Sequence analysis and structural implications of rotavirus capsid proteins

N. PARBHOO, J. B. DEWAR, S. GILDENHUYS*

Department of Life and Consumer Sciences, College of Agriculture and Environmental Sciences, University of South Africa, Private Bag X6, Florida, Roodepoort, South Africa, 1710

Received September 28, 2015; revised November 20, 2015; accepted June 23, 2016

Summary. - Rotavirus is the major cause of severe virus-associated gastroenteritis worldwide in children aged 5 and younger. Many children lose their lives annually due to this infection and the impact is particularly pronounced in developing countries. The mature rotavirus is a non-enveloped triple-layered nucleocapsid containing 11 double stranded RNA segments. Here a global view on the sequence and structure of the three main capsid proteins, VP2, VP6 and VP7 is shown by generating a consensus sequence for each of these rotavirus proteins, for each species obtained from published data of representative rotavirus genotypes from across the world and across species. Degree of conservation between species was represented on homology models for each of the proteins. VP7 shows the highest level of variation with 14-45 amino acids showing conservation of less than 60%. These changes are localised to the outer surface alluding to a possible mechanism in evading the immune system. The middle layer, VP6 shows lower variability with only 14-32 sites having lower than 70% conservation. The inner structural layer made up of VP2 showed the lowest variability with only 1-16 sites having less than 70% conservation across species. The results correlate with each protein's multiple structural roles in the infection cycle. Thus, although the nucleotide sequences vary due to the error-prone nature of replication and lack of proof reading, the corresponding amino acid sequence of VP2, 6 and 7 remain relatively conserved. Benefits of this knowledge about the conservation include the ability to target proteins at sites that cannot undergo mutational changes without influencing viral fitness; as well as possibility to study systems that are highly evolved for structure and function in order to determine how to generate and manipulate such systems for use in various biotechnological applications.

Keywords: rotavirus; capsid protein; amino acid sequence conservation; protein structure; consensus; vaccine candidate

Introduction

Rotavirus infects many species from mammals to birds resulting in gastroenteritis, with group A rotavirus predominantly affecting humans. Rotavirus is the cause of ~450 000 deaths annually worldwide (Parashar *et al.*, 2003, 2009; Tate *et al.*, 2012) with the highest mortality occurring in the developing world (Tate *et al.*, 2012). Rotavirus infection has also been implicated in the onset of type I diabetes in infants (Honeyman *et al.*, 2000). Surveillance of circulat-

Abbreviations: VP = viral protein; PDB = Protein Data Bank

ing rotavirus strains report a continuous change in human strains highlighting the need for continuous characterization of new strains (Collins *et al.*, 2015) and occurrence of coinfection, re-assortment as well as interspecies transmission (Jere *et al.*, 2001; Papp *et al.*, 2014) from porcine to human (Nagai *et al.*, 2015; Nyaga *et al.*, 2015) or bovine to human (Nyaga *et al.*, 2015; Nemoto *et al.*, 2015). Nyaga *et al.* (2015) also showed a clear evolutionary relationship between porcine, bovine and ovine rotavirus sequences, indicating that relatively recent interspecies transmission and re-assortment has taken place. Such reports raise concerns as these changes in circulating rotaviruses may have implications on rotavirus vaccine efficacy. The currently administered live rotavirus vaccine, Rotarix* was developed using a strain isolated more than 20 years ago (Ciarlet and Schödel, 2009; Ward and Bern-

^{*}Corresponding author. E-mail: gildes@unisa.ac.za; phone:+27114713295.

stein, 2009; Plosker, 2010). It has been effective in lowering hospital admissions of diarrhoeal patients (Plosker, 2010; Wang et al., 2010) but its efficacy is lower in the developing world, including Africa and Asia (Molbak et al., 2000; Madhi et al., 2010; Cunliffe et al., 2012; Sow et al., 2012; Zaman et al., 2010). Therefore, solutions need to be sought for future vaccines and treatment strategies. Currently, treatment is limited to symptomatic relief and oral rehydration to prevent dehydration and maintenance of the correct fluid and electrolyte balance. It must be mentioned that there are studies where treatment of the virus with antivirals such as nitazoxanide (Rossignol et al., 2006) and immunoglobulin (Sarker et al., 1998) has been tested and showed to be effective. The treatment with nitazoxanide was based on its ability to inhibit a broad range of viruses (Rossignol et al., 2006) and is, therefore, not a targeted drug designed to act on specific rotavirus targets. More recently, however, novel pyrrole and pyrrolopyrimidine compounds have been synthesized and showed anti-viral activity but have yet to be clinically tested (Mohamed et al., 2015).

