Genomic Concatemerization/Deletion in Rotaviruses: a New Mechanism for Generating Rapid Genetic Change of Potential Epidemiological Importance†

Y. TIAN,[1]‡ O. TARLOW,[1] A. BALLARD,[2] U. DESSELBERGER,[2]§ AND M. A. McCRAE[1]*

Department of Biological Sciences, University of Warwick, Coventry CV4 7AL,[1] and Regional Virus Laboratory, East Birmingham Hospital, Birmingham BS9 5ST,[2] United Kingdom

Three variants of group A rotavirus with large changes in their gene 5 structures have been analyzed at the molecular level. The first of these, P9Δ5, was obtained during plaque purification undertaken as part of the biological cloning of a field isolate of virus. The gene 5 homolog in this isolate migrated just ahead of the normal segment 6 RNA, giving an estimated size of 1,300 bp. Molecular cloning and sequencing of this homolog revealed it to have a single 308-bp deletion in the center of the normal gene 5 sequence extending between nucleotides 460 and 768 of the normal gene sequence. This deletion caused a frameshift in the gene such that a stop codon was encountered 8 amino acids downstream of the deletion point, giving a predicted size for the protein product of this gene of 150 amino acids compared with the 490 amino acids of its normal-size counterpart. Attempts to detect this shortened protein in virus-infected cells were not successful, indicating that it was much less stable than the full-length protein and/or had suffered a large change in its antigenicity. The second two variants, brvA and brvE, were generated in an earlier study following the high-multiplicity passage of the UKtc strain of bovine rotavirus. Polyacrylamide gel electrophoresis analysis of these nondefective variants showed that brvA had a gene 5 homolog approximately equal in size to the normal RNA segment 2 (~2,700 bp) and that brvE had a size of approximately 2,300 bp. Both variants showed changes in their gene 5 protein products, with brvA mimicking P9Δ5 in failing to produce a detectable product whereas brvE produced a new virus-specific protein approximately 80 kDa in size. Full-length cDNA clones of the brvE gene 5 homolog were isolated, and analysis of their structure revealed a head-to-tail concatemerization of the normal gene 5 sequence with the first copy of the concatemer covering nucleotides 1 to 808 and the second covering nucleotides 92 to 1579, giving a total length of 2,296 bp. Sequencing across the junction region of the two copies of the gene showed that they were joined in frame to give a predicted combined open reading frame of 728 amino acids with the amino-terminal region consisting of amino acids 1 to 258 fused at the carboxy terminus to amino acids 21 to 490. The biological phenotypes, in terms of virus yield and plaque size, of brvA and brvE were compared with those of the parent UKtc strain. This showed that both variants gave a somewhat reduced virus yield and that brvE had plaques covering an average 2.5-fold smaller area whereas brvA gave plaques whose area was reduced by approximately 50-fold. These results demonstrate that in rotaviruses rapid genomic change through both deletion and concatemerization can generate nondefective virus variants whose protein coding potential has been altered, leading to changes in biological phenotype. This new mechanism for effecting rapid genomic change has now to be considered one by which rotaviruses are able to rapidly change the nature of the proteins they encode, which may in turn lead to important epidemiological changes.

The group A rotaviruses are the predominant etiological agents of acute viral gastroenteritis in the young of a wide range of avian and mammalian species, including humans (6, 11). As a consequence, they are major medical and veterinary pathogens, for which a clear definition of the mechanisms by which they are able to induce variations in their genomes forms an integral part of attempts to develop vaccines aimed at combatting them.

The rotavirus genome is composed of 11 discrete segments of double-stranded RNA whose overall migrational pattern following fractionation by electrophoresis on polyacrylamide gels (PAGE) is highly conserved and in the case of group A viruses is 4, 2, 3, 2 (4, 25). Slight changes in the migrational positions of individual RNA segments have been attributed to genomic variation being generated either by simple point mutations resulting in possible antigenic drift events or through the occurrence of genome segment reassortment events between genetically distinct virus strains. The latter of these is well recognized, particularly from studies on type A influenza viruses (33), as a mechanism by which viruses with segmented genomes are able to effect sudden large changes in their genome which may give rise to antigenic shift events of major epidemiological importance for disease prevention (24).

In addition to the minor changes in rotavirus genome profile, two more profound types of change have been reported. In the first, the tight triplet of RNA segments (genes 7 to 9) is reduced to a doublet with the displaced segment in some cases migrating more slowly on PAGE and in other cases migrating more rapidly (25, 26). Virus isolates exhibiting this type of genomic change also failed to react in immunological assays to detect the rotavirus group antigen previously thought to be present in all rotaviruses irrespective of their species of origin (25, 26). A combination of serologically and nucleic

acid-based assays have been employed to show that these atypical viruses fall into a number of distinct groups sufficiently different from each other that they are likely to form separate gene pools in a manner analogous to that found in the type A, B, and C influenza viruses. As a consequence, the rotaviruses are now recognized as being made up of a number of distinct groups, of which five (A to E) have been clearly defined to date (19), with the group A viruses being currently by far the most important agents of disease.

The second type of gross genomic abnormality, first recognized in group A rotaviruses recovered from profoundly immunodeficient children, is that in which either additional RNA segments are present on the genome profile or a particular RNA segment has disappeared from its normal migrational position and been replaced by an RNA migrating much more slowly (2, 9, 14, 27). The initial studies carried out on the isolates made from the immunodeficient children showed that the additional RNA segments were not derived as expected by deletion of sequence from larger RNAs but rather resulted from some form of genomic concatemerization (27). Detailed sequence analysis has been reported from both human and animal virus isolates in which such concatemerizations have taken place (1, 12, 17, 23, 30). The general picture that has emerged is exemplified by the situation found in those isolates having such concatemerizations in RNA segment 11, where the mobility shift was found to be due to a partial concatemerization of gene 11 with the duplication occurring downstream of the normal open reading frame such that a normal protein product could be produced (12, 17, 18, 30). Inasmuch as these concatemerizations failed to affect the coding potential of the gene in question, they remained of somewhat academic interest.

In a previous study from one of our groups, we reported that similar genomic concatemerization events could be stimulated by high-multiplicity passage in tissue culture of the UKtc strain of bovine rotavirus (13). Among the viruses isolated in that study were some in which a concatemerization in RNA segment 5 appeared to affect the coding capacity of the gene (13). In the present study, we have both extended this earlier work to define the nature of the concatemerization in one case and studied a newly tissue culture-adapted field isolate with an apparent large deletion in RNA segment 5 which also affects its coding potential.

