Volume 12 Number 20 1984 Nucleic Acids Research

The complete nucleotide sequence of a common cold virus: human rhinovirus 14

Glyn Stanway, Pamela J.Hughes, Roger C.Mountford, Philip D.Minor+ and Jeffrey W.Almond

Department of Microbiology, University of Leicester, University Road, Leicester LE1 7RH, and +National Institute for Biological Standards and Control, Holly Hill, London NW3 6RB, UK

Received 6 August 1984; Revised and Accepted 24 September 1984

ABSTRACT

The complete nucleotide sequence of the single-stranded RNA genome of human rhinovirus 14, one of the causative agents of the common cold, has been determined from cDNA cloned in *E.coli*. The genome is typical of the picornaviridae family, comprising a 5′ non-coding region of 624 nucleotides, a long open reading frame of 6537 nucleotides (90.8 % of the genome) and a 3′ non-coding region of 47 nucleotides. Comparison of the nucleotide sequence and the predicted amino acid sequence with those of the polioviruses reveals a surprising degree of homology which may allow recognition of regions of antigenic importance and prediction of the virus polyprotein cleavage sites. The results presented here imply a closer genetic relationship between the rhinovirus and enterovirus genera than previously suspected.

INTRODUCTION

Human rhinoviruses are the major causative agents of the upper respiratory tract infections collectively known as the common cold, one of the most common virus infections of man (1). The high incidence of the disease can be explained, at least partially, by the fact that several of the 115 immunologically distinct known rhinovirus serotypes can co-circulate within a community (2,3). Since this serotype diversity effectively precludes a vaccination program based on conventional methods, the elucidation of its molecular basis is one of the most important problems in rhinovirus research.

Rhinoviruses, including human rhinoviruses, form one genus of the family picornaviridae (4). They share the common features of this family, namely a 25nm capsid of icosohedral symmetry, made up of 60 copies of each of 4 virus coded proteins (VP1-4) and enclosing a single-stranded RNA genome of approximately 7500 nucleotides (5-7). The RNA is of positive polarity, is poly-adenylated at its 3′ terminus and has a small protein, VPg, covalently attached to the 5′ terminus (8,9). The genomes of representatives of each of the other three genera of picornaviridae, enterovirus (10-14), aphthovirus

(15), and cardiovirus (16), have been sequenced, but to date no complete sequence of a rhinovirus genome has been published. As part of a study into the molecular basis of serotype diversity of rhinoviruses, we have determined the complete nucleotide sequence of human rhinovirus 14 (HRV-14). The sequence presented here allows comparisons to be made between each of the picornaviridae genera and our results indicate that HRV-14 is closely related to the enteroviruses.

MATERIALS AND METHODS

Virus

Human rhinovirus 14 (HRV-14), obtained from the MRC Common Cold Unit, Salisbury, U.K., was propagated at 33°C in Ohio HeLa cells grown in roller tubes. The virus was purified by adding 1% NP40 to the cleared tissue culture supernatant and centrifuging (100,000g, 4hours) through a 15-45% sucrose gradient.

Molecular cloning

The purification of RNA from HRV-14 and the cloning of cDNA.RNA hybrids into *E.coli* JA221 were as described previously (17,18). Transformants of phenotype $Tet^{r}Amp^{s}$ were tested for the presence of HRV-14 specific cDNA by hybridization using the method of Grunstein and Hogness (19). Initially, a radioactive probe was synthesized by using random oligonucleotides, generated by DNase1 digestion of salmon sperm DNA, to prime reverse transcription of HRV-14 RNA. Subsequently, probes were produced by labelling cDNA by nick translation.

Nucleotide Sequence Analysis

Five overlapping cDNAs, together representing the entire genome of HRV-14, were excised from the plasmid vector pAT 153 with PstI and the virus specific fragments isolated. The cDNA fragment representing the 3′ terminus was sub-cloned into PstI-digested M13 mp9. The four remaining fragments were each circularized by ligation and then sheared by sonication. The random fragments thus generated were end-repaired by treatment with T_4 DNA polymerase I and fractionated on a 1.5% agarose gel. DNA in the size range 300-1000 base pairs was electroeluted, sub-cloned into SmaI-digested, phosphatase-treated M13 mp8 and sequenced by the dideoxynucleotide method (20,21).

Nucleotide sequence data thus obtained were collated and assembled with the aid of a Digital PDP 11/44 computer using programs developed by Staden (22).

RESULTS AND DISCUSSION

Molecular Cloning

RNA extracted from HRV-14 was analyzed by gel electrophoresis on a 1% agarose gel and shown to consist of a discrete band corresponding to full length RNA, although considerable degradation was also observed (data not shown). An estimated 1.5µg of this material was reverse transcribed, yielding 350ng of cDNA. After dC-homopolymeric tailing of the cDNA.RNA hybrid molecules, 40% of the material (140ng of cDNA) was annealed to 300ng of dG-tailed, PstI-cut pAT 153 and used to transform competent *E.coli* JA221. 674 colonies of phenotype Tet^r Amp^s were produced, equivalent to 1.1×10^3 /µg of RNA. This result is typical for the cloning of RNA virus genomes using this method when *E.coli* JA221, with a routine transformation efficiency of 5×10^5/µg of supercoiled pAT 153, is used as the host strain (17,18).

Analysis of the colonies by filter hybridization, using the randomly-primed cDNA probe prepared as described in Methods, identified the clones pAM 1-4 and pAM 6-9 represented in Fig.1 (in addition to several others not shown). These were positioned relative to each other by cross hybridization and restriction enzyme mapping and were found to overlap to give a contiguous stretch of DNA approximately 7200 nucleotides in length. The orientation of the cloned DNA was determined by using a cDNA probe enriched for 3′ sequences (17). pAM 4 was found, by sequence analysis, to terminate short of the 3′ terminus since it did not contain a poly A tract. When pAM 5, selected on the basis of its hybridization with the 3′-enriched

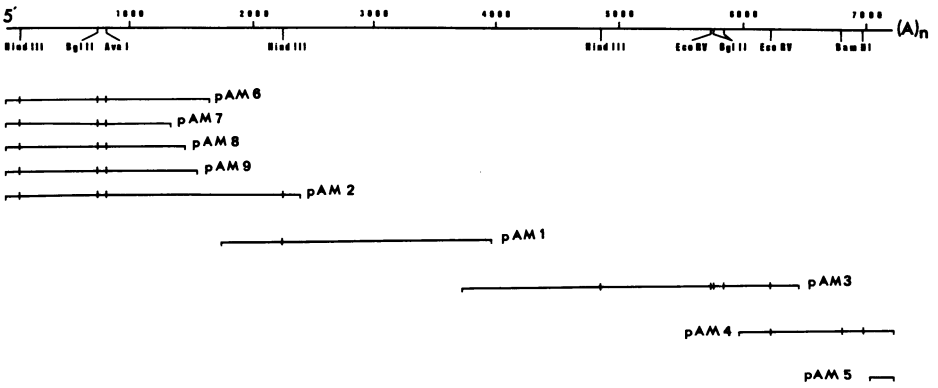


Figure 1.Overlapping cDNA clones produced by the cDNA.RNA method, spanning the genome of HRV-14. pAM 6-9 terminate at the same point, presumed to be the 5′ terminus. pAM 1-5 were used to determine the complete genome sequence.

probe, was sequenced, it was found to contain a poly A tract of 20 residues and was therefore deemed to include the 3′ terminus. Five of the clones analyzed were found, by restriction enzyme mapping, to form a nested set at the 5′ terminus of the cDNA. Sequence analysis revealed that these terminated at a point, which by its extensive homology with poliovirus (the final 10 bases are identical [10-14]) was presumed to be the 5′ terminus of the genome.