Rotavirus belongs to the Reoviridae family and its genome comprises 11 segments of double stranded RNA (dsRNA) encoding six structural and six non-structural proteins. The genome is enclosed in a capsid formed by three layers of proteins (Estes and Cohen, 1989). Each dsRNA segment encodes a single protein, apart from gene 11 which encodes two proteins. The inner capsid layer of the virion encapsulates the genome and is mainly comprised of VP2. The trimeric VP6 protein makes up the middle layer of the virion and is used to classify rotavirus subgroups. The outer layer of the virion is made up of VP4, the spike protein that binds to the host cell and, particularly, VP7, the glycoprotein that determines the serotype of rotavirus (Estes and Cohen, 1989). As described by Boyle and Holmes (1986), the structural organization of rotaviruses means that the inner leaflet of the VP2 capsid protein interfaces mainly with viral gene segments while its outer surface associates with the inner surface of the intermediate VP6 capsid layer. The outer surface of this intermediate capsid layer, in turn, associates with the inner surface of the outer capsid, VP7 protein layer. A sequenced-based classification system is now used for rotaviruses (Matthijnssens et al., 2008, 2011).

A plethora of rotavirus sequences have been used in order to identify conserved regions in the rotavirus proteins within and across species. These sites of variability are then analyzed according to their effect on protein structure by overlaying sequence information onto the structure for each protein to determine the role they may play in the molecular function of the virus. Implications are discussed for future drug design, vaccine development and molecular research involving structure and function of the capsid system in infection and biotechnological applications. Although knowledge of genomic sequence data for current and future rotavirus strains is crucial for subtyping and understanding the morphogenesis of the virus, it is equally important to determine to what extent the virus protein products are affected by such genome changes and what the implications may be of such protein variability.

Materials and Methods

Sequence retrieval. Amino acid sequences of human, cow, pig and horse rotavirus VP2, VP6 and VP7 proteins were retrieved from the NCBI database in FASTA format using the search string "VPx AND rotavirus A AND (organism) AND (amino acid full length)[sequence length]".

Consensus sequence deduction. The consensus sequence and conservation score for each protein from each species was determined using CLC Genomics Workbench software. The FASTA format for each species and capsid protein was imported into CLC Genomics Workbench software. The alignment score and consensus sequences were generated using the "create alignment" tool.

Homology model generation and sequence conservation mapping. Homology models of each of the proteins were constructed using the homology builder in the Accelrys Discovery Studio v3.5 suite of tools using PDB codes: 3KZ4 (VP2); 1QHD (VP6) and 3FMG (VP7) as the model templates. An alignment variability score was assigned using ProtSkin (Ritter *et al.*, 2004). Structure renderings according to conservation were performed using PyMOL (The PyMOL Molecular Graphics System, Version 0.99rc6 Schrödinger, LLC).

Results

Rotavirus sequence data indicate changes occurring over time within circulating viruses. We set out to determine the extent to which these sequence changes have impact on the amino acid level of the major capsid proteins constituting the triple layered particle, especially VP2, VP6 and VP7 across all strains of group A rotaviruses. Rather than focusing on a particular species of rotavirus or from a specific region, the dataset encompasses all full-length protein sequences currently available on the NCBI database for human as well as for some livestock (pig, horse and cow) for group A rotaviruses, allowing us to dissect potential effects of interspecies zoonosis and reassortment if they were to occur. Fig. 1 indicates comparative amino acid consensus sequences for the three rotavirus capsid proteins, VP2, VP6 and VP7. The number of data sequences for each capsid protein for each of host organisms is indicated in Table 1. There is no clear evidence that large segments or stretches of amino acids readily interchange as would be expected if reassortment had taken place. It is interesting to note that the amino acid substitutions at the variable regions for each protein