MATERIALS AND METHODS

Viruses and cells. The UKtc Compton strain of bovine rotavirus and the brvA and brvE derivatives of it (13) were propagated in either BSC-1 or MA-104 cells as previously described (20). Stocks of the brvA and brvE rearranged RNA segment variants were prepared at a low multiplicity of infection (MOI) (<0.1) from virus that had been plaque purified three times. The P9 and P9Δ5 used in this study were kindly provided by P. Isa, Moredun Research Institute; these isolates were made by plaque picking from passage 3 of the P9 isolate of virus. This isolate was made originally by tissue culture adaptation of virus from a diarrheic foal (31). A low-passage (<5) stock of this isolate was passaged (twice) at a low MOI in MA-104 cells before being subjected to plaque isolation. Each of these isolates was purified by three cycles of plaque-to-plaque picking before virus stocks were prepared at a low MOI in MA-104 cells (13, 20).

Nucleic acid analyses. Genome segment profiles of the viruses used in this study were prepared by PAGE analysis carried out as previously described (4) with the exception that RNA bands were detected by staining of gels with ethidium bromide. Northern (RNA) blotting of fractionated genomic RNA was performed as described by Pedley et al. (27), and gene 5 sequences were detected by using a cDNA clone of the UKtc gene 5 (21) made radioactive by nick translation (28).

The isolation of cDNA clones of the rearranged genes 5 from the different viruses was undertaken with infected cell RNA as a starting point. Confluent monolayers of BSC-1 cells were infected at a low MOI (<0.1), and when approximately 10% of cells showed a cytopathic effect (3 to 4 days), RNA was prepared from the cytoplasm as previously described (35). Infected cell RNA was reverse transcribed and polymerase chain reaction amplified with terminal primers as previously described (34, 35). The sequence of the gene 5 primers used in the cloning was as follows: 5' end primer, CCCGGGATCC ATGGCCGGCTTTTTTTATGA; and complementary 3' end primer, CGATCGCGAATTCTGCAGGTCACATTTTAT, with the region underlined in each case being that complementary to the terminus of the appropriate strand of gene 5. The polymerase chain reaction-amplified cDNA either was used directly for DNA sequencing or was inserted into BamHI-EcoRI-digested Bluescribe and propagated in Escherichia coli (35). DNA sequencing was carried out by the dideoxy chain termination method (29) as previously described (35).

Protein analysis. Radioactive labelling of virus-infected cells was carried out as previously described (20, 22). Infected cell extracts were fractionated on 5 to 11% gradient polyacrylamide gels (20, 22). Immunoprecipitation studies were carried out on both infected cell lysates and supernatants made from them at 100,000 × g (22). Immunoprecipitation and washing of precipitates collected with protein A-Sepharose were carried out as described by Watson et al. (32). Western blotting (immunoblotting) of fractionated infected cell lysates was carried out (16), with viral proteins being detected with either hyperimmune convalescent bovine serum or monospecific serum against VP5 kindly supplied by J. Cohen.

Nucleotide sequence accession numbers. The sequence of gene 5 of the UKtc strain of bovine rotavirus has the EMBL data base accession number Z12108. The sequence of the P9Δ5 isolate has been submitted to the EMBL data base and given the accession number Z24736. The junction sequence of the gene 5 homolog of the brvE isolate has also been submitted to the EMBL data base and given the accession number Z24735.

RESULTS

Overall genome analysis of rearranged variants. In a previous study, a number of replication-competent variants of the UKtc bovine rotavirus were derived by high-multiplicity passage in MA-104 cells (13). Two of these, brvA and brvE, were selected for further analysis in this study. When the genome profiles of these two viruses were examined by PAGE analysis, a large change in the pattern of RNA segments was observed (Fig. 1). In both variants, the RNA segment migrating in the normal gene 5 position had disappeared and a new band, comigrating with RNA segment 2 in the case of the brvA variant and just ahead of RNA segment 4 in the case of brvE, was evident on the electropherogram (Fig. 1). In the previous work, the new band from one of these variants (brvA) was analyzed by two-dimensional RNA fingerprinting which indicated that it contained gene 5-specific sequences. To confirm and extend this result, Northern blot analysis was performed on the fractionated RNA segments with a full-length cDNA clone of gene 5 from the UKtc strain of bovine rotavirus as the probe. Results from this (Fig. 2) showed that the new bands present on both the brvA and the brvE electropherograms contained gene 5 sequences.
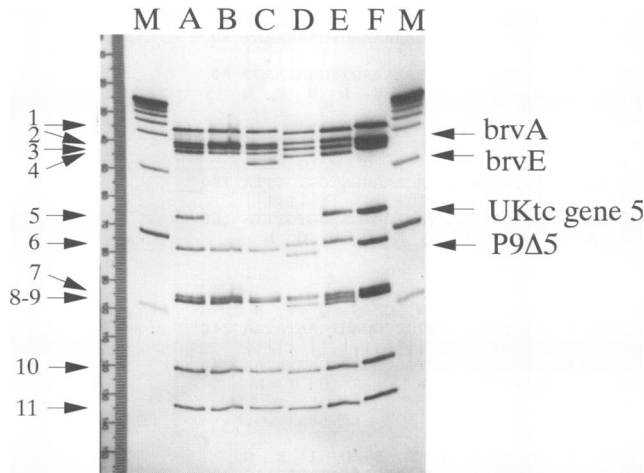
FIG. 1. PAGE fractionation of genomic double-stranded RNAs. Viral genomic double-stranded RNA was prepared from virus-infected cells and fractionated on a 6% polyacrylamide gel as described in Materials and Methods. Following electrophoresis, the RNAs were stained with ethidium bromide and photographed under UV light. The arrows and labelling down the left-hand side of the gel indicate the migrational positions of the various genomic RNA segments of the UKtc strain of bovine rotavirus. The arrows and labelling down the right-hand side of the gel indicate the migrational positions of the normal UKtc gene 5 and of bands with altered migration in the three isolates studied. Lanes: M, DNA size markers (the visible separated bands have sizes, in ascending order from the bottom of the gel, of 1.0, 1.6, 2.0, 3.0, 4.0, and 5.0 kb); A and F, UKtc strain of bovine rotavirus; B, brvA isolate; C, brvE isolate; D, P9Δ5 isolate; E, P9 isolate.

