

## Human T-cell leukemia virus type I

### Aufgabe 2

[<https://www.ncbi.nlm.nih.gov/nuccore/9626453>]

*Die erste 100 Basenpaare der kodierende Sequenz.*

```
a tgggccaaat cttttcccg agcgctagcc  
ctattccgcg gccgccccgg gggctggccg ctcactctg gcttaacttc ctccaggcgg  
catatcgcc
```

### Aufgabe 3

*Die erste 1000 Basenpaare der kodierenden Sequenz.*

```
a  
tgggccaaat cttttcccg agcgctagcc ctattccgcg gccgccccgg gggctggccg ctcactctg  
gcttaacttc ctccaggcgg catatcgcc agaaccgggt cctccagtt acgatttcca ccagttaaaa  
aaatttttta aaatagcttt agaaacaccg gtctggatct gcccattaa ctactcctc ctaccagcc  
tactccaaa aggatacccc ggccgggtga atgaaatctt acacatactc atccaaaccc aagccagat  
cccgtccgc cccgcgcgc cgcgcgcgc atcctccacc cagaccccc cggatttga cccacaaatc  
ccccctcct atgttgagcc tacagcccc caagtcttc cagtcatgca cccacatggt gcccctcca  
accaccgcc atggcaaatg aaagacctac aggcattaa gcaagaagtc tccaagcgg ccctggaag  
ccccagttt atgcagacca tccggcttgc ggtgcagcag ttgaccca ctgccaaga cctccaagac  
ctctgcagt accttgcct ctcctctgtg gttccctcc atcaccagca gctagatagc cttatatcag  
aggccgaaac tcgaggtatt acaggttata accccttagc cgtgccctc cgtgtccaag ccaacaatcc  
acaacaacaa ggattaaggc gagaatacca gcaactctgg ctgccgcct tcgccgcct gccagggagt  
gccaaagacc ctctctgggc ctctatctc caaggcctgg aggagcctta ccacgcctc gtagaacgcc  
tcaacatage tcttgacaat gggctgccag aaggcagcc caaagacccc atttacgtt ccttagccta  
ctctaagca aacaaagaat gccaaaaatt actacaggcc cgagggcaca ctaatagccc tctaggagat  
atgttgccgg cttgtcagg
```

*Ersten 30 Aminosäuren von 1.5'3' Frame*

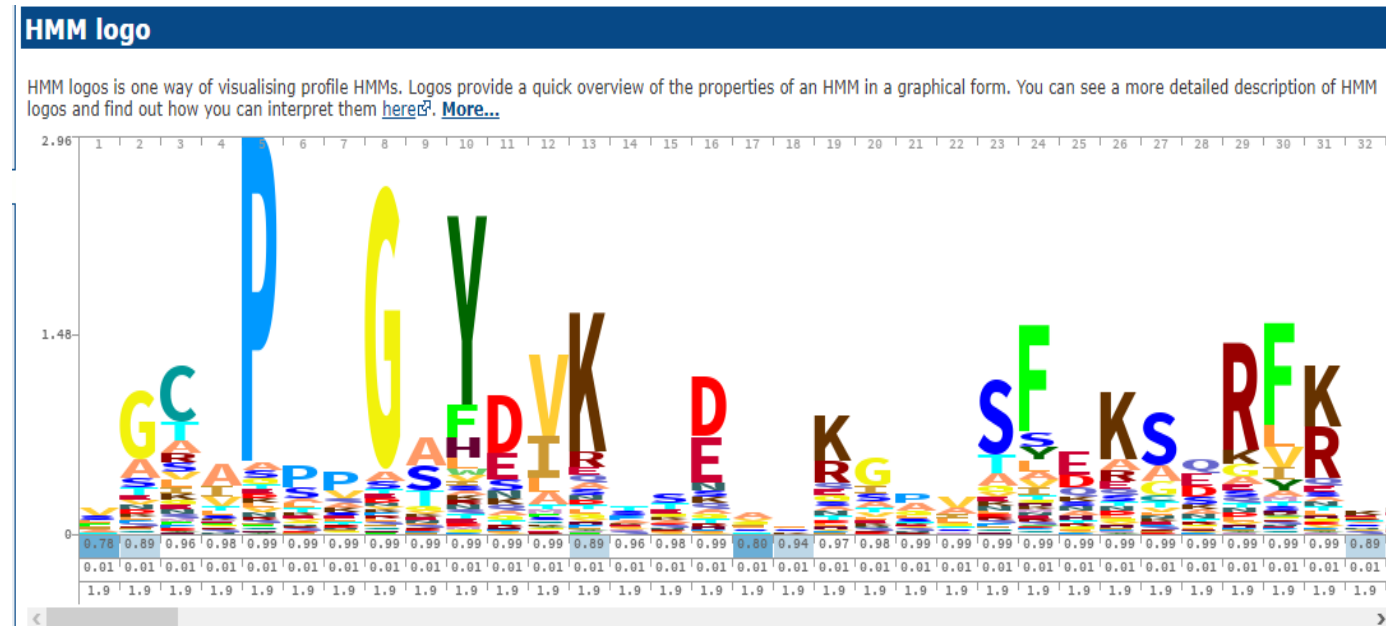
[<https://web.expasy.org/cgi-bin/translate/dna2aa.cgi>]