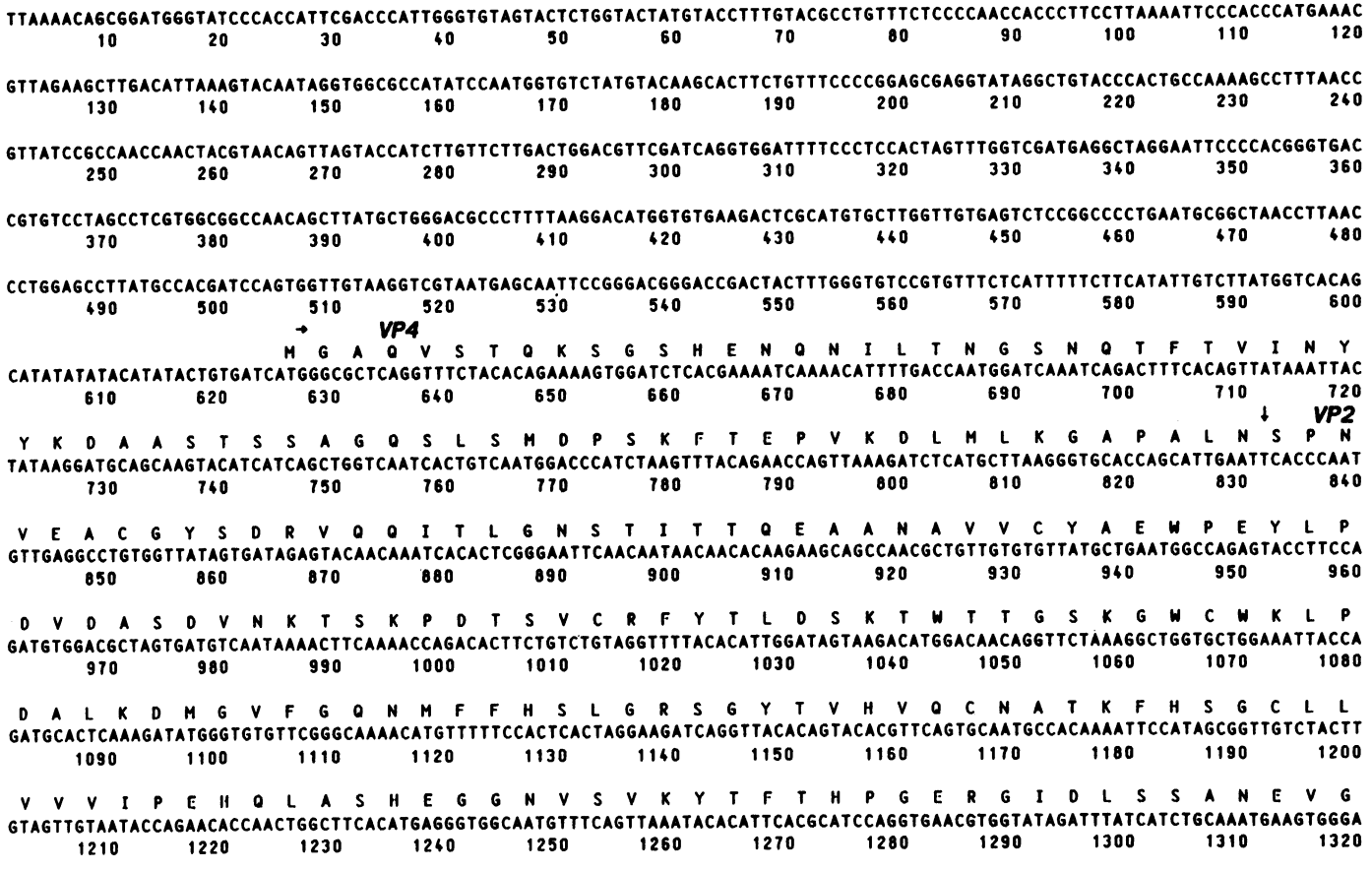
Nucleotide sequencing

The majority of the nucleotide sequence of the cloned cDNA was determined from the clones pAM 1-4. The sequence of the 3′ terminal 15 bases was obtained from pAM 5. The complete genome sequence of HRV-14, together with the predicted amino acid sequence of the major open reading frame is presented in Fig.2 (protein nomenclature is in accordance with the L434 system proposed by Rueckert and Wimmer [23]). The nucleotide sequence was derived from 140 individual gel readings and approximately 75% of the genome was sequenced in both orientations. Other regions were sequenced at least twice and contained no compressions or otherwise ambiguous regions. In addition, comparison with the other picornaviruses completely sequenced in our laboratory provided a useful reading frame check (13,14,unpublished). We are therefore confident that the presented sequence is accurate. The HRV-14 genome consists of 7208 nucleotides plus a 3′ poly A tract previously estimated to be 74-150 residues in length (8,24-26). The overall proportion of the nucleotides is A=32.1%, C=20.2%, G=20.4% and T=27.3%.