The P9Δ5 variant was isolated during plaque picking carried out as part of adapting the P9 isolate to growth in tissue culture. PAGE analysis carried out on virus stocks grown from individual P9 plaques revealed a number of electrophoretic
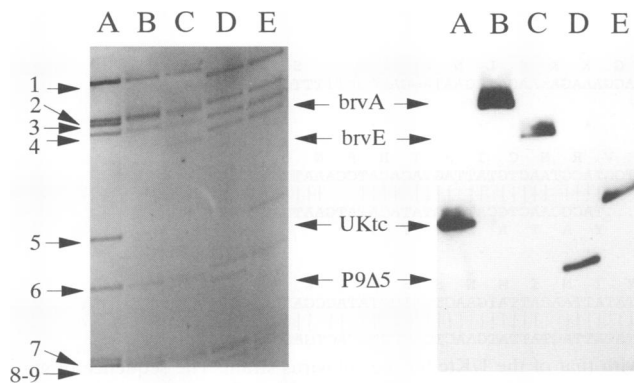
FIG. 2. Northern blot analysis of gene 5 variants. Viral genomic RNAs were fractionated on a 6% polyacrylamide gel, transblotted onto nylon membrane, and probed with a ^{32}P-labelled cDNA probe of the UKtc gene 5 all as described in Materials and Methods. The left-hand panel of the figure shows a photograph of the ethidium bromide-stained gel covering the gene 1 to 9 region, and the right-hand panel shows an autoradiogram of the Northern blot of the gel after having been probed with gene 5 cDNA. The arrows and labelling down the left-hand side of the figure indicate the migrational positions of the various genomic RNAs of the UKtc strain of bovine rotavirus. The arrows and labelling between the two panels indicate the positions on the two panels of gene 5 of the UKtc strain and the gene 5 homologs in the other virus isolates. Lanes: A, UKtc strain of bovine rotavirus; B, brvA isolate; C, brvE isolate; D, P9Δ5 isolate; E, P9 isolate.

variants, indicating that the original field isolate was almost certainly a mixed population (14a). The predominant virus population had the electropherogram designated as P9 in Fig. 1 with a normal pattern of RNA segments. In common with the brvA and brvE isolates, P9Δ5 did not have an RNA segment migrating in the normal gene 5 position, but in this case, the new band present on the electropherogram did not migrate more slowly than the normal gene 5 but rather migrated faster to a position just ahead of RNA segment 6 (Fig. 1). It is evident from comparison of the P9 and P9Δ5 profiles that in addition to the major difference in gene 5 migration small migrational differences can be seen for other segments (Fig. 1), indicating that P9Δ5 is probably not derived directly from P9, and certainly its isolation did not require high-multiplicity passage in tissue culture. Northern blot analysis of P9Δ5 showed that in this case the new band appearing on the electropherogram also contained gene 5 sequences (Fig. 2).