(a)		20		40		60		80	
VP2HumanCons	MAYRKRGARR	EANINNNDRM	QEKDDEKQDQ	NNKMQLSDKV	LSKKEEVVTD	SQEEIKITDE	VKKSTKEESK	QLLEVLKTKE	80
VP2BovineCons.	MAYRKRGARR	EANINNNDRM	QEKDDEKQDQ	NNRMQLSDKV	LSKKEEVVTD	SQEEIKIADE	VKKSTKEESK	QLLEVLKTKE	80
VP2PorcineCons. VP2EquineCons.	MAYRKRGARR	EVNVNNDRM	QEKDDEKQDQ	NNKMQLSDKV NNKMQLSDKV	LSKKEEVITD	SOEKIEIVDE	VKKSTKEESK	QLLEVLKTKE	80 80
Consensus	MAYRKRGARR	EANINNDRM	QEKDDEKQDQ	NNKMQLSDKV	LSKKEEVXTD	SQEEIKIADE	VKKSTKEESK	QLLEVLKTKE	~
		100		120		140		160	
VP2HumanCons	EHQKEIQYEI	LQKTIPTFEP	KESILKKLED	IKPEQAKKQT	KLFRIFEPRQ	LPIYRANGEK	ELRNRWYWKL	KKDTLPDGDY	160
VP2BovineCons.	EHQKEIQYEI	LQKTIPTFEP	KESILKKLED	IKPEQAKKQT	KLFRIFEPRQ	LPIYRANGEK	ELRNRWYWKL	KKDTLPDGDY	160
VP2PorcineCons. VP2EquineCons.	EHOKELOYEL	LOKTIPTEP	KESILKKLED	IKPEQAKKQT	KLFRIFEPRQ	LPIYRANGEK	ELRNRWYWKL	KKDTLPDGDY	160
Consensus	EHQKEIQYEI	LQKTIPTFEP	KESILKKLED	IKPEQAKKQT	KLFRIFEPRQ	LPIYRANGEK	ELRNRWYWKL	KKDTLPDGDY	
		180		200		220		240	
VP2HumanCons	DVREYFLNLY	DQVLTEMPDY	LLLKDMAVEN	KNSRDAGKVV	DSETASICDA	IFQDEETEGA	VRRFIAEMRQ	RVQADRNVVN	240
VP2BovineCons.	DVREYFLNLY	DQVLTEMPDY	LLLKDMAVEN	KNSRDAGKVV	DSETASICDA	I FODEETEGA	VRRFIAEMRQ	RVQADRNVVN	240
VP2EquineCons.	DVREYFLNLY	DQVLTEMPDY	LLLKDMAVEN	KNSRDAGKVV	DSETASICDA	IFQDEETEGA	VRRFIAEMRQ	HVQADRNVVN	240
Consensus	DVREYFLNLY	DQVLTEMPDY	LLLKDMAVEN	KNSRDAGKVV	DSETASICDA	IFQDEETEGA	VRRFIAEMRQ	RVQADRNVVN	
		260		280		300		320	
VP2HumanCons	YPSILHPIDY	AFNEYFLQHQ	LVEPLNNDII	FNYIPERIRN	DVNYILNMDR	NLPSTARYIR	PNLLQDRLNL	HDNFESLWDT	320
VP2BovineCons.	YPSILHPIDY	AFNEYFLQHQ	LVEPLNNDII	FNYIPERIRN	DVNYILNMDR	NLPSTARYIR	PNLLQDRLNL	HDNFESLWDT	320
VP2EquineCons.	YPSILHPIDY	AFNEYFLQHQ	LVEPLNNDII	FNYIPERIRN	DVNYILNMDR	NLPSTARYIR	PNLLQDRLNL	HDNFESLWDT	320
Consensus	YPSILHPIDY	AFNEYFLQHQ	LVEPLNNDII	FNYIPERIRN	DVNYILNMDR	NLPSTARYIR	PNLLQDRLNL	HDNFESLWDT	
		340		360		380		400	
VP2HumanCons	ITTSNYILAR	SVVPDLKELV	STEAQIQKMS	QDLQLEALTI	QSETQFLTGI	NSQAANDCFK	TLIAAMLSQR	TMSLDFVTTN	400