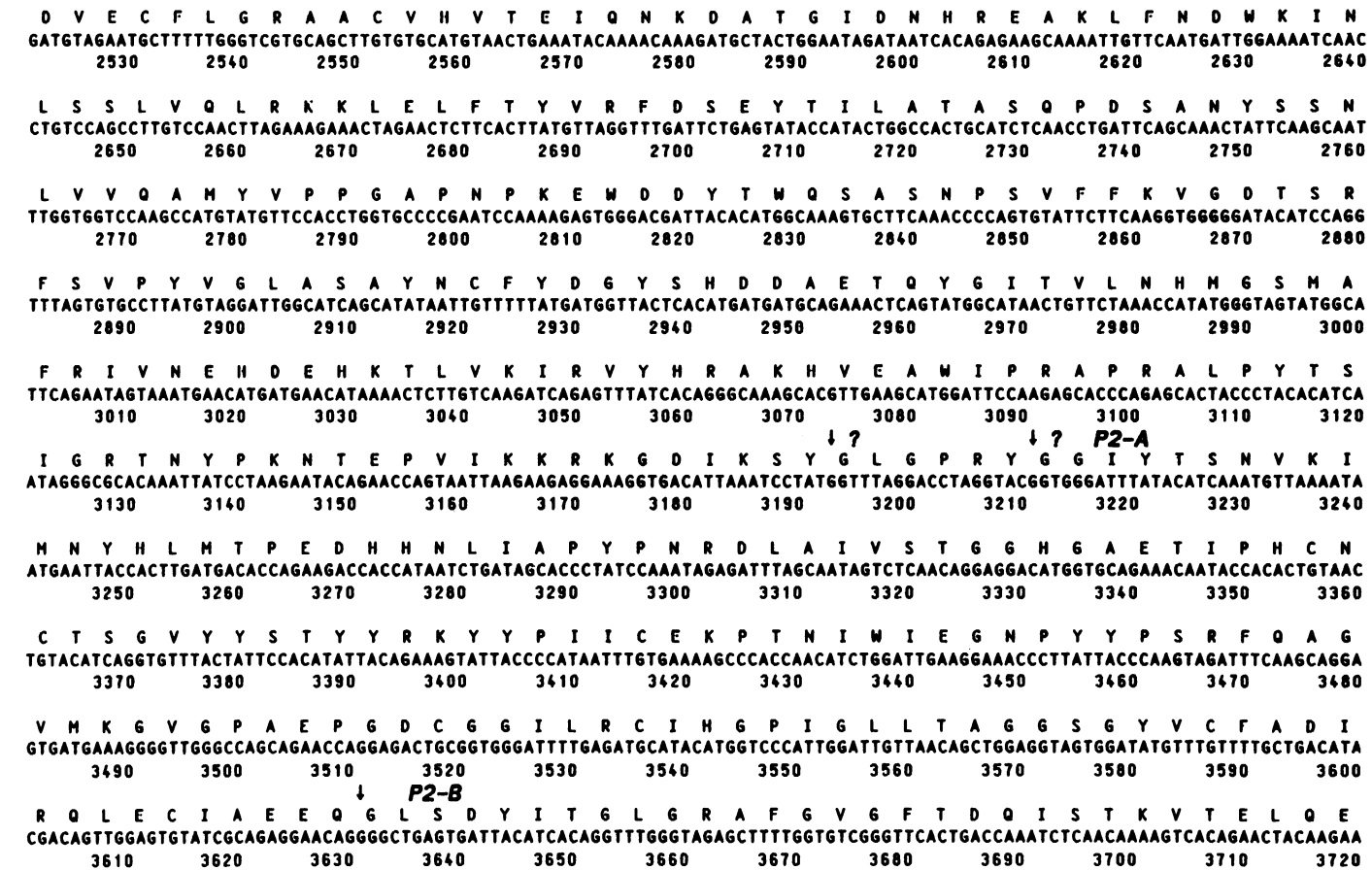
Comparison with other picornaviruses

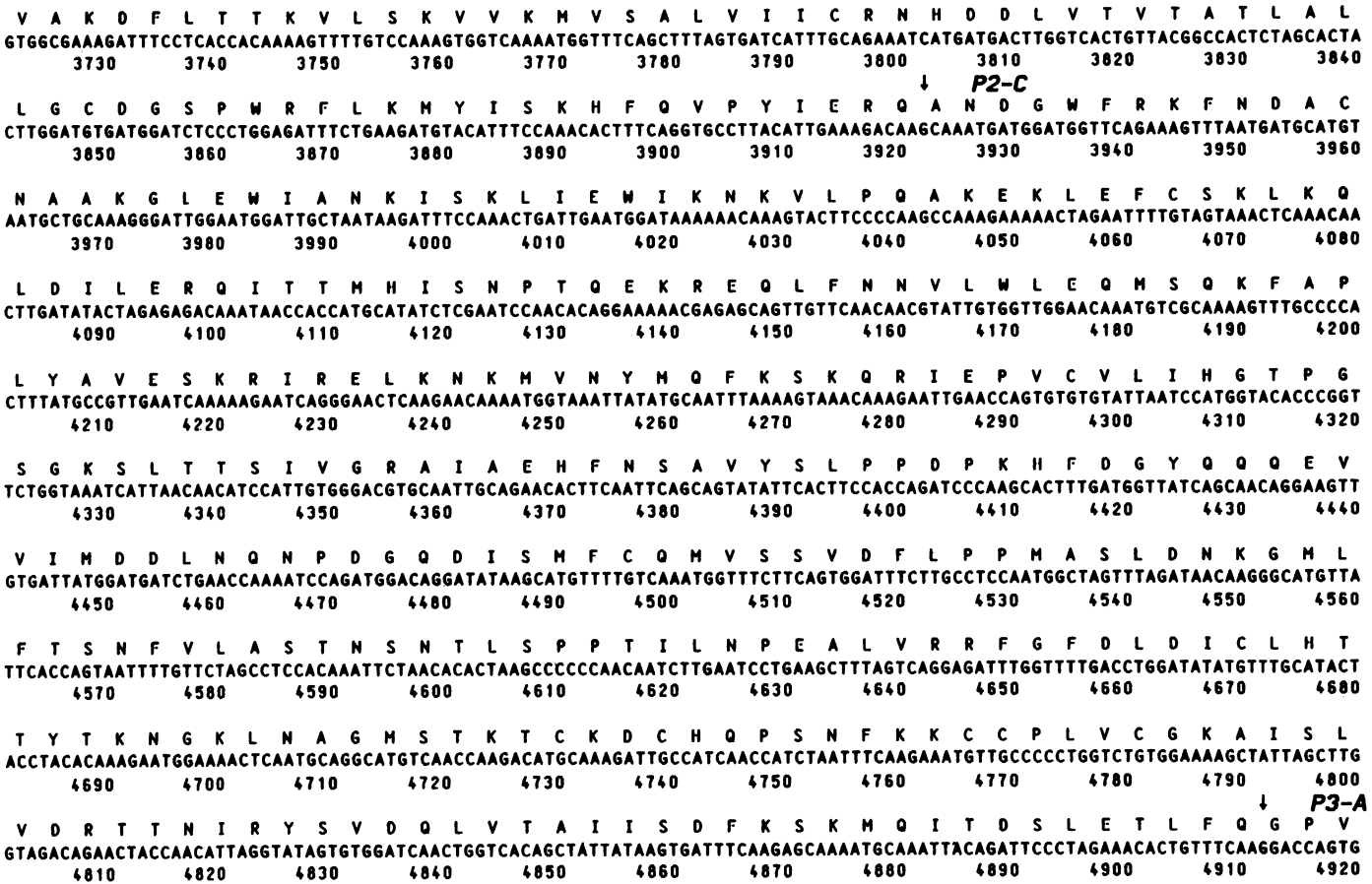
The presented nucleotide sequence allows detailed comparisons with representatives of the other genera of picornaviridae. There is no detectable nucleotide sequence homology with encephalomyocarditis virus (EMCV)(16) or foot-and-mouth-disease virus (FMDV)(15) but the genome is highly homologous to all 3 serotypes of poliovirus, members of the enterovirus genus (10-14). The similarity to these viruses implies that the genome organization is identical and that the corresponding viral proteins are closely related. A schematic representation of the HRV-14 genome, based on this comparison is shown in Fig.3. In the following discussion, the comparisons refer to poliovirus type 3 (13,14), although similar conclusions could be drawn from comparisons with either of the other two serotypes.

The 5′ terminal region of the genome is of unknown function but by analogy with the polioviruses this region is presumed to be non-coding (10). This is the first nucleotide sequence to be presented for the 5′ region of a