Protein studies on rearranged variants. To examine the effects of the rearrangements in gene 5 on its protein coding potential, pulse-chase studies were performed. The normal
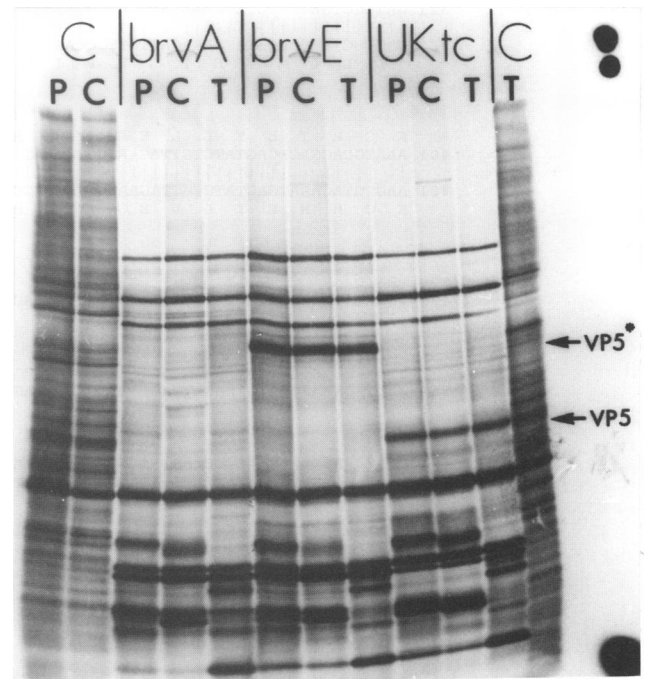
FIG. 3. Protein analysis of gene 5 variants. Protein labelling of virus-infected cells and subsequent fractionation of those cells on 5 to 11% linear gradient polyacrylamide gels were carried out as described in Materials and Methods. The lanes labelled P were produced from cells that had been labelled for 15 min with [^{35}S]methionine and then immediately harvested for gel analysis. The lanes labelled C were produced by labelling the cells as for the P lanes but then chasing the samples by overlaying them with medium containing 100 times the normal level of methionine and incubating for 2 h before harvesting. The lanes labelled T were produced by maintaining the cells in the presence of 5 μg of the N-linked glycosylation inhibitor tunicamycin per ml up to and including the labelling period which was done exactly as for the pulse-labelled (P) lanes. The two left-hand lanes and the right-hand lane of the gel represent controls made by labelling uninfected cells. The particular virus used to generate each group of three lanes is indicated above each block. The positions of the normal gene 5 product (VP5) and of its homolog produced by the brvE variant (VP5*) are indicated on the right-hand side of the gel.

```
                                  M   A   T   F   K   D   A   C   Y   H   Y   K   K   L   N   K   L
  1 GGCTTTTTTTATGAAAAGTCTTGTGTGAGCCATGGCGACTTTTAAAGACGCTTGCTATCATTACAAAAAACTGAATAAAT 80
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| || ||||| ||||||| |
  1 GGCTTTTTTTATGAAAAGTCTTGTGTGAGCCATGGCGACTTTTAAAGACGCTTGTTATCACTATAAAAAGTTGAATAAGT 80
                                  M   A   T   F   K   D   A   C   Y   H   Y   K   K   L   N   K   L


                 N   S   L   V   L   K   L   G   A   N   D   E   W   R   P   A   P   V   T   K   Y   K   G   W   C   L
 81 TAAATAGCTTAGTACTTAAACTAGGAGCAAATGATGAATGGAGACCAGCACCAGTGACAAAATACAAAGGGTGGTGTTTA 160
    ||||| |||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||| ||||| |||||||||
 81 TAAACAGCTTAGTACTCAAACTAGGAGCAAATGATGAATGGAGGCCAGCACCAGTGACAAAATATAAAGGATGGTGTTTA 160
                 N   S   L   V   L   K   L   G   A   N   D   E   W   R   P   A   P   V   T   K   Y   K   G   W   C   L


                 D   C   C   Q   Y   T   N   L   T   Y   C   R   G   C   A   L   Y   H   V   C   Q   W   C   S   Q   Y   N
161 GACTGCTGTCAATATACAAATTTAACATATTGCAGAGGATGTGCTTTGTACCATGTATGTCAGTGGTGCAGTCAATATAA 240
    || ||||||||||||||||||| ||||||||||| ||||| | ||||||||||||||||||||||||||||||| |||||
161 GATTGCTGTCAATATACAAATTTGACATATTGCAGGGGATGCGCCCTATACCATGTATGTCAGTGGTGCAGTCAGTATAA 240
                 D   C   C   Q   Y   T   N   L   T   Y   C   R   G   C   A   L   Y   H   V   C   Q   W   C   S   Q   Y   N


                 R   C   F   L   D   E   E   P   H   L   L   R   M   R   T   F   K   D   V   I   T   K   E   D   I   E   G
241 TAGGTGTTTCTTAGATGAAGAACCTCACTTGTTAAGAATGCGAACATTCAAGGATGTAATAACGAAAGAGGATATAGAAG 320
    ||||||||||||||||||||||||||||||||||| |||||||||||||| ||||||||||| ||| |||||||||||||
241 CAGGTGTTTCTTAGATGAAGAACCTCATTTGTTAAGAATGCGAACATTTAAAGATGTAATAACAAAAGAAGACATAGAAG 320
                 R   C   F   L   D   E   E   P   H   L   L   R   M   R   T   F   K   D   V   I   T   K   E   D   I   E   G


                 L   L   T   M   Y   E   I   L   F   P   I   N   E   K   L   V   N   K   F   I   N   S   V   K   Q   R
321 GATTGCTGACCATGTATGAAATATTGTTTCCAATAAATGAAAAATTAGTGAATAAATTCATAAATTCTGTGAAACAACGT 400
    || ||||| |||||||||||||||||||| |||||| ||||||||||||| ||||||||||||||||||| ||||| || |||
321 GACTGCTAACCATGTATGAAACATTGTTCCCAATAAATGAGAAGTTAGTGAATAAATTCATAAACTCTGTAAAGCAGCGT 400
                 L   L   T   M   Y   E   T   L   F   P   I   N   E   K   L   V   N   K   F   I   N   S   V   K   Q   R


                 K   C   R   T   E   Y   L   L   E   W   Y   N   H   L   L   M   P   I   T   L   Q   A   L   T   I   K   L
401 AAATGCAGGACTGAGTATCTGTTAGAATGGTATAATCATCTACTGATGCCAATAACATTACAAGCATTAACTATAAAACT 480
    || || || | |||||||||||||||||||||||||||||||| | || |||||| ||||||||
401 AAGTGTAGAAATGAGTATCTGTTAGAATGGTATAATCATTTGCTAATGCCAGTAACATTA..................... 460
                 K   C   R   N   E   Y   L   L   E   W   Y   N   H   L   L   M   P   V   T   L   Y   A   T   A   S   V


                 E   D   S   T   Y   Y   I   F   G   Y   Y   D   C   M   E   H   E   N   Q   T   P   F   Q   F   I   N   L
481 TGAAGATAGCACATACTATATATTTGGATACTATGACTGTATGGAACATGAGAACCAGACACCATTCCAGTTTATTAATT 560


                 L   E   K   Y   D   K   L   L   L   D   D   R   N   F   N   R   M   L   H   L   P   T   I   L   Q   Q
561 TATTAGAAAAGTATGACAAATTACTATTAGATGATAGAAATTTCAATAGAATGTTACATTTACCAACAATATTGCAGCAA 640


                 E   Y   A   L   R   Y   F   S   K   S   R   F   L   S   K   G   K   K   R   L   N   R   N   D   F   S   D
641 GAATATGCGTTAAGATATTTTTCAAAATCAAGATTTTTAAGTAAAGGAAAGAAAAGATTGAATAGGAACGATTTTTCAGA 720


                 N   L   V   E   D   R   H   S   P   T   S   L   I   Q   V   V   R   N   C   I   S   T   H   P   N   D   Y
721 TAATCTTGTGGAAGACAGACATAGTCCAACATCATTAATACAGGTGGTACGTAACTGTATTAGTACACATCCAAATGATT 800
                                                 |||| |||||| || |||| |||| ||||||
461 ...............................................TACGCAACTGCATCAGTATACATATGAATGATT 493
                                                 Y   A   T   A   S   V   Y   I   *


                 E   W   N   K   A   C   T   F   V   V   D   A   R   N   Y   I   N   I   M   N   S   S   Y   T   E   H
801 ATGAATGGAATAAGGCATGTACGTTTGTGGTTGATGCTAGAAATTATATTAACATTATGAACTCATCGTATACCGAACAT 880
    ||||||||||| |||||||||| ||| |||||||| |||||||||||||||| |||||||||| || || || |||
494 GTGAATGGAATAAAGCATGTACGCTTATAGTTGATATTAGAAATTATATTAGTATTATGAACTCATCCTACACTGAGCAT 573
```

FIG. 4. Comparison of the sequence of the gene 5 homolog from P9Δ5 with that of the UKtc bovine rotavirus strain. The sequence of gene 5 of the UKtc strain of bovine rotavirus given in the upper row of the comparison was obtained in this laboratory (33a). The sequence of the P9Δ5 isolate given in the lower row of the comparison was determined from cDNA clones as described in Materials and Methods. The alignment of the two sequences was carried out with the Bestfit program which forms part of the Wisconsin sequence analysis package (7).

