The nucleotide sequence of the 5' and 3' ends of rotavirus SA11 gene 4

Susana López* and Carlos F. Arias

Departamento de Biología Molecular, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Mexico City 04510, México
Submitted April 27, 1987

Rotaviruses, important pathogens of infantile gastroenteritis, have a genome composed of 11 dsRNA segments contained in a double-layered capsid. One of the surface proteins, an 88 Kd protease sensitive protein (VP3), encoded by dsRNA segment 4, has been reported to be the hemagglutinin, and its cleavage is associated with enhanced infectivity and the ability of the virus to replicate in tissue culture.

The partial nucleotide sequence of rotavirus SA11 gene 4 has been previously reported (1). The missing sequences at the 5' and 3' ends of the gene were obtained by primer extension using synthetic oligonucleotides. The 5' end sequence was determined using an oligonucleotide complementary to nucleotides 85 to 100 of the positive strand of the gene, and total mRNA transcribed in vitro as template. To obtain the 3' end sequence, an oligonucleotide representing nucleotides 756 to 775 of the gene 4 coding strand was used, with total genomic dsRNA as template. The following sequence was obtained:

5' 1 GGCTATAAA ATG GCT GCA ... ...

1 M A A

... A TTG CGT GAA TTT ATT AAT CAG GAT AAT CCA ATA ATA CGA AAT AGA

V L R E F I N Q D N P I I R N R

ATA GAA AGT TTG ATA ATG CAA TGT CGC TTA TAA GCAACTGACAGGAGTGACC 2362 3'

I E S L I M Q C R L X 776

This information, together with that previously reported, shows that the entire gene is 2362 nt long containing an uninterrupted sequence of 2328 nt which specifies a protein of 776 amino acids, with a MW of 86,733 d. The open reading frame starts at nucleotide 10 and ends at nucleotide 2337, leaving 5'- and 3'-noncoding sequences of 9 and 25 nt, respectively. The two trypsin cleavage sites of VP3, associated with the enhancement of rotavirus infectivity, are at amino acid positions 241 and 246. Taking as reference the first cleavage site, the two trypsin cleavage products, VP8 and VP5, are 241 (MW 26,787) and 535 (MW 59,964) amino acids long, respectively.

Interestingly, the deduced amino acid sequence at the NH2 terminus of VP3 does not have the characteristics of a signal peptide, suggesting that this protein is synthesized in free ribosomes in the cytoplasm, where it interacts with the newly assembled subviral particles during the morphogenesis of the virion, in agreement with the localization of this protein by ultrastructural label with colloidal gold (2).

*To whom correspondence should be addressed

References: 1. López, S. et al. (1985) Virology 144: 11-19.
2. Petrie, B.L. et al. (1984) Virus Res. 1: 133-152.