TTAAAACAGCGGATGGGTATCCCACCATTCGACCCATTGGGTGTAGTACTCTGGTACTATGTACCTTTGTACGCCTGTTTCTCCCCAACCACCCTTCCTTAAAATTCCCACCCATGAAAC
10 20 30 40 50 60 70 80 90 100 110 120
GTTAGAAGCTTGACATTAAAGTACAATAGGTGGCGCCATATCCAATGGTGTCTATGTACAAGCACTTCTGTTTCCCCGGAGCGAGGTATAGGCTGTACCCACTGCCAAAAGCCTTTAACC
130 140 150 160 170 180 190 200 210 220 230 240
GTTATCCGCCAACCAACTACGTAACAGTTAGTACCATCTTGTTCTTGACTGGACGTTCGATCAGGTGGATTTTCCCTCCACTAGTTTGGTCGATGAGGCTAGGAATTCCCCACGGGTGAC
250 260 270 280 290 300 310 320 330 340 350 360
CGTGTCCTAGCCTCGTGGCGGCCAACAGCTTATGCTGGGACGCCCTTTTAAGGACATGGTGTGAAGACTCGCATGTGCTTGGTTGTGAGTCTCCGGCCCCTGAATGCGGCTAACCTTAAC
370 380 390 400 410 420 430 440 450 460 470 480
CCTGGAGCCTTATGCCACGATCCAGTGGTTGTAAGGTCGTAATGAGCAATTCCGGGACGGGACCGACTACTTTGGGTGTCCGTGTTTCTCATTTTTCTTCATATTGTCTTATGGTCACAG
490 500 510 520 530 540 550 560 570 580 590 600
→ VP4
M G A Q V S T Q K S G S H E N Q N I L T N G S N Q T F T V I N Y
CATATATATACATATACTGTGATCATGGGCGCTCAGGTTTCTACACAGAAAAGTGGATCTCACGAAAATCAAAACATTTTGACCAATGGATCAAATCAGACTTTCACAGTTATAAATTAC
610 620 630 640 650 660 670 680 690 700 710 720
↓ VP2
Y K D A A S T S S A G Q S L S M D P S K F T E P V K D L M L K G A P A L N S P N
TATAAGGATGCAGCAAGTACATCATCAGCTGGTCAATCACTGTCAATGGACCCATCTAAGTTTACAGAACCAGTTAAAGATCTCATGCTTAAGGGTGCACCAGCATTGAATTCACCCAAT
730 740 750 760 770 780 790 800 810 820 830 840
V E A C G Y S D R V Q Q I T L G N S T I T T Q E A A N A V V C Y A E W P E Y L P
GTTGAGGCCTGTGGTTATAGTGATAGAGTACAACAAATCACACTCGGGAATTCAACAATAACAACACAAGAAGCAGCCAACGCTGTTGTGTGTTATGCTGAATGGCCAGAGTACCTTCCA
850 860 870 880 890 900 910 920 930 940 950 960
D V D A S D V N K T S K P D T S V C R F Y T L D S K T W T T G S K G W C W K L P
GATGTGGACGCTAGTGATGTCAATAAAACTTCAAAACCAGACACTTCTGTCTGTAGGTTTTACACATTGGATAGTAAGACATGGACAACAGGTTCTAAAGGCTGGTGCTGGAAATTACCA
970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
D A L K D M G V F G Q N M F F H S L G R S G Y T V H V Q C N A T K F H S G C L L
GATGCACTCAAAGATATGGGTGTGTTCGGGCAAAACATGTTTTTCCACTCACTAGGAAGATCAGGTTACACAGTACACGTTCAGTGCAATGCCACAAAATTCCATAGCGGTTGTCTACTT
1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
V V V I P E H Q L A S H E G G N V S V K Y T F T H P G E R G I D L S S A N E V G
GTAGTTGTAATACCAGAACACCAACTGGCTTCACATGAGGGTGGCAATGTTTCAGTTAAATACACATTCACGCATCCAGGTGAACGTGGTATAGATTTATCATCTGCAAATGAAGTGGGA
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320

```
 G  P  V  K  D  V  I  Y  N  M  N  G  T  L  L  G  N  L  L  I  F  P  H  Q  F  I  N  L  R  T  N  N  T  A  T  I  V  I  P  Y
GGGCCTGTCAAGGATGTCATATACAATATGAATGGTACTTTATTAGGAAATCTGCTCATTTTCCCTCACCAGTTCATTAATCTAAGAACCAATAATACAGCCACAATAGTGATACCATAC
      1330       1340       1350       1360       1370       1380       1390       1400       1410       1420       1430       1440

 I  N  S  V  P  I  D  S  M  T  R  H  N  N  V  S  L  M  V  I  P  I  A  P  L  T  V  P  T  G  A  T  P  S  L  P  I  T  V  T
ATAAACTCAGTACCCATTGATTCAATGACACGTCACAACAATGTCTCACTGATGGTCATCCCTATTGCCCCTCTTACAGTACCAACTGGAGCAACTCCCTCACTCCCTATAACAGTCACA
      1450       1460       1470       1480       1490       1500       1510       1520       1530       1540       1550       156
                                                                ↓ VP3
 I  A  P  M  C  T  E  F  S  G  I  R  S  K  S  I  V  P  Q  G  L  P  T  T  T  L  P  G  S  G  Q  F  L  T  T  D  D  R  Q  S
ATAGCACCTATGTGCACTGAGTTCTCTGGGATAAGGTCCAAGTCAATTGTGCCACAAGGTTTGCCAACTACAACTTTGCCGGGGTCAGGACAATTCTTGACCACAGATGACAGGCAATCC
      1570       1580       1590       1600       1610       1620       1630       1640       1650       1660       1670       1680

 P  S  A  L  P  N  Y  E  P  T  P  R  I  H  I  P  G  K  V  H  N  L  L  E  I  I  Q  V  D  T  L  I  P  M  N  N  T  H  T  K
CCCAGTGCACTGCCAAATTATGAGCCAACTCCAAGAATACACATACCAGGGAAAGTTCATAACTTGCTAGAAATTATACAGGTAGATACACTCATTCCTATGAACAACACGCATACAAAA
      1690       1700       1710       1720       1730       1740       1750       1760       1770       1780       1790       1800

 D  E  V  N  S  Y  L  I  P  L  N  A  N  R  Q  N  E  Q  V  F  G  T  N  L  F  I  G  D  G  V  F  K  T  T  L  L  G  E  I  V
GATGAGGTTAACAGTTACCTCATACCACTAAATGCAAACAGGCAAAATGAGCAGGTTTTTGGGACAAACCTGTTTATTGGTGATGGGGTCTTCAAAACTACTCTTCTGGGTGAAATTGTT
      1810       1820       1830       1840       1850       1860       1870       1880       1890       1900       1910       1920

 Q  Y  Y  T  H  W  S  G  S  L  R  F  S  L  M  Y  T  G  P  A  L  S  S  A  K  L  I  L  A  Y  T  P  P  G  A  R  G  P  Q  D
CAGTACTATACACATTGGTCTGGATCACTTAGATTCTCTTTGATGTATACTGGTCCTGCCTTGTCCAGTGCTAAACTCATTCTAGCATACACCCCGCCTGGTGCTCGTGGTCCACAGGAC
      1930       1940       1950       1960       1970       1980       1990       2000       2010       2020       2030       2040

 R  R  E  A  M  L  G  T  H  V  V  W  D  I  G  L  Q  S  T  I  V  M  T  I  P  W  T  S  G  V  Q  F  R  Y  T  D  P  D  T  Y
AGGAGAGAAGCAATGCTAGGTACTCATGTTGTCTGGGATATTGGTCTGCAATCCACCATAGTAATGACAATACCATGGACATCAGGGGTGCAGTTTAGATATACTGATCCAGATACATAC
      2050       2060       2070       2080       2090       2100       2110       2120       2130       2140       2150       2160

 T  S  A  G  F  L  S  C  W  Y  Q  T  S  L  I  L  P  P  E  T  T  G  Q  V  Y  L  L  S  F  I  S  A  C  P  D  F  K  L  R  L
ACCAGTGCTGGCTTTCTATCATGTTGGTATCAAACTTCTCTTATACTTCCCCCAGAAACGACCGGCCAGGTCTACTTATTATCATTCATAAGTGCATGTCCAGATTTTAAGCTTAGGCTG
      2170       2180       2190       2200       2210       2220       2230       2240       2250       2260       2270       2280
           ↓ ?        ↓ ?                   ↓ ?   VP1
 M  K  D  T  Q  T  I  S  Q  T  V  A  L  T  E  G  L  G  D  E  L  E  E  V  I  V  E  K  T  K  Q  T  V  A  S  I  S  S  G  P
ATGAAAGATACTCAAACTATCTCACAGACTGTTGCACTCACTGAAGGCTTAGGTGATGAATTAGAAGAAGTCATCGTTGAGAAAACGAAACAGACGGTGGCCTCAATCTCATCTGGTCCA
      2290       2300       2310       2320       2330       2340       2350       2360       2370       2380       2390       2400

 K  H  T  Q  K  V  P  I  L  T  A  N  E  T  G  A  T  M  P  V  L  P  S  D  S  I  E  T  R  T  T  Y  M  H  F  N  G  S  E  T
AAACACACACAAAAAGTCCCCATACTAACTGCAAACGAAACAGGGGCCACAATGCCTGTTCTTCCATCAGACAGCATAGAAACCAGAACTACCTACATGCACTTTAATGGTTCAGAAACT
      2410       2420       2430       2440       2450       2460       2470       2480       2490       2500       2510       2520
```


D V E C F L G R A A C V H V T E I Q N K D A T G I D N H R E A K L F N D W K I N
GATGTAGAATGCTTTTTGGGTCGTGCAGCTTGTGTGCATGTAACTGAAATACAAAACAAAGATGCTACTGGAATAGATAATCACAGAGAAGCAAAATTGTTCAATGATTGGAAAATCAAC
2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640
L S S L V Q L R K K L E L F T Y V R F D S E Y T I L A T A S Q P D S A N Y S S N
CTGTCCAGCCTTGTCCAACTTAGAAAGAAACTAGAACTCTTCACTTATGTTAGGTTTGATTCTGAGTATACCATACTGGCCACTGCATCTCAACCTGATTCAGCAAACTATTCAAGCAAT
2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760
L V V Q A M Y V P P G A P N P K E W D D Y T W Q S A S N P S V F F K V G D T S R
TTGGTGGTCCAAGCCATGTATGTTCCACCTGGTGCCCCGAATCCAAAAGAGTGGGACGATTACACATGGCAAAGTGCTTCAAACCCCAGTGTATTCTTCAAGGTGGGGGATACATCCAGG
2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880
F S V P Y V G L A S A Y N C F Y D G Y S H D D A E T Q Y G I T V L N H M G S M A
TTTAGTGTGCCTTATGTAGGATTGGCATCAGCATATAATTGTTTTTATGATGGTTACTCACATGATGATGCAGAAACTCAGTATGGCATAACTGTTCTAAACCATATGGGTAGTATGGCA
2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
F R I V N E H D E H K T L V K I R V Y H R A K H V E A W I P R A P R A L P Y T S
TTCAGAATAGTAAATGAACATGATGAACATAAAACTCTTGTCAAGATCAGAGTTTATCACAGGGCAAAGCACGTTGAAGCATGGATTCCAAGAGCACCCAGAGCACTACCCTACACATCA
3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120
↓ ? ↓ ? P2-A
I G R T N Y P K N T E P V I K K R K G D I K S Y G L G P R Y G G I Y T S N V K I
ATAGGGCGCACAAATTATCCTAAGAATACAGAACCAGTAATTAAGAAGAGGAAAGGTGACATTAAATCCTATGGTTTAGGACCTAGGTACGGTGGGATTTATACATCAAATGTTAAAATA
3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240
M N Y H L M T P E D H H N L I A P Y P N R D L A I V S T G G H G A E T I P H C N
ATGAATTACCACTTGATGACACCAGAAGACCACCATAATCTGATAGCACCCTATCCAAATAGAGATTTAGCAATAGTCTCAACAGGAGGACATGGTGCAGAAACAATACCACACTGTAAC
3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360
C T S G V Y Y S T Y Y R K Y Y P I I C E K P T N I W I E G N P Y Y P S R F Q A G
TGTACATCAGGTGTTTACTATTCCACATATTACAGAAAGTATTACCCCATAATTTGTGAAAAGCCCACCAACATCTGGATTGAAGGAAACCCTTATTACCCAAGTAGATTTCAAGCAGGA
3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480
V M K G V G P A E P G D C G G I L R C I H G P I G L L T A G G S G Y V C F A D I
GTGATGAAAGGGGTTGGGCCAGCAGAACCAGGAGACTGCGGTGGGATTTTGAGATGCATACATGGTCCCATTGGATTGTTAACAGCTGGAGGTAGTGGATATGTTTGTTTTGCTGACATA
3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
↓ P2-B
R Q L E C I A E E Q G L S D Y I T G L G R A F G V G F T D Q I S T K V T E L Q E
CGACAGTTGGAGTGTATCGCAGAGGAACAGGGGCTGAGTGATTACATCACAGGTTTGGGTAGAGCTTTTGGTGTCGGGTTCACTGACCAAATCTCAACAAAAGTCACAGAACTACAAGAA
3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720