protein product encoded by RNA segment 5 is a protein of approximately 53 kDa designated VP5 in the case of the UKtc strain of bovine rotavirus (20) and NS53 in the simian SA11 strain (10). The normal gene 5 protein product was synthesized in amounts that could be easily detected when the proteins labelled during a 15-min pulse with [35S]methionine carried out at 6 h postinfection were analyzed by PAGE (Fig. 3). The mobility and quantity of this protein were not affected by either

allowing a 2-h chase before harvesting the sample or carrying out the pulse-labelling in cells that had been maintained after infection in the presence of the N-linked glycosylation inhibitor tunicamycin. This confirmed that VP5 is not rapidly turned over in infected cells, nor does it appear to be posttranslationally modified by N-linked glycosylation. Corresponding analysis carried out on the brvA and brvE variants revealed that in both cases the only protein product whose mobility appeared

```
          Y   S   V   S   Q   R   C   K   L   F   T   K   Y   K   F   G   I   I   S   K   L   V   K   P   N   Y   I
      881 TATAGTGTTTCACAAAGATGTAAGCTGTTTACTAAATATAAATTTGGGATTATATCAAAATTGGTAAAGCCAAATTACAT 960
          |||||||| || |||||||||||  ||||||||| ||||||||||||||||||||| ||| || || || ||||||||
      574 TATAGTGTGTCGCAAAGATGTAAATTGTTTACTAAGTATAAATTTGGGATTATATCAAGATTAGTGAAACCGAATTACAT 653


          F   S   S   H   E   S   C   A   L   N   V   H   N   C   R   W   C   Q   I   N   S   H   Y   K   V   W   E
      961 TTTTTCTAGTCATGAATCATGTGCATTAAATGTACATAATTGTAGATGGTGTCAGATAAACAGCCATTACAAAGTGTGGG 1040
          |||||| || ||| |||||||  ||||||||||||||||||||| |||||| || || |||| |||||||||||||||||
      654 TTTTTCCAGCCATAAATCATGCGCATTAAATGTACATAATTGTAAATGGTGCCAAATCAACAACCATTACAAAGTGTGGG 733


          D   F   R   L   K   K   I   Y   N   N   V   M   D   F   I   R   A   L   V   K   S   N   G   N   V   G
     1041 AAGATTTTAGACTGAAGAAGATATATAATAATGTGATGGATTTTATCAGGGCACTTGTAAAGTCAAACGGAAATGTTGGG 1120
          |||||||||||| ||||| |||||||||||| | |||||||||||||||||||||||||||| ||||| ||||||||||||
      734 AAGATTTTAGACTTAAGAAAATATATAATAATATAATGGATTTTATCAGGGCACTTGTGAAGTCGAACGGAAATGTTGGA 813


          Q   C   S   S   Q   E   S   V   Y   K   C   I   P   D   I   F   L   I   C   K   M   E   K   W   N   E   A
     1121 CAGTGTTCGTCACAAGAATCAGTGTATAAATGTATACCAGATATATTTTTAATTTGTAAAATGGAAAAATGGAATGAAGC 1200
          || |||||||||||  |||||||||||||||||  |||||||||||||||||||  |||||| ||||||||||||||||||
      814 CATTGTTCGTCACAGGAATCAGTGTATAAATATATACCAGATATATTTTTAATCTGTAAAACGGAAAAATGGAATGAAGC 893


          V   E   V   L   F   N   Y   L   E   P   V   D   I   N   G   T   E   Y   V   L   L   D   Y   E   V   N   W
     1201 TGTCGAAGTATTATTTAATTATCTAGAACCAGTGGATATAAATGGAACGGAGTATGTGCTATTGGACTATGAAGTGAACT 1280
          ||||||| | ||||||||||||||||||||||||||||||||||||||||||  ||||||||  |||||||||||||||||
      894 TGTCGAAATCTTATTTAATTATCTAGAACCAGTGGATATAAATGGAACAGAGTATGTATTATTGGACTATGAAGTGAACT 973


          E   V   R   G   L   V   M   Q   S   M   D   G   K   V   P   R   I   L   N   I   N   D   T   K   K   I
     1281 GGGAAGTGAGAGGATTAGTTATGCAGAGTATGGATGGAAAAGTACCAAGAATTTTGAACATAAATGATACAAAGAAGATA 1360
          ||||||||||||| |||| ||||| | |||||||| |||||||| |||||||| || |||||||||||||||||||||||
      974 GGGAAGTGAGAGGACTAGTCATGCAAAATATGGATGGGAAAGTACCGAGAATTTTGAATATGAATGATACAAAGAAGATA 1053


          L   S   T   I   I   F   D   W   F   D   V   R   Y   M   R   E   T   P   M   T   T   S   T   T   N   Q   L
     1361 TTAAGTACAATAATATTTGATTGGTTTGACGTAAGATACATGAGAGAAACGCCAATGACGACGTCAACAACGAATCAGCT 1440
          |||| ||||||||||||||| |||||||||  |||||| ||||||||||| |||||||| |||||||| || || ||
     1054 CAAAGTGCAATAATATTTGACTGGTTTGACACAAGATATATGAGAGAAACACCAATGACAACGTCAACAACAAACCAACT 1133


          R   T   L   N   K   R   N   E   L   I   D   E   Y   D   L   E   L   S   D   V   E   *
     1441 TAGAACACTGAATAAAAGGAATGAGCTCATAGATGAGTACGACTTAGAACTTTCAGATGTTGAATAAATATGGTCACAAA 1520
          | | || || |||||| | |||||||||  ||||||||||||| ||||||||||||||||||||  ||||||||||||||
     1134 TCGGACGCTAAATAAGAAGAATGAGCTTATAGATGAGTACGATTTAGAACTTTCAGATGTTGAATAAGTATGGTCACAAC 1213


     1521 TACGCCAGCACAAGACCGTCTGCACTAGAGTAGCGCCTAGGCAGCAATAAAATGTGACC 1579
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1214 TACGCCAGCACAAGACCGTCTGCACTAGAGTAGCGCCTAGGCAGCAATAAAATGTGACC 1271
```

FIG. 4—*Continued.*

altered was VP5; in the case of brvA, the normal VP5 disappeared from the protein profile and no new protein bands were detected, whereas for brvE, the normal VP5 was replaced by a much larger protein (VP5*) migrating just ahead of VP4 on PAGE (Fig. 3). Allowing for its difference in methionine content (see below), about the same amount of VP5* was synthesized and, in common with VP5, its amount and mobility were not affected by either a 2-h chase prior to harvesting or the use of tunicamycin (Fig. 3). Similar studies were carried out on P9Δ5, but in this case the polyacrylamide gels were run for a shorter time than that for Fig. 3 to ensure that the predicted 150-amino-acid VP5 homolog would not have migrated off the bottom of the gel. Despite this, the results obtained mirrored what was found for brvA, with no new protein band being evident on the gel profile (data not shown). In an attempt to identify a VP5 homolog(s) in brvA- and P9Δ5-infected cells that may have been masked by other viral proteins, both immunoprecipitation and Western blot studies were carried out with a monospecific VP5 serum. These experiments (data not shown) gave a strong specific signal in the case of wild-type UKtc, P9, and the brvE variant but failed to identify any VP5

analog in the other two variants, indicating that they either were much less stable than the normal protein or alternatively were so altered in their antigenicity that they were no longer recognized by the monospecific serum.