MGQIFSRASPIRPPRGLAAHHWLNFLQA

- Bei Aminosäuresequenzen wird ersichtlich, wo sich die kodierenden Sequenzen/ Exons befinden. Zudem kann man erkennen, wo das Exon beginnt- bei einer Aminosäure ist dies eher zu sehen. Es wird sichtbar, ob es verschiedene Reading-Frames gibt.
- Einerseits gibt es drei Reading-Frames, jeweils von 5' -> 3' Richtung. Es ist jedoch vorzuziehen, dass auch die Sequenz von 3' -> 5' zu übersetzen.

## Aufgabe 4

[<http://pfam.xfam.org/family/PF15905.4#tabview=tab4>]



**Abb.1. HMM Logo des gefundenen Profils aus Pfam Datenbank**

Im vergleich mit

MGQIFSRASPIRPPRGLAAHHWLNFLQA

-GCAPPPGAYDVK--D--KG--SFEKS-RF, es gibt nur es gleiche Aminosäure an 2.und 9. Stelle.

(,,-“ = Die Wahrscheinlichkeit der vorkommende Aminosäure an dieser Stellen könnte nicht bestimmen werden). Die anderen vorkommenden Aminosäuren besitzen höheren Wahrscheinlichkeit und wären erkennbar.

## Aufgabe 5

### **5.2 Hepatitis C virus genotype 1,(NC 004102.1)**

[ <https://www.ncbi.nlm.nih.gov/nucore/22129792> ]

*Meine erste 100 Basenpaare*

```
atgagcacg aatcctaaac ctcaaagaaa aaccaaactg aacaccaacc gtcgcccaca ggacgtcaag  
ttcccgggtg gcgggtcagat cgttggtgga g
```

### 5.3 Meine erste 1000 Baasenpaare

```
atgagcacg aatcctaaac|ctcaaagaaa aaccaaacgt aacaccaacc gtcgcccaca ggacgtcaag
ttcccgggtg|gcggtcagat cgttgggtgga gtttacttgt tgccgcgcag gggccctaga ttgggtgtgc
gcgcgacgag gaagacttcc gagcggtcgc aacctcgagg tagacgtcag cctatcccca
aggcacgtcg gcccaggggc aggacctggg ctcagcccgg gtacccttgg cccctctatg
gcaatgaggg ttgcgggtgg gcgggatggc tectgtctcc ccgtggctct cggcctagct|ggggccccac
agacccccgg cgtaggtcgc gcaatttggg taaggtcacc gataccctta|cgtgcggctt cgccgacctc
atggggtaca taccgctcgt cggcgccct cttggaggcg|ctgccagggc cctggcgcac ggcgccggg
ttctggaaga cggcgtgaac tatgaacag|ggaaccttcc tggttgctct ttctctatct tccttctggc
cctgctctct tgctgactg|tgcccgttc agcctacca gtcgcgaatt cctcgggggt ttaccatgtc
accaatgatt|gccctaactc gagtatttg tacgaggcgg ccgatgcat cctgcacact ccggggtgtg
tccttgcgt tcgcgagggt aacgcctcga ggtgttgggt ggcggtgacc cccacggtgg|ccaccaggga
cggcaaactc cccacaacgc agcttcgacg tcatatgat ctgcttgcg|ggagcgccac cctctgctcg
gccctctacg tgggggacct gtgcgggtct gtcttcttg|ttgtcaact gttacctc tcctccaggc
gccactggac gacgcaagac tgcaattgtt|ctatctatcc cggccatata acgggtcacc gcattggcatg
ggatatgatg atgaactggt|cccctacggc agcgttggtg g
```