V A K D F L T T K V L S K V V K M V S A L V I I C R N H D D L V T V T A T L A L
GTGGCGAAAGATTTCCTCACCACAAAAGTTTTGTCCAAAGTGGTCAAAATGGTTTCAGCTTTAGTGATCATTTGCAGAAATCATGATGACTTGGTCACTGTTACGGCCACTCTAGCACTA
3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830 3840
P2-C
L G C D G S P W R F L K M Y I S K H F Q V P Y I E R Q A N D G W F R K F N D A C
CTTGGATGTGATGGATCTCCCTGGAGATTTCTGAAGATGTACATTTCCAAACACTTTCAGGTGCCTTACATTGAAAGACAAGCAAATGATGGATGGTTCAGAAAGTTTAATGATGCATGT
3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960
N A A K G L E W I A N K I S K L I E W I K N K V L P Q A K E K L E F C S K L K Q
AATGCTGCAAAGGGATTGGAATGGATTGCTAATAAGATTTCCAAACTGATTGAATGGATAAAAACAAAGTACTTCCCCAAGCCAAAGAAAAACTAGAATTTTGTAGTAAACTCAAACAA
3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080
L D I L E R Q I T T M H I S N P T Q E K R E Q L F N N V L W L E Q M S Q K F A P
CTTGATATACTAGAGAGACAAATAACCACCATGCATATCTCGAATCCAACACAGGAAAAACGAGAGCAGTTGTTCAACAACGTATTGTGGTTGGAACAAATGTCGCAAAAGTTTGCCCCA
4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
L Y A V E S K R I R E L K N K M V N Y M Q F K S K Q R I E P V C V L I H G T P G
CTTTATGCCGTTGAATCAAAAAGAATCAGGGAACTCAAGAACAAAATGGTAAATTATATGCAATTTAAAAGTAAACAAAGAATTGAACCAGTGTGTGTATTAATCCATGGTACACCCGGT
4210 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320
S G K S L T T S I V G R A I A E H F N S A V Y S L P P D P K H F D G Y Q Q Q E V
TCTGGTAAATCATTAACAACATCCATTGTGGGACGTGCAATTGCAGAACACTTCAATTCAGCAGTATATTCACTTCCACCAGATCCCAAGCACTTTGATGGTTATCAGCAACAGGAAGTT
4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440
V I M D D L N Q N P D G Q D I S M F C Q M V S S V D F L P P M A S L D N K G M L
GTGATTATGGATGATCTGAACCAAAATCCAGATGGACAGGATATAAGCATGTTTTGTCAAATGGTTTCTTCAGTGGATTTCTTGCCTCCAATGGCTAGTTTAGATAACAAGGGCATGTTA
4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560
F T S N F V L A S T N S N T L S P P T I L N P E A L V R R F G F D L D I C L H T
TTCACCAGTAATTTTGTTCTAGCCTCCACAAATTCTAACACACTAAGCCCCCCAACAATCTTGAATCCTGAAGCTTTAGTCAGGAGATTTGGTTTTGACCTGGATATATGTTTGCATACT
4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680
T Y T K N G K L N A G M S T K T C K D C H Q P S N F K K C C P L V C G K A I S L
ACCTACACAAAGAATGGAAAACTCAATGCAGGCATGTCAACCAAGACATGCAAAGATTGCCATCAACCATCTAATTTCAAGAAATGTTGCCCCCTGGTCTGTGGAAAAGCTATTAGCTTG
4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
P3-A
V D R T T N I R Y S V D Q L V T A I I S D F K S K M Q I T D S L E T L F Q G P V
GTAGACAGAACTACCAACATTAGGTATAGTGTGGATCAACTGGTCACAGCTATTATAAGTGATTTCAAGAGCAAAATGCAAATTACAGATTCCCTAGAAACACTGTTTCAAGGACCAGTG
4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920

```
 Y  K  D  L  E  I  D  V  C  N  T  P  P  P  E  C  I  N  D  L  L  K  S  V  D  S  E  E  I  R  E  Y  C  K  K  K  K  W  I  I
TATAAAGATTTAGAGATTGATGTTTGCAACACACCACCTCCAGAATGTATCAACGATTTACTGAAATCTGTAGATTCAGAAGAGATTAGGGAATATTGTAAGAAGAAGAAATGGATTATA
     4930      4940      4950      4960      4970      4980      4990      5000      5010      5020      5030      5040

 P  E  I  P  T  N  I  E  R  A  M  N  Q  A  S  M  I  I  N  T  I  L  M  F  V  S  T  L  G  I  V  Y  V  I  Y  K  L  F  A  Q
CCTGAAATTCCTACCAACATAGAAAGAGCTATGAATCAAGCCAGCATGATTATTAATACTATTCTGATGTTTGTCAGTACATTAGGTATTGTTTATGTCATTTATAAATTGTTTGCTCAA
     5050      5060      5070      5080      5090      5100      5110      5120      5130      5140      5150      5160
     ↓ VPG                                                                                   ↓ PROTEASE
 T  Q  G  P  Y  S  G  N  P  P  H  N  K  L  K  A  P  T  L  R  P  V  V  V  Q  G  P  N  T  E  F  A  L  S  L  L  R  K  N  I
ACTCAAGGACCATATTCTGGTAACCCGCCTCACAATAAACTAAAAGCCCCAACTTTACGCCCAGTTGTTGTGCAAGGACCAAACACAGAATTTGCACTATCCCTGTTAAGGAAAAACATA
     5170      5180      5190      5200      5210      5220      5230      5240      5250      5260      5270      5280

 M  T  I  T  T  S  K  G  E  F  T  G  L  G  I  H  D  R  V  C  V  I  P  T  H  A  Q  P  G  D  D  V  L  V  N  G  Q  K  I  R
ATGACTATAACAACCTCAAAGGGAGAGTTCACAGGGTTAGGCATACATGATCGTGTCTGTGTGATACCCACACACGCACAGCCTGGTGATGATGTACTAGTGAATGGTCAGAAAATTAGA
     5290      5300      5310      5320      5330      5340      5350      5360      5370      5380      5390      5400

 V  K  D  K  Y  K  L  V  D  P  E  N  I  N  L  E  L  T  V  L  T  L  D  R  N  E  K  F  R  D  I  R  G  F  I  S  E  D  L  E
GTTAAGGATAAGTACAAATTAGTAGATCCAGAGAACATTAATCTAGAGCTTACAGTGTTGACTTTAGATAGAAATGAAAAATTCAGAGATATCAGGGGATTTATATCAGAAGATCTAGAA
     5410      5420      5430      5440      5450      5460      5470      5480      5490      5500      5510      5520

 G  V  D  A  T  L  V  V  H  S  N  N  F  T  N  T  I  L  E  V  G  P  V  T  M  A  G  L  I  N  L  S  S  T  P  T  N  R  M  I
GGTGTGGATGCCACTTTGGTAGTACATTCAAATAACTTTACCAACACTATCTTAGAAGTTGGCCCTGTAACAATGGCAGGACTTATTAATTTGAGTAGCACCCCCACTAACAGAATGATT
     5530      5540      5550      5560      5570      5580      5590      5600      5610      5620      5630      5640

 R  Y  D  Y  A  T  K  T  G  Q  C  G  G  V  L  C  A  T  G  K  I  F  G  I  H  V  G  G  N  G  R  Q  G  F  S  A  Q  L  K  K
CGTTATGATTATGCAACAAAAACTGGGCAGTGTGGAGGTGTGCTGTGTGCTACTGGTAAGATCTTTGGTATTCATGTTGGCGGTAATGGAAGACAAGGATTTTCAGCTCAACTTAAAAAA
     5650      5660      5670      5680      5690      5700      5710      5720      5730      5740      5750      5760
               ↓ POLYMERASE
 Q  Y  F  V  E  K  Q  G  Q  V  I  A  R  H  K  V  R  E  F  N  I  N  P  V  N  T  P  T  K  S  K  L  H  P  S  V  F  Y  D  V
CAATATTTTGTAGAGAAACAAGGCCAAGTAATAGCTAGACATAAGGTTAGGGAGTTTAACATAAATCCAGTCAACACGCCAACCAAGTCAAAATTACATCCCAGTGTATTCTATGATGTT
     5770      5780      5790      5800      5810      5820      5830      5840      5850      5860      5870      5880

 F  P  G  D  K  E  P  A  V  L  S  D  N  D  P  R  L  E  V  K  L  T  E  S  L  F  S  K  Y  K  G  N  V  N  T  E  P  T  E  N
TTCCCAGGTGACAAGGAACCTGCTGTATTGAGTGACAATGATCCCAGACTGGAAGTTAAATTGACTGAATCATTATTCTCTAAGTACAAGGGGAATGTAAATACGGAACCCACTGAAAAT
     5890      5900      5910      5920      5930      5940      5950      5960      5970      5980      5990      6000

 M  L  V  A  V  D  H  Y  A  G  Q  L  L  S  L  D  I  P  T  S  E  L  T  L  K  E  A  L  Y  G  V  D  G  L  E  P  I  D  I  T
ATGCTTGTGGCTGTAGACCATTATGCAGGGCAACTATTATCACTAGATATCCCCACTTCTGAACTTACACTAAAAGAAGCATTATATGGAGTAGATGGACTAGAACCTATAGATATTACA
     6010      6020      6030      6040      6050      6060      6070      6080      6090      6100      6110      6120
```

```
 T  S  A  G  F  P  Y  V  S  L  G  I  K  K  R  D  I  L  N  K  E  T  Q  D  T  E  K  M  K  F  Y  L  D  K  Y  G  I  D  L  P
ACCAGTGCAGGATTTCCCTATGTGAGTCTTGGGATCAAAAAGAGAGACATTCTGAATAAAGAGACCCAGGACACAGAAAAGATGAAGTTTTATCTAGACAAGTATGGCATTGACTTGCCT
      6130       6140       6150       6160       6170       6180       6190       6200       6210       6220       6230       6240

 L  V  T  Y  I  K  D  E  L  R  S  V  D  K  V  R  L  G  K  S  R  L  I  E  A  S  S  L  N  D  S  V  N  M  R  M  K  L  G  N
CTAGTTACATATATTAAGGATGAATTAAGAAGTGTTGACAAAGTCCGATTAGGGAAAAGTAGATTAATTGAAGCCTCCAGTTTGAATGATTCTGTTAACATGAGAATGAAACTAGGCAAC
      6250       6260       6270       6280       6290       6300       6310       6320       6330       6340       6350       6360

 L  Y  K  A  F  H  Q  N  P  G  V  L  T  G  S  A  V  G  C  D  P  D  V  F  W  S  V  I  P  C  L  M  D  G  H  L  M  A  F  D
CTTTACAAAGCATTCCATCAAAATCCCGGTGTTCTGACTGGATCAGCAGTGGGTTGTGATCCTGATGTGTTTTGGTCTGTCATCCCTTGTTTAATGGATGGGCACCTGATGGCATTTGAT
      6370       6380       6390       6400       6410       6420       6430       6440       6450       6460       6470       6480

 Y  S  N  F  D  A  S  L  S  P  V  W  F  V  C  L  E  K  V  L  T  K  L  G  F  A  G  S  S  L  I  Q  S  I  C  N  T  H  H  I
TACTCTAATTTTGATGCCTCTTTGTCACCAGTTTGGTTTGTCTGTCTAGAGAAGGTTTTGACCAAGTTAGGCTTTGCAGGCTCTTCATTAATTCAATCAATTTGTAATACCCATCATATC
      6490       6500       6510       6520       6530       6540       6550       6560       6570       6580       6590       6600

 F  R  D  E  I  Y  V  V  E  G  G  M  P  S  G  C  S  G  T  S  I  F  N  S  M  I  N  N  I  I  I  R  T  L  I  L  D  A  Y  K
TTTAGGGATGAAATATATGTGGTTGAAGGTGGCATGCCCTCAGGGTGTTCAGGAACCAGCATATTCAATTCCATGATCAACAACATAATCATTAGGACTTTGATATTAGATGCATATAAA
      6610       6620       6630       6640       6650       6660       6670       6680       6690       6700       6710       6720

 G  I  D  L  D  K  L  K  I  L  A  Y  G  D  D  L  I  V  S  Y  P  Y  E  L  D  P  Q  V  L  A  T  L  G  K  N  Y  G  L  T  I
GGAATAGATTTAGACAAACTTAAAATCTTAGCTTACGGTGATGATTTGATTGTTTCTTATCCTTATGAACTGGATCCACAAGTGTTGGCAACTCTTGGTAAAAATTATGGACTAACCATC
      6730       6740       6750       6760       6770       6780       6790       6800       6810       6820       6830       6840

 T  P  P  D  K  S  E  T  F  T  K  M  T  W  E  N  L  T  F  L  K  R  Y  F  K  P  D  Q  Q  F  P  F  L  V  H  P  V  M  P  M
ACACCCCCAGACAAATCTGAAACTTTTACAAAAATGACATGGGAAAACTTGACATTTTTAAAGAGATACTTCAAGCCTGATCAACAATTTCCCTTTTTGGTTCACCCAGTTATGCCCATG
      6850       6860       6870       6880       6890       6900       6910       6920       6930       6940       6950       6960

 K  D  I  H  E  S  I  R  W  T  K  D  P  K  N  T  Q  D  H  V  R  S  L  C  M  L  A  W  H  S  G  E  K  E  Y  N  E  F  I  Q
AAAGATATACATGAGTCAATCAGATGGACAAAGGATCCTAAAAACACACAGGATCACGTCCGATCATTATGCATGTTAGCATGGCACTCAGGAGAAAAAGAGTACAATGAATTCATTCAG
      6970       6980       6990       7000       7010       7020       7030       7040       7050       7060       7070       7080

 K  I  R  T  T  D  I  G  K  C  L  I  L  P  E  Y  S  V  L  R  R  R  W  L  D  L  F  *
AAGATCAGAACTACTGACATTGGAAAATGTCTAATTCTCCCAGAATACAGCGTACTTAGGAGGCGCTGGTTGGACCTCTTTTAGGTTAACAATATAGACACTTAATTTGAGTAGAAGTAG
      7090       7100       7110       7120       7130       7140       7150       7160       7170       7180       7190       7200

GAGTTTAT  - poly A
 7205
```