Molecular cloning and DNA sequence analysis. To definitively characterize the nature of the gene rearrangements in these gene 5 variants, molecular cloning with a polymerase chain reaction-based strategy (34) was undertaken. Full-length cDNA clones of the rearranged genomic segments were obtained from both P9Δ5 and brvE, but we were not successfully able to obtain such a clone from brvA, although this has recently been achieved by others (24a). Dealing with P9Δ5 first, we found that the full-length clones were approximately 1,300 nucleotides in length, and because this was an independent variant not derived by tissue culture passage under unusual conditions, we decided to completely sequence the variant gene 5. This revealed that P9Δ5's gene 5 homolog was 1,272 nucleotides long compared with the 1,579 nucleotides of the normal gene from the UKtc bovine strain. When these two sequences were compared (Fig. 4), a high level of sequence homology, 90% at the nucleotide level and 98% at the protein

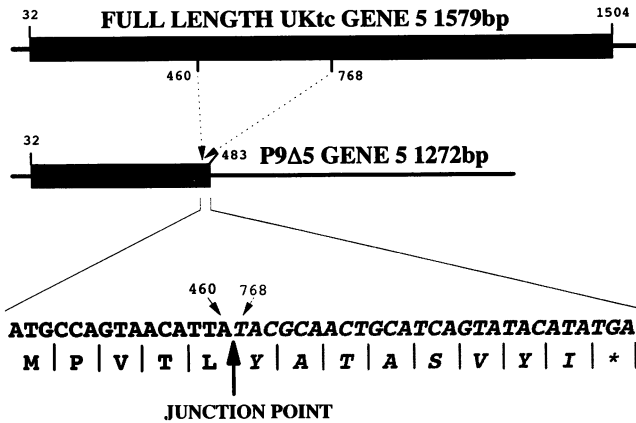FIG. 5. Overall gene arrangement and junction sequence of the gene 5 homolog of the P9Δ5 isolate.

FIG. 6. Overall gene arrangement and junction sequence of the gene 5 homolog of the brvE isolate.

level in the normal reading frame, was found for the aligned sequences. This alignment also showed that the relationship between the two sequences was that P9Δ5 had a single deletion of 308 bases extending between nucleotides 460 and 768 of the normal gene 5 sequence (Fig. 4). The effect of this deletion was to shift the reading frame of the gene such that a termination triplet would be encountered 8 amino acids downstream of the point of deletion (Fig. 5), leading to a predicted protein product of only 150 amino acids compared with the 490 of the normal VP5.

Full-length cDNA clones of the brvE gene 5 homolog were approximately 2,300 nucleotides in length (Fig. 6). As the sequence of the normal gene 5 from the UKtc bovine strain, which was the progenitor of brvE, had already been completed in parallel studies being carried out in this laboratory (33a), the initial characterization of brvE clones was done by restriction enzyme analysis (data not shown). This revealed that in common with P9Δ5 the rearrangement in brvE occurred at a single point but in this case appeared to involve a duplication of sequence with the normal sequence proceeding for the first approximately 800 nucleotides and then returning to approximately nucleotide 90 before proceeding to the normal 3' terminus. To define the exact nature of the rearrangement, the region extending from a unique XmnI site at nucleotide 370 in the first copy of the sequence to the corresponding base position in the second copy was subcloned from three independent full-length clones into m13. In these subclones, the region corresponding to nucleotides 750 to 900 of the rearranged gene was sequenced (data not shown). This revealed that they all had the same junction point occurring between nucleotide 808 of the first copy of the sequence and nucleotide 92 of the second copy (Fig. 6). The consequence of this fusion was to join the coding sequences of the two partial copies of gene 5 together in frame, giving a predicted protein size for VP5* of 728 amino acids.

Biological properties of bovine rotavirus variants. The three rearranged variants examined in this study were all nondefective, able to grow to high titer in tissue culture, and capable of giving virus plaques. Comparison of the genome profiles of P9 and P9Δ5 (Fig. 1) revealed small differences in migrational positions of RNA segments other than gene 5, suggesting that these two viruses may not be closely related to each other and hence that any comparison of their biological properties with a view to defining possible biological functions for gene 5 would not be valid. The brvA and -E isolates were, however, derived
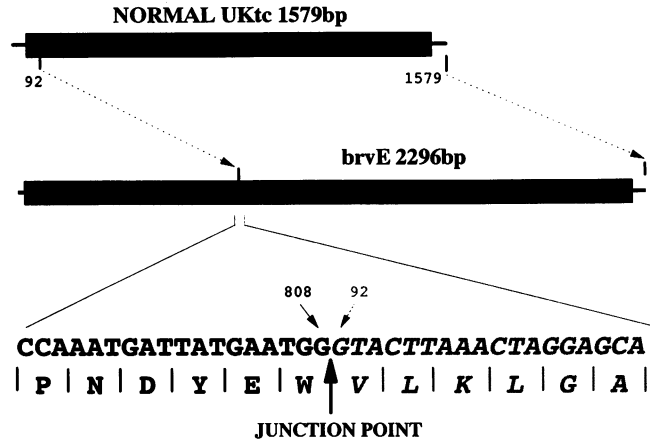
directly from the UKtc bovine strain, so some comparison of their biological phenotypes was undertaken. The results (Table 1) showed that the two variants showed a 9- to 60-fold lower yield in single-step growth experiments although the kinetics of virus growth were not changed (data not shown). In addition, both viruses gave smaller plaques in MA-104 cells, with the brvE variant showing an approximately 2.5-fold reduction in plaque area and brvA showing a more than 50-fold reduction in mean plaque area (Table 1).

DISCUSSION

This study has focused on three virus isolates showing the loss of an RNA segment 5 migrating at its normal position on PAGE and its replacement by a new RNA band which in two cases, brvA and -E, migrated more slowly on the gel, indicating a concatemerization of sequence, and in the third, P9Δ5, migrated more rapidly, suggesting deletion of sequence. These three isolates were of particular interest because protein analyses carried out on them in this and previous work demonstrated that the changes seen in gene 5 were reflected in alterations in the protein that it encodes, namely, VP5. The Northern blot analysis confirmed and extended previous work done solely on brvA, showing that the new RNA bands seen on PAGE analysis did contain gene 5 sequences.

