

# Nucleotide sequence of group antigen (VP6) of the UK tissue culture adapted strain of Bovine Rotavirus

O.Tarlow and M.A.McCrae\*

Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK

Submitted July 16, 1990

EMBL accession no. X53667

A full length c-DNA clone corresponding to RNA segment 6, which encodes the group antigen VP6 (1), of the UK tissue culture adapted strain of Bovine Rotavirus was isolated from a pAT153 library (2). The nucleotide sequence of the viral insert was determined by dideoxy sequencing following subcloning into m13. The final sequence obtained was 1356 base pairs and contained the terminally conserved sequences found on all group A rotavirus genes (3). The gene contained a single long open reading frame (bases 24–1214) giving rise to a group antigen (VP6) of 397 amino acids with a calculated molecular weight of 44873. Comparison of the UKtc gene 6 sequence with that from the other strain of rotavirus subject to intensive molecular analysis, namely the simian virus SA11 (4), showed that the UKtc gene had a single base deletion in the 3' non-coding sequence but the overall sequence conservation was high (87.8% at the nucleotide level and 97.5% at the amino acid level). The gene 6 sequence of the RF strain of bovine rotavirus (5) was even more similar to that of the UKtc reported here (97.8% at the nucleotide level and 99% at the amino acid level).

## ACKNOWLEDGEMENTS

This work was supported by grants from the MRC. MAM is a Lister Institute Fellow.

## REFERENCES

1. McCrae, M.A. and McCorquodale, J.G. (1982) *Virology* **117**, 435–443.
2. McCrae, M.A. and McCorquodale, J.G. (1982) *J. Virol.* **44**, 1076–1079.
3. McCrae, M.A. and McCorquodale, J.G. (1983) *Virology* **126**, 204–212.
4. Estes, M.K., Mason, B.B., Crawford, S. and Cohen, J. (1984) *Nucl. Acids Res.* **12**, 1875–1887.
5. Cohen, J.B., Lefevre, F., Estes, M.K. and Bremont, M. (1984) *Virology* **138**, 178–182.

```

          H D V L Y S L S K T L K D
1  GGCTTTTAAACGAAGTCTTCAACATGGATGCTCTGACTCCTTGTCAAAACCTTTAAAG 60
      A R D K I V E G T L Y S N V S D L I Q Q
61 ATGCTAGAGACAAAATTGTCGAAGGCACATTATACTCCAATGTAAGTGATCTAATTCAAC 120
      F N Q M I I T M N G N E F Q T G G I G N
121 AATTTAATCAAATGATAATTACTATGAATGGAAATGAGTCCAAAACGGAGGAATTGGTA 180
      L P I R N W N F D F G L L G T T L L N L
181 ATCTACCGATTAGAAATGGAAATTTGATTTGGATTACTCGGAACAACTCTACTAAAT 240
      D A N Y V E T A R N T I D Y F V D F V D
241 TGGATGCCAACTACGTCGAAACGGCCCGCAATACAATGATATTTTGTAGATTTGTAG 300
      N V C M D E M V R E S Q R N G I A P Q S
301 ATAATGTATGTATGGATGAAATGGTATAGAAATCACAAGAAATGGAATTGCCACAAT 360
      D S L R K L S G I K F K R I N F D N S S
361 CAGATTCACCTAGAAAGTTGTCAGGTATTAATTCAAAAGAATAAAATTTGACAATTCAT 420
      E Y I E N W N L Q N R R Q R T G F T F H
421 CAGAATACATAGAGAAGTGGAAATTTGCAAAACAGAGACAAGAACGGGTTTACATTC 480
      K P N I F P Y S A S F T L N R S Q P A H
481 ATAAACCAACATTTCCCTTACTCAGCGTCATTCACACTGAACAGATCACAACCCAGCTC 540
      D N L M G T M W L N A G S E I Q V A G F
541 ATGATAACTTGATGGTACGATGTTGGTCAATCGGGATCAGAATTCAGGTCCGCTGGAT 600
      D Y S C A I N A P A N T Q Q F E H I V Q
601 TCGATTATTCATGTGCAATCAATGGCCGCAATACACAACAAATTTGAGCATATTTGATC 660
      L R R V L T T A T I T L L P D A E R F S
661 AGCTCGAAGAGTTGTGACTACAGCTACAATAACTCTTTTACAGATGCGAAGAGATTTA 720
      F P R V I N S A D G A T T W Y F N P V I
721 GTTTTCAAGAGTGATTAATTCAGCTCAGCGACTACTACATGGTACTTCAACCCAGTGA 780
      L R P N N V E V E F L L N G Q I I N T Y
781 TTCTTAGACCAATAACGTTGAAGTAGAGTTTCTACTAAACGGGCAGATAATAAACTT 840
      Q A R F G T I I A R N F D T I R L S F Q
841 ACCAAGCAAGATTGGAAACGATCATAGCTAGAAAATTTGATACAATTAGATTGTCATTTC 900
      L M R P P N H T P A V A A L F P N A Q P
901 AGTTGATGAGACCACCAATATGACACCAGCGGTAGCGCGGTATTTCAAAATGCCGAGC 960
      F E H Q A T V G L T L R I E S A V C E S
961 CATTGCAACATCAGGCAACAGTAGGACTCAGCTTAGAATTTGAATCTGACAGTTGTGAA 1020
      V L A D A S E T M L A N V T S V R Q E Y
1021 CAGTGCTTGCCGACGCAAGTGAACAATGCTAGCAAATGTGACATCTGTAGACAAGAAT 1080
      A I P V G P V F P P G M N W T D L I T N
1081 ACGCGATACCAGTTGGACCACTTTTCCACCAGGTATGAATTGGACTGATTGATCACTA 1140
      Y S P S R E D N L Q R V F T V A S I R S
1141 ACTATTCCATCTAGAGAGGATAATTTGCAGCGTATTTACAGTGGCTTCCATTAGAA 1200
      H L V K
1201 GCATGCTTGTCAAATGAGGACCAAGCTAACCACTTGGTATCCGACTTTGGTGAGTATGTA 1260
1261 GCTACGTCAGACTGTTGAACTCTGTAAGTAAGGATGCGTCTACGTATTGCTACACAGA 1320
1321 GTAATCACTCAGATGGCGTAGTGAGAGGATCTGACC 1356

```

\* To whom correspondence should be addressed