picornavirus other than the polioviruses and thus provides important information about the sequence conservation of this region which may give an indication of function. The homology between HRV-14 and poliovirus type 3 in the non-coding region is remarkable. The nucleotide sequences can be aligned, by taking into account several small deletions or insertions, to give 63% homology. This includes completely conserved stretches of 20, 23 and 27 nucleotides starting at positions 61, 452 and 542 respectively in HRV-14. These matches are all in regions which are highly homologous between polioviruses, the latter two being identical in the three serotypes (10-14). No matches greater than 17 nucleotides have been observed in the coding regions of HRV-14 and poliovirus type 3. The level of nucleotide homology of the 5′ presumed non-coding region is similar to that of the region of the genome coding for the polymerase protein, the most highly conserved protein at the amino acid level. Here the nucleotide sequence homology is 60%. These results imply that there is a strong pressure to conserve the nucleotide sequence of the non-coding region. The reason for this is not clear since the function of the region is unknown. One possible explanation is that there is a hitherto unsuspected coding capacity. In HRV-14 this region does contain several open reading frames, including one of 77 codons (starting position 492) and one of 64 codons (starting position 165). In neither case is there an equivalent open reading frame in poliovirus and these are therefore of questionable significance. One potential reading frame (starting at position 433), coding for a peptide of 29 amino acids, is conserved between HRV-14 and poliovirus type 3 and this contains the stretch of 23 identical nucleotides (13,14). However, the initiating methionine codon is absent in the other two poliovirus serotypes and therefore this observation is again difficult to interpret (12). The fact that insertions or deletions have to be inserted to align the 5′ sequences of poliovirus type 3 and HRV-14 suggests that the region is under pressure to be conserved for reasons other than protein coding. In this connection it is interesting to note that the nucleotide composition of the region (A=22.4%, C=26.0%, G=22.8%, T=28.8%) is different from that of the rest of the genome (A=33.0%, C=19.6%, G=20.2%, T=27.2%). The higher G+C content possibly suggests that secondary structure in the 5′ untranslated region plays some part in the