Cloning and sequencing of the first of these isolates, P9Δ5, revealed that the gene had a 308-bp deletion extending from nucleotides 460 to 768. The effect of this deletion was to shift the reading frame of the translated protein such that a termination triplet would be encountered 8 amino acids beyond the point of deletion and hence the protein product would have a predicted length of 150 amino acids. This truncated protein had a predicted pI of 8.25 compared with the pI of 7.76 of the full-length protein and contains the amino-terminal region zinc finger motif which lies between amino acids 53 and 69 and has been found in all nine of the group A genes 5 sequenced to date (33a). With the exception of this deletion of sequence from the center of the gene, the segment 5 homolog P9Δ5 showed a very high level of sequence homology with the UKtc gene 5, and consequently, it seems reasonable to suppose that the transcription of the deleted gene occurred at levels similar to that previously established for the UKtc virus (15). Despite this, attempts to demonstrate the synthesis of the VP5 homolog were not successful in straightforward pulse-chase labelling studies, and the use of either

TABLE 1. In vitro growth properties of UKtc and bovine
rotavirus variants

Virus	Log PFU/ml at time postinfection[a]:		Plaque diam (mm) at 7 days (mean ± SD)	Mean plaque size (mm²)
	30 h	46 h		
UKtc	8.9	8.5	7.6 ± 0.8 (n = 9)	45.4
brvE	7.8	7.5	4.8 ± 0.7 (n = 16)	18.1
brvA	7.2	7.1	1.0 ± 0.7 (n = 50)	0.4

[a] Single-step growth experiments were carried out in MA-104 cells infected at an MOI of 10 PFU per cell.

immunoprecipitation or Western blotting failed to help in its detection. This strongly suggested that the VP5 homolog is either turned over much more rapidly than its normal counterpart or alternatively is sufficiently changed in its antigenicity that it is no longer able to interact with the monospecific serum used. The fact that P9Δ5 is nondefective for growth in tissue culture also indicates that any essential role that VP5 plays in the growth of the virus is encoded in the amino-terminal 150 amino acids of the protein; formally of course, the possibility exists that at least for replication in vitro this protein may be completely dispensable. Given the recent evidence suggesting that gene 5 may be involved in the pathogenicity of rotaviruses (3), it will be interesting to examine its effect on this viral property when reassorted into a pathogenic strain background. In contrast to P9Δ5, the gene 5 of the brvE isolate had not suffered a deletion event but a concatemerization of sequence with the first copy of the gene extending from nucleotides 1 to 808 of the normal sequence, at which point it was fused to a partial second copy of the sequence extending from nucleotides 92 to 1579, giving a total size for the gene 5 homolog of 2,296 bp, which was commensurate with its migration position on PAGE analysis just ahead of gene 4, which has a size of 2,359 bp (10). The junction point between the two copies of the gene occurred in frame, and hence, the effect of the concatemerization was to generate an extended open reading frame in which amino acids 21 to 490 of the normal protein were fused onto the carboxy terminus of amino acids 1 to 258, giving this VP5 homolog a total length of 728 amino acids, an approximately 50% increase over that of the normal VP5. In this case, protein labelling experiments showed that the protein is made in approximately the same amount as the normal VP5 and there was no major change in its stability. As in the case of P9Δ5, the nondefective nature of this isolate indicates that this extended VP5 homolog is still able to fold in a way that preserves any essential function(s) that it may possess for virus replication in vitro. In this study, we were unable to generate a full-length cDNA of the altered-mobility RNA segment from the brvA isolate. Therefore, apart from demonstrating in Northern blot analysis that it contained sequence derived from gene 5 and showing that it did not synthesize a detectable VP5 homolog in virus-infected cells, we were unable to further characterize this variant. However, others have recently been able to obtain full-length cDNA clones of the gene 5 homolog in brvA and complete the sequencing of it (24a). This revealed that in this case the homolog was 2,693 bp in length, with the same general head-to-tail partial duplication as seen in brvE with the first copy of the gene extending from nucleotides 1 to 1450 and the second copy extending from nucleotides 337 to 1579. In this case, the two copies were not joined in frame, but this fact was somewhat irrelevant as a point mutation at nucleotide 808 of the first copy of the gene introduced a

termination codon, giving rise to a predicted size for the VP5 homolog of 258 amino acids. Also in contrast to our results, Patton et al. were able to detect this protein product in Western blots made from virus-infected cells (24a).

The three rotavirus variants examined in this study are the first to be studied in detail in which sudden rearrangements of gene sequence either by deletion or by concatemerization have resulted in changes in the protein coding potential of the gene involved. The main genetic mechanism by which sudden epidemiologically important changes are thought to occur in the genome of rotaviruses is gene reassortment, and there is ample evidence in the literature to indicate that when two genetically distinct rotaviruses infect a single cell in tissue culture reassortment occurs at high frequency. However, the evidence that such events play a role in the genetic diversity of viruses circulating in animal and human populations is more scanty (5). By contrast, there have now been several reports of genomic rearrangements having taken place in rotaviruses present either in the clinical sample itself or in viruses that have been isolated from it by standard tissue culture adaptation techniques (1, 2, 9, 12, 14, 17, 18, 23, 27, 30). In addition to rotaviruses, gene rearrangements of a similar type in orbiviruses have also been reported (8). Where these events have been analyzed in detail, they have been found to result from gene concatemerization events that do not affect the normal coding potential of the gene in question and are hence of somewhat academic interest as far as providing a route to rapid epidemiological change is concerned. The present study has shown that replication-competent viruses with altered protein coding potential can be generated by rearrangement events, both deletions and concatemerizations. In two of the examples studied here, the rearranged variants have a measurable difference in a biological phenotype. This change in phenotype may well have resulted directly from the rearrangement event although additional work involving both complete sequencing of the brvE gene 5 homolog and reassortant studies to show that the altered phenotype cosegregates with the segment 5 homologs will be required before a definitive conclusion can be reached.

Therefore, although further work will clearly be needed before any assessment of the importance of rearrangement events to the generation of overall genetic diversity in rotaviruses can be made, nevertheless it must now be considered another mechanism available to viruses for bringing about rapid changes in their genomes of potential epidemiological importance.