### Erste 30 Aminosäuren von 1.5'3' Frame

[<https://web.expasy.org/cgi-bin/translate/dna2aa.cgi>]

MSTNPKPQRKTKRNTNRRPQDVKFPGGGQI

### 5.4

[<http://pfam.xfam.org/family/PF15905.4#tabview=tab4>]

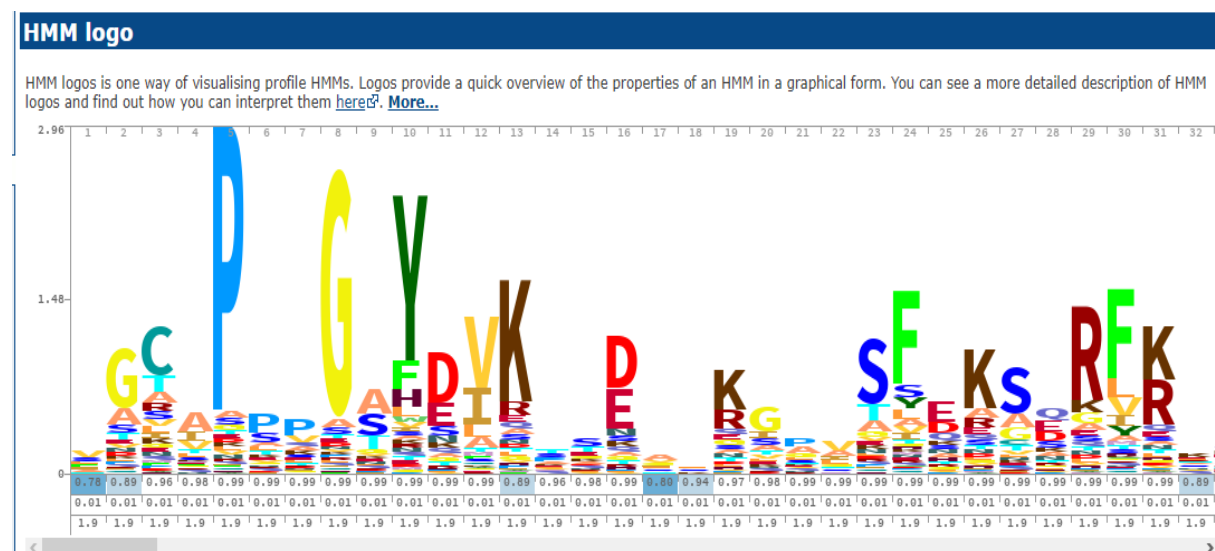


Abb.2. HMM Logo des gefundenen Profils aus Pfam Datenbank

Im Vergleich mit

MSTNPKPQRKTKRNTNRRPQDVKFPGGGQI

-GCAPPPGAYDVK--D--KG--SFEKS-RF, es gibt nur es gleiche Aminosäure an 5., 7. und 24. Stelle.

(„-“ = Die Wahrscheinlichkeit der vorkommende Aminosäure an dieser Stellen könnte nicht bestimmen werden). Die anderen vorkommenden Aminosäuren besitzen höheren Wahrscheinlichkeit und wären erkennbar.