Figure 2. The complete nucleotide sequence of the cDNA representing the genome of HRV-14 together with the predicted amino acid sequence of polypeptides encoded by the major open reading frame. Vertical arrows indicate the positions of probable polyprotein cleavage sites based on homology with poliovirus. Proteins are named in accordance with the L434 nomenclature proposed by Rueckert and Wimmer (23).

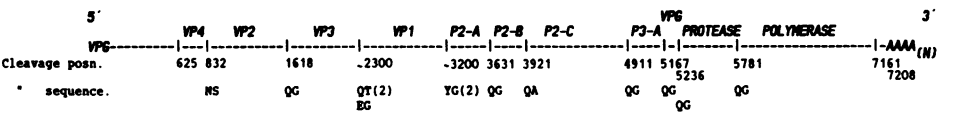


Figure 3. Schematic representation of the HRV-14 genome showing the position of the virus proteins and the predicted cleavage sites.

replicative cycle of the virus.

The 5′ regions of HRV-14 and poliovirus type 3 align for approximately 600 nucleotides. Around this point however, there is a divergence of sequence and at position 625 in HRV-14 there is a methionine codon at the start of the long open reading frame. In poliovirus type 3 the non-coding region extends for a further 140 nucleotides. This relative deletion in HRV-14 is a notable feature in view of the otherwise close homology and it was therefore thought necessary to ensure that this was not an artefact produced during cDNA cloning. Restriction enzyme mapping of pAM 6-9 over this region confirmed that these clones are identical in this respect to pAM 2, from which the sequence was derived, and therefore there is a radical difference between HRV-14 and poliovirus type 3 at this point.

A methionine codon located at position 625 initiates a long open reading frame of 2178 codons and this is sufficient to code for all the known virus proteins (27). The most striking feature of the coding region is again the close homology with the polioviruses. The overall amino acid homology with poliovirus type 3 varies from 43% in VP1 and VP3 to 65% in the polymerase protein suggesting that the virus proteins have closely related functions. The homologies are slightly less than those between the polioviruses and another enterovirus, enterovirus 70 (Ryan *et al* in preparation), but much greater than those between HRV-14 and EMCV (16) or FMDV (15)(see Table 1). Previous studies based on RNA hybridization have suggested that there is little homology between poliovirus and human rhinoviruses, including HRV-14, at the nucleotide sequence level (28) and therefore the extensive amino acid homology found here is unexpected. However, it has been shown that HRV-14 and some of the other rhinoviruses share a common cellular receptor with the enterovirus, coxsackie A21 (29). These viruses may therefore be related to some extent, at least in the region of the cellular receptor binding site.

The homology with the polioviruses facilitates the prediction of the cleavage sites in the polyprotein recognized by the virus protease to generate the viral proteins and these are summarized in Table 2. As in the

Table 1. Homologies between the proteins of HRV-14 and the equivalent proteins of other picornaviruses.

HRV-14 protein	% Homology with: poliovirus type 3 (enterovirus)	EMCV (cardiovirus)	FMDV (aphthovirus)
VP4	58	n.h.d.	n.h.d.
VP2	56	"	"
VP3	43	"	"
VP1	43	"	"
P2-A	49	"	"
P2-B	49	"	-
P2-C	61	27	22
P3-A	52	19	n.h.d.
VPg	48	30	35(VPg[1])
Protease	46	21	14
Polymerase	65	33	33

n.h.d. =no homology detected.

case of polioviruses, many of these seem to occur between glutamine and glycine residues (10). This sequence is not present, however, at the VP3/VP1 and P2-B/P2-C boundaries. In the case of the former, the sequence glutamine-threonine occurs twice and glutamic acid-glycine once in the area likely to be the cleavage junction, but it is not known which if any of these is recognized by the HRV-14 protease. In FMDV (15), the VP3/VP1 cleavage is at glutamine-threonine and this possibly suggests that one of these two potential cleavage sites is used in HRV-14. The most likely cleavage site in the predicted region of the P2-B/P2-C boundary would seem to be glutamine-alanine. These results imply that the rhinovirus encoded protease is less stringent in its substrate specificity than that of the

Table 2. Predicted cleavage sites of the HRV-14 polyprotein.

Protein boundary	Probable cleavage
VP4/VP2	ALN|SPN
VP2/VP3	VPQ|GLP
VP3/VP1	DTQ|TIS?
	ISQ|TVA?
	LTE|GLG?
VP1/P2-A	KSY|GLG?
	PRY|GGI?
P2-A/P2-B	EEQ|GLS
P2-B/P2-C	ERQ|AND
P2-C/P3-A	LFQ|GPV
P3-A/VPg	QTQ|GPY
VPg/Protease	VVQ|GPN
Protease/Polymerase	EKQ|GQV

polioviruses. In general, the regions around the cleavage sites show the greatest divergence between HRV-14 and poliovirus type 3 with the presence of several small deletions or insertions and this also may be a reflection of the differing substrate specificities. It is noteworthy that there is only 46% homology at the amino acid level between HRV-14 and poliovirus type 3 in the region of the protease. In contrast, at the VP4/VP2 boundary, whose processing is unique in that it occurs during the maturation of the virion (30), 18 amino acids around the cleavage junction are exactly conserved between HRV-14 and poliovirus type 3 (13,14). The distinctive nature of the cleavage site (asparagine-serine) and the time-scale of the processing suggest that this is performed by a second virus protease or one specified by the host. As with the polioviruses, the VP1/P2-A cleavage probably occurs at tyrosine-glycine. This sequence appears twice in the appropriate region of HRV-14 and comparison with enterovirus 70 (Ryan *et al*, in preparation) suggests that the most C-terminal is used.

In poliovirus type 3, a major antigenic site for virus neutralization has been located at amino acid position 92-99 of VP1 (31). By aligning homologous amino acids flanking this sequence, a corresponding region can be identified in HRV-14 (nucleotide position 2557-2613) which has a completely different primary structure. Hydrophilicity profiles of the capsid proteins of HRV-14 and poliovirus type 3 are remarkably similar and show this region to be hydrophilic in both cases (32, data not shown). These observations suggest that this region may be an antigenic site in HRV-14 against which neutralizing antibodies are directed. There is evidence for the involvement of at least two other antigenic regions of the capsid in neutralization of polioviruses 1 and 3 (33, P.D. Minor, unpublished). These are located at position 55-70 in VP3 and 284-291 in VP1 of poliovirus. In HRV-14 the corresponding regions are again hydrophilic and have a primary structure distinct from those of the polioviruses. It will be interesting to test whether synthetic peptides corresponding to these regions induce antibodies which neutralize HRV-14.

