKD (15) [-1.48]
                                            KCGSPDHQMKDCQKQ (15) [-1.89]
                                                PDHQMKDCQKQVNFL (15) [-1.23]
                                                    MKDCQKQVNFLGFGP (15) [-0.41]
                                                        QKQVNFLGFGPWGRG (15) [-0.63]
                                                            NFLGFGPWGRGKPRN (15) [-1.08]
                                                                FGPWGRGKPRNFPLT (15) [-0.97]
                                                                    GRGKPRNFPLTSIRP (15) [-1.13]
                                                                        PRNFPLTSIRPTAPP (15) [-0.65]
                                                                            PLTSIRPTAPPMERD (15) [-0.84]
```

```
   481                 501                 521                 541
IRPTAPPMERDYSRPEENWYADRPPTRGPGPDDPATALLKQYAVQGRRQRQQRQSCPPQQSPYEEAYSSLRSLFGEDQ
IRPTAPPMERDYSRP (15) [-1.43]
    APPMERDYSRPEENW (15) [-2.04]
        ERDYSRPEENWYADR (15) [-2.57]
            SRPEENWYADRPPTR (15) [-2.28]
                ENWYADRPPTRGPGP (15) [-1.85]
                    ADRPPTRGPGPDDPA (15) [-1.69]
                        PTRGPGPDDPATALL (15) [-0.59]
                            PGPDDPATALLKQYA (15) [-0.57]
                                DPATALLKQYAVQGR (15) [-0.38]
                                    ALLKQYAVQGRRQRQ (15) [-1.18]
                                        QYAVQGRRQRQQRQS (15) [-2.37]
                                            QGRRQRQQRQSCPPQ (15) [-2.73]
                                                QRQQRQSCPPQQSPY (15) [-2.35]
                                                    RQSCPPQQSPYEEAY (15) [-1.78]
                                                        PPQQSPYEEAYSSLR (15) [-1.51]
                                                            SPYEEAYSSLRSLFG (15) [-0.47]
                                                                EAYSSLRSLFGEDQ (15) [-0.69]
```

```
GAG.VERVET
                  Word length: 15
                  Overlap consecutive peptides by: 11
                  Shorten by: 3
                  Lengthen by: 2
                  Number of peptides generated: 128 (-1) = 127

1                   21                  41                  61                  81
MGAATSALNRRQLDQFEHIRLRPNGKKKYQIKHLIWAGKEMERFGLHERLLETEEGCKKIIEVLYPLEPTGSEGLKSLFNLVCVLYC
MGAATSALNRRQLDQ (15) [-0.67]
    TSALNRRQLDQFEHI (15) [-0.97]
        NRRQLDQFEHIRLRP (15) [-1.69]
            LDQFEHIRLRPNGKK (15) [-1.41]
                EHIRLRPNGKKKYQI (15) [-1.66]
                    LRPNGKKKYQIKHLI (15) [-1.13]
                        GKKKYQIKHLIWAGK (15) [-0.97]
                            YQIKHLIWAGKEMER (15) [-0.81]
                                HLIWAGKEMERFGLH (15) [-0.33]
                                    AGKEMERFGLHERLL (15) [-0.63]
                                        MERFGLHERLLETEE (15) [-0.98]
                                            GLHERLLETEEGCKK (15) [-1.14]
                                                RLLETEEGCKKIIEV (15) [-0.27]
                                                    TEEGCKKIIEVLYPL (15) [0.07]
                                                        CKKIIEVLYPLEPTG (15) [0.19]
                                                            IEVLYPLEPTGSEGL (15) [0.19]
                                                                YPLEPTGSEGLKSLF (15) [-0.29]
                                                                    PTGSEGLKSLFNLVC (15) [0.35]
                                                                        EGLKSLFNLVCVLYC (15) [1.20]
    81                  101                 121                 141                 161
SLFNLVCVLYCIHKEQKVKDTEEAVATVRQCCHLVEKEKNATENTTEPSSGQKKNDKGTTAPPGGSQNFPAQQQGNAWVHVPLSPRTLNAW
SLFNLVCVLYCIHKE (15) [1.06]
    LVCVLYCIHKEQKVK (15) [0.43]
        LYCIHKEQKVKDTEE (15) [-1.29]
            HKEQKVKDTEEAVAT (15) [-1.45]
                KVKDTEEAVATVRQC (15) [-0.60]
                    TEEAVATVRQCCHLV (15) [0.36]
                        VATVRQCCHLVEKEK (15) [-0.23]
                            RQCCHLVEKEKNATE (15) [-1.26]
                                HLVEKEKNATENTTE (15) [-1.62]
                                    KEKNATENTTEPSSG (15) [-1.95]
                                        ATENTTEPSSGQKKN (15) [-1.95]
                                            TTEPSSGQKKNDKGT (15) [-2.12]
                                                SSGQKKNDKGTTAPP (15) [-1.83]
                                                    KKNDKGTTAPPGGSQ (15) [-1.80]
                                                        KGTTAPPGGSQNFPA (15) [-0.85]
                                                            APPGGSQNFPAQQQG (15) [-1.19]
                                                                GSQNFPAQQQGNAWV (15) [-0.97]
```