VP2BovineCons. VP2PorcineCons.	ITTSNYILAR	SVVPDLKELV	STEAQIQKMS	QDLQLEALTI ODLQLEALTI	QSETQFLTGI OSETOFLTGI	NSQAANDCFK	TLIAAMLSQR	TMSLDFVTTN	400
VP2EquineCons.	ITTSNYILAR	SVVPDLKELV	STEAQIQKMS	QDLQLEALTI	QSETQFLTGI	NSQAANDCFK	TLIAAMLSQR	TMSLDFVTTN	400
Consensus	ITTSNYILAR	SVVPDLKELV	STEAQIQKMS	QDLQLEALTI	QSETQFLTGI	NSQAANDCFK	TLIAAMLSQR	TMSLDFVTTN	
		420		440		450		480	
VP2HumanCons	YMSLISGMWL	LTVVPNDMFI	RESLVACQLA	IVNTITYPAF	GMQRMHYRNG	DPQTPFQIAE	QQIQNFQVAN	WLHFVNNNQF	480
VP2BovineCons. VP2PorcineCons.	YMSLISGMWL	LTVVPNDMFI	RESLVACQLA	IVNTIVYPAF	GMORMHYRNG	DPQTPFQIAE	QQIQNFQVAN	WLHFVNNNQF	480
VP2EquineCons.	YMSLISGMWL	LTVTPNDMFI	RESLVACQLA	IVNTIVYPAF	GMQRMHYRNG	DPQTPFQIAE	QQIQNFQVAN	WLHFVNNNQF	480
Consensus	YMSLISGMWL	LTVVPNDMFI	RESLVACQLA	IVNTIXYPAF	GMQRMHYRNG	DPQTPFQIAE	QQIQNFQVAN	WLHFVNNNQF	
		500 I		520 I		540 I		560	
VP2HumanCons VP2RovineCoort	ROVVIDGVLN	QVLNDNIRNG	HVINQLMEAL	MQLSRQQFPT	MPVDYKRSIQ	RGILLLSNRL	GQLVDLTRLL	AYNYETLMAC	560
VP2PorcineCons.	RQVVIDGVLN	QVLNDNIRNG	HVINQLMEAL	MQLSRQQFPT	MPVDYKRSIQ	RGILLLSNRL	GQLVDLTRLL	AYNYETLMAC	560
VP2EquineCons.	RQVVIDGVLN	QVLNDNIRNG	HVINQLMEAL	MQLSRQQFPT	MPVDYKRSIQ	RGILLLSNRL	GQLVDLTRLL	AYNYETLMAC	560
Consensus	RQVVIDGVLN	QVLNDNIRNG	HVINQLMEAL	MQLSRQQFPT	MPVDYKRSIQ	RGILLLSNRL	GQLVDLTRLL	AYNYETLMAC	
10011	TTUNNOUVOT	T		ATVIDODOTI	EUVVNUNUE	I			~ * *
VP2HumanCons VP2BovineCons.	VTMNMQHVQT	LTTEKLQLTS	VTSLCMLIGN	ATVIPSPOTL	FHYYNVNVNF	HSNYNERIND	AVAIITAANR	LNLYQKKMKA	640
VP2PorcineCons.	TMNMQHVQT	LTTEKLQLTS	VTSLCMLIGN	ATVIPSPQTL	FHYYNVNVNF	HSNYNERIND	AVAIITAANR	LNLYQKKMKA	640
VP2EquineCons.	VTMNMQHVQT	LTTEKLQLTS	VTSLCMLIGN	ATVIPSPQTL	FHYYNVNVNF	HSNYNERIND	AVAIITAANR	LNLYQKKMKT	640
Consensus	XIMNMQHVQI	660	VISLOMLIGN	ATVIPSPULL 680	FHTTNVNVNF	TON THER IND	AVAIIIAANK	LNLTQKKMKA 720	
VP2HumanCons			OMYRLEDELE	LLPVEVERLD			OGVILAYEDM	OLERDEMYGY	720
VP2BovineCons.	IVEDFLKRLH	IFDVARVPDD