Comparison with human rhinovirus 2

On the basis of neutralization by reference antisera, at least 115 serotypes of human rhinoviruses are believed to exist (3). Some of these serotypes can be grouped on the basis of low-level, one-way or reciprocal cross-reactivity and there is some evidence for the existance of intertypes which are related to two serotypes (34,35). However, little is known about the overall degree of homology between serotypes at the nucleotide or amino

acid sequence level although one hybridization study has indicated that HRV-1A, HRV-2 and HRV-14 share no more homology with one another than each does with poliovirus type 2 (28). Recently the sequence of 1425 nucleotides from the 3′ terminus of the genome of HRV-2 has been determined (36). This comprises the polymerase gene and the 3′ non-coding region. Comparison with the sequence of HRV-14 presented here provides a preliminary estimate of diversity within the rhinovirus genus. In contrast to the different poliovirus serotypes which are 97% homologous to each other in the polymerase gene (17), HRV-2 and HRV-14 are only 55% homologous at the amino acid level. This is to be expected from the previous report of the low level of nucleotide sequence homology among the rhinoviruses (28). More surprising however, is the fact that HRV-14 is more homologous, at the level of predicted amino acid sequence, to poliovirus (65%) than it is to HRV-2. These results would seem to indicate that in contrast to the wide diversity between the enteroviruses, cardioviruses and aphthoviruses (15,16), there is a considerable degree of overlap between the enterovirus and rhinovirus genera in terms of nucleotide and amino acid sequence homology. This finding casts doubt on the genetic basis for separating the enteroviruses and rhinoviruses and it may be more appropriate to consider them as members of one genus of the picornaviridae family.

One of the interesting features evident on comparison of the sequences of the two rhinoviruses is the homology of the 3′ non-coding regions. At 47 nucleotides this region of HRV-14 is of a similar size to HRV-2 (42 nucleotides) and is in contrast to poliovirus (72 nucleotides). Moreover, despite the divergence in the polymerase gene, there is some homology in the non-coding region between the two rhinoviruses. There are blocks of identity of 7 (at position 7172 in HRV-14) and 5 nucleotides (at position 7196), together with some homology in the intervening region. There are no such blocks of homology with the polioviruses where this region is almost perfectly conserved between the three serotypes (10-14). The comparative features of the non-coding regions which have been described in this paper, namely the relative deletion in the HRV-14 5′ non-coding region and the homology between HRV-2 and HRV-14 in the 3′ region, raise the intriguing possibility that these regions of the rhinovirus genome are involved in defining the distinctive characteristics of the rhinoviruses. This point may be elucidated as more complete rhinovirus sequences become available.

ACKNOWLEDGEMENTS

We are indebted to Dr.G.C.Schild for helpful discussions and for his encouragement and would like to thank Jane Stanway for proof-reading the manuscript. This work was supported by the Medical Research Council of Great Britain, grant number G83242556CB.

REFERENCES

1. Stott,E.J. and Killington,R.A. (1972) Ann. Rev. Microbiol. 26, 503-524.
2. Gwaltney,J.M. (1975) Yale J. Biol. Med. 48, 17-45.
3. Melnick,J.L. (1980) Prog. Med. Virol. 26, 214-232.
4. Cooper,P.D., Agol,V.I., Bachrach,H.L., Brown,F., Ghendon,Y., Gibbs,A.J., Gillespie,J.H., Lonberg-Holm,K., Melnick,J.L., Mohanty,S.B., Povey,R.C., Rueckert,R.R., Schaffer,F.L. and Tyrrell,D.A.J. (1978) Intervirology 10, 165-180.
5. Medappa,K.C., McLean,C. and Rueckert,R.R. (1971) Virology 44, 259-270.
6. Dimmock,N.J. (1966) Nature 209, 792-794.
7. Brown,F., Newman,J.F.E. and Stott,E.J. (1970) J. Gen. Virol. 8, 145-148.
8. Nair,C.N. and Owens,M.J. (1974) J. Virol. 13, 535-537.
9. Gauntt,C.J. (1980) Intervirology 13, 7-15.
10. Kitamura,N., Semler,B.L., Rothberg,P.G., Larsen, D.R., Adler,C.J., Dorner,A.J., Emini,E.A., Hanecak,R., Lee,J.L., van der Werf,S., Anderson,C.W. and Wimmer,E. (1981) Nature 291, 547-553.
11. Racaniello,V.R. and Baltimore,D. (1981) P.N.A.S. U.S.A. 78, 4887-4891.
12. Toyoda,H., Kohara,M., Kataoka,Y., Suganuma,T., Omata,T., Imura,N. and Nomoto,A. (1984) J. Mol. Biol. 174, 561-586.
13. Stanway,G., Cann,A.J., Hauptmann,R., Hughes,P., Clarke,L.D., Mountford,R.C., Minor,P.D., Schild,G.C. and Almond,J.W. (1983) Nucleic Acids Research 11, 5629-5643.
14. Stanway,G., Hughes,P.J., Mountford,R.C., Reeve,P., Minor,P.D., Schild,G.C. and Almond,J.W. (1984) P.N.A.S. U.S.A. 81, 1539-1543.
15. Carroll,A.R., Rowlands,D.J. and Clarke,B.E. (1984) Nucleic Acids Research 12, 2461-2472.
16. Palmenberg,A.C., Kirby,E.M., Janda,M.R., Drake,N.L., Duke,G.M., Potratz,K.F. and Collett,M.S. (1984) Nucleic Acids Research 12, 2969-2985.
17. Cann,A.J., Stanway,G., Hauptmann,R., Minor,P.D., Schild,G.C., Clarke,L.D., Mountford,R.C. and Almond,J.W. (1983) Nucleic Acids Research 11, 1267-1281.
18. Stanway,G., Mountford,R.C., Cox,S., Minor,P.D., Schild,G.C. and Almond,J.W. (1984) Archives of Virology. 81, 67-78.
19. Grunstein,M. and Hogness,D.S.(1975) P.N.A.S. U.S.A. 72, 3961-3965.
20. Deininger,P. (1983) Anal. Biochem. 129, 216-223.
21. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) P.N.A.S. U.S.A. 74, 5463-5467.
22. Staden,R. (1980) Nucleic Acids Research 8, 3673-3694.
23. Rueckert,R.R. and Wimmer,E. (1984) J.Virol. 50, 957-959.
24. Macnaughton,M.R., Dimmock,N.J. (1976) J. Gen. Virol. 30, 263-267.
25. Nair,C.N. and Panicali,D.L. (1976) J.Virol. 20, 170-176.
26. Butterworth,B.E. (1973) Virology 56, 439-453.
27. Ahlquist,P., Kaesberg,P. (1976) Nucleic Acids Research 7, 1195-1204.
28. Yin,F.H., Lonberg-Holm,K. and Chan,S.P. (1973) J. Virol. 12, 108-113.
29. Lonberg-Holm,K., Crowell,R.L., and Philipson,L. (1976) Nature 259 679-681.

30. Rueckert,R.R. (1976) Compreh. Virol. 6, 131-213.
31. Minor,P.D., Schild,G.C., Bootman,J., Evans,D.M.A., Ferguson,M., Reeve,P., Spitz,M., Stanway,G., Cann,A.J., Hauptmann,R., Clarke,L.D., Mountford,R.C. and Almond,J.W. (1983) Nature, 301, 674-679.
32. Hopp,T.P. and Woods,K.R. (1981) P.N.A.S. U.S.A. 78, 3824-3828.
33. Wimmer,E., Jameson,B.A. and Emini,E.A. (1984) Nature 308, 19.
34. Cooney,M.K., Fox,J.P. and Kenny,G.E. (1982) Infect. Immun. 37, 642-647.
35. Halfpap,L.M. and Cooney,M.K. (1983) Infect. Immun. 40, 213-218.
36. Skern,T., Sommergruber,W., Blaas,D., Pieler,C. and Kuechler,E. (1984) Virology 136, 125-132.