REFERENCES

1. **Ballard, A., M. A. McCrae, and U. Desselberger.** 1992. Nucleotide sequences of normal and rearranged RNA segments 10 of human rotaviruses. J. Gen. Virol. **73:**633–638.
2. **Besselaar, T. G., A. Rosenblatt, and A. H. Kidd.** 1986. Atypical rotavirus from South African neonates. Arch. Virol. **87:**327–330.
3. **Broome, R. L., P. T. Vo, R. L. Ward, H. F. Clark, and H. B. Greenberg.** 1993. Murine rotavirus genes encoding outer capsid proteins VP4 and VP7 are not major determinants of host range restriction and virulence. J. Virol. **67:**2448–2455.
4. **Clarke, I. N., and M. A. McCrae.** 1981. A rapid and sensitive method for analysing the genome profiles of field isolates of rotavirus. J. Virol. Methods **2:**203–209.

5. Clarke, I. N., and M. A. McCrae. 1982. Structural analysis of electrophoretic variation in the genome profiles of field isolates of rotavirus. Infect. Immun. **36:**392–397.

6. Cukor, G., and N. R. Blacklow. 1984. Human viral gastroenteritis. Microbiol. Rev. **48:**157–179.

7. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.

8. Eaton, B. T., and A. R. Gould. 1987. Isolation and characterisation of orbivirus genotypic variants. Virus Res. **6:**363–382.

9. Eiden, J., G. A. Losonski, J. Johnson, and R. Yolken. 1985. Rotavirus RNA variation during chronic infection of immunocompromised children. Pediatr. Infect. Dis. J. **4:**632–637.

10. Estes, M. K., and J. Cohen. 1989. Rotavirus gene structure and function. Microbiol. Rev. **53:**410–449.

11. Flewett, T. H., and G. N. Woode. 1979. Rotaviruses. Arch. Virol. **57:**1–25.

12. Gorziglia, M., K. Nishikawa, and N. Fukuhara. 1989. Evidence of duplication and deletion in super short segment 11 of rabbit rotavirus Alabama strain. Virology **170:**587–590.

13. Hundley, F., B. Biryahwaho, M. Gow, and U. Desselberger. 1985. Genome rearrangements of bovine rotavirus after serial passage at high multiplicity of infection. Virology **143:**88–103.

14. Hundley, F., M. McIntyre, B. Clark, G. Beards, D. Wood, I. Chrystie, and U. Desselberger. 1987. Heterogeneity of genome rearrangements in rotaviruses isolated from a chronically infected immunodeficient child. J. Virol. **61:**3365–3372.

14a.Isa, P. Personal communication.

15. Johnson, M. A., and M. A. McCrae. 1989. Molecular biology of rotaviruses. VIII. Quantitative analysis of regulation of gene expression during virus replication. J. Virol. **63:**2048–2055.

16. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

17. Matsui, S. M., E. R. Mackow, S. Matsuno, P. S. Paul, and H. B. Greenberg. 1990. Sequence analysis of gene 11 equivalents from "short" and "super short" strains of rotavirus. J. Virol. **64:**120–124.

18. Mattion, N., S. A. Gonzalez, O. Burrone, R. Bellinzoni, J. L. La Torre, and E. A. Scodeller. 1988. Rearrangements of genomic segment 11 in two swine rotavirus strains. J. Gen. Virol. **69:**695–698.

19. McCrae, M. A. 1987. Nucleic acid based analysis on non-group A rotaviruses. CIBA Symp. **128:**22–48.

20. McCrae, M. A., and G. P. Faulkner-Valle. 1981. Molecular biology of rotaviruses. I. Characterization of basic growth parameters and pattern of macromolecular synthesis. J. Virol. **39:**490–496.

21. McCrae, M. A., and J. G. McCorquodale. 1982. Molecular biology of rotaviruses. IV. Molecular cloning of the bovine rotavirus genome. J. Virol. **44:**1076–1079.

22. McCrae, M. A., and J. G. McCorquodale. 1987. Expression of a major bovine neutralisation antigen (VP7c) in *Escherichia coli*. Gene **55:**9–18.

23. Mendez, E., C. F. Arias, and S. Lopez. 1992. Genomic rearrangements in human rotavirus strain Wa: analysis of rearranged RNA segment 7. Arch. Virol. **125:**331–338.

24. Murphy, B. R., and R. G. Webster. 1990. Orthomyxoviruses, p. 1091–1152. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology, 2nd ed. Raven Press, New York.

24a.Patton, J. T., et al. Personal communication.

25. Pedley, S., J. C. Bridger, J. F. Brown, and M. A. McCrae. 1983. Molecular characterisation of rotaviruses with distinct group antigens. J. Gen. Virol. **64:**2093–2101.

26. Pedley, S., J. C. Bridger, D. Chasey, and M. A. McCrae. 1986. Molecular definition of two new groups of atypical rotaviruses. J. Gen. Virol. **67:**131–137.

27. Pedley, S., F. Hundley, I. Chrystie, M. A. McCrae, and U. Desselberger. 1984. The genomes of rotaviruses isolated from chronically infected immunodeficient children. J. Gen. Virol. **65:**1141–1150.

28. Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity *in-vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113:**237–251.

29. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.

30. Scott, G. S., O. Tarlow, and M. A. McCrae. 1989. Detailed structural analysis of a genome rearrangement in bovine rotavirus. Virus Res. **14:**119–128.

31. Thouless, M. E. 1979. Rotavirus polypeptides. J. Gen. Virol. **44:**187–197.

32. Watson, R. J., J. H. Weis, J. S. Salstrom, and L. W. Enquist. 1982. Herpes simplex virus type 1 glycoprotein D gene: nucleotide sequence and expression in *E. coli*. Science **218:**381–384.

33. Webster, R. G., C. H. Campbell, and A. Granoff. 1971. The "in-vivo" production of "new" influenza A viruses. Virology **44:**317–328.

33a.Xu, L., et al. Unpublished data.

34. Xu, L., D. Harbour, and M. A. McCrae. 1990. The application of the polymerase chain reaction to the detection of rotaviruses in faeces. J. Virol. Methods **27:**29–38.

35. Xu, L., D. Harbour, and M. A. McCrae. 1991. Sequence of the gene encoding the major neutralisation antigen (VP7) of serotype 10 rotavirus. J. Gen. Virol. **72:**177–180.