```
                                                                    FPAQQQGNAWVHVPL (15) [-0.21]
                                                                        QQGNAWVHVPLSPRT (15) [-0.68]
                                                                            AWVHVPLSPRTLNAW (15) [0.13]

    161                 181                 201                 221                 241
VPLSPRTLNAWVKAVEEKKFGAEIVPMFQALSEGCTPYDINQMLNVLGDHQGALQIVKEIINEEAAQWDVTHPPPAGPLPAGQLRDPRGSD
VPLSPRTLNAWVKAV (15) [0.42]
    PRTLNAWVKAVEEKK (15) [-0.94]
        NAWVKAVEEKKFGAE (15) [-0.69]
            KAVEEKKFGAEIVPM (15) [-0.20]
                EKKFGAEIVPMFQAL (15) [0.22]
                    GAEIVPMFQALSEGC (15) [0.64]
                        VPMFQALSEGCTPYD (15) [0.01]
                            QALSEGCTPYDINQM (15) [-0.52]
                                EGCTPYDINQMLNVL (15) [-0.05]
                                    PYDINQMLNVLGDHQ (15) [-0.62]
                                        NQMLNVLGDHQGALQ (15) [-0.38]
                                            NVLGDHQGALQIVKE (15) [-0.21]
                                                DHQGALQIVKEIINE (15) [-0.35]
                                                    ALQIVKEIINEEAAQ (15) [0.13]
                                                        VKEIINEEAAQWDVT (15) [-0.37]
                                                            INEEAAQWDVTHPPP (15) [-0.99]
                                                                AAQWDVTHPPPAGPL (15) [-0.35]
                                                                    DVTHPPPAGPLPAGQ (15) [-0.54]
                                                                        PPPAGPLPAGQLRDP (15) [-0.71]
                                                                            GPLPAGQLRDPRGSD (15) [-1.13]

    241                 261                 281                 301                 321
AGQLRDPRGSDIAGTTSTVQEQLEWIYTANPRVDVGAIYRRWIILGLQKCVKMYNPVSVLDIRQGPKEPFKDYVDRFYKAIRAEQASGEVK
AGQLRDPRGSDIAGT (15) [-0.79]
    RDPRGSDIAGTTSTV (15) [-0.77]
        GSDIAGTTSTVQEQL (15) [-0.28]
            AGTTSTVQEQLEWIY (15) [-0.35]
                STVQEQLEWIYTANP (15) [-0.61]
                    EQLEWIYTANPRVDV (15) [-0.53]
                        WIYTANPRVDVGAIY (15) [0.22]
                            ANPRVDVGAIYRRWI (15) [-0.25]
                                VDVGAIYRRWIILGL (15) [1.05]
                                    AIYRRWIILGLQKCV (15) [0.71]
                                        RWIILGLQKCVKMYN (15) [0.22]
                                            LGLQKCVKMYNPVSV (15) [0.38]
                                                KCVKMYNPVSVLDIR (15) [0.15]
                                                    MYNPVSVLDIRQGPK (15) [-0.40]
                                                        VSVLDIRQGPKEPFK (15) [-0.51]
                                                            DIRQGPKEPFKDYVD (15) [-1.55]
                                                                GPKEPFKDYVDRFYK (15) [-1.54]
                                                                    PFKDYVDRFYKAIRA (15) [-0.67]
                                                                        YVDRFYKAIRAEQAS (15) [-0.66]
                                                                            FYKAIRAEQASGEVK (15) [-0.56]
```

```
    321                 341                 361                 381                 401
IRAEQASGEVKQWMTESLLIQNANPDCKVILKGLGMHPTLEEMLTACQGVGGPSYKAKVMAEMMQNMQTQNMVQQGGGRGRQRPPLKCYNC
IRAEQASGEVKQWMT (15) [-0.73]
    QASGEVKQWMTESLL (15) [-0.40]
        EVKQWMTESLLIQNA (15) [-0.25]
            WMTESLLIQNANPDC (15) [-0.21]
                SLLIQNANPDCKVIL (15) [0.57]
                    QNANPDCKVILKGLG (15) [-0.24]
                        PDCKVILKGLGMHPT (15) [0.10]
                            VILKGLGMHPTLEEM (15) [0.45]
                                GLGMHPTLEEMLTAC (15) [0.37]
                                    HPTLEEMLTACQGVG (15) [0.03]
                                        EEMLTACQGVGGPSY (15) [-0.13]
                                            TACQGVGGPSYKAKV (15) [-0.16]
                                                GVGGPSYKAKVMAEM (15) [-0.03]
                                                    PSYKAKVMAEMMQNM (15) [-0.44]
                                                        AKVMAEMMQNMQTQN (15) [-0.68]
                                                            AEMMQNMQTQNMVQQ (15) [-1.01]
                                                                QNMQTQNMVQQGGGR (15) [-1.53]
                                                                    TQNMVQQGGGRGRQR (15) [-1.81]
                                                                        VQQGGGRGRQRPPLK (15) [-1.65]
                                                                            GGRGRQRPPLKCYNC (15) [-1.42]
```

```
    401                 421                 441                 461                 481
RQRPPLKCYNCGKFGHMQRQCPEPRKIKCLKCGKPGHLAKDCRGQVNFLGYGRWMGAKPRNFPAATLGAEPSAPPPPTPYDPAKKLLQQYA
RQRPPLKCYNCGKFG (15) [-1.17]
    PLKCYNCGKFGHMQR (15) [-0.85]
        YNCGKFGHMQRQCPE (15) [-1.31]
            KFGHMQRQCPEPRKI (15) [-1.49]
                MQRQCPEPRKIKCLK (15) [-1.28]
                    CPEPRKIKCLKCGKP (15) [-0.87]
                        RKIKCLKCGKPGHLA (15) [-0.45]
                            CLKCGKPGHLAKDCR (15) [-0.56]
                                GKPGHLAKDCRGQVN (15) [-1.10]
                                    HLAKDCRGQVNFLGY (15) [-0.35]
                                        DCRGQVNFLGYGRWM (15) [-0.51]
                                            QVNFLGYGRWMGAKP (15) [-0.39]
                                                LGYGRWMGAKPRNFP (15) [-0.85]
                                                    RWMGAKPRNFPAATL (15) [-0.51]
                                                        AKPRNFPAATLGAEP (15) [-0.50]
                                                            NFPAATLGAEPSAPP (15) [-0.10]
                                                                ATLGAEPSAPPPPTP (15) [-0.43]
                                                                    AEPSAPPPPTPYDPA (15) [-1.04]
                                                                        APPPPTPYDPAKKLL (15) [-0.78]
                                                                            PTPYDPAKKLLQQYA (15) [-1.01]
```

```
    481                 501                 521
DPAKKLLQQYAEKGKQLREQRKRPPAVNPDWTEGYSLNSLFGEDQ
DPAKKLLQQYAEKGK (15) [-1.45]
    KLLQQYAEKGKQLRE (15) [-1.48]
        QYAEKGKQLREQRKR (15) [-2.59]
            KGKQLREQRKRPPAV (15) [-1.97]
                LREQRKRPPAVNPDW (15) [-1.82]
                    RKRPPAVNPDWTEGY (15) [-1.70]
                        PAVNPDWTEGYSLNS (15) [-0.82]
                            PDWTEGYSLNSLFGE (15) [-0.70]
                                EGYSLNSLFGEDQ (14) [-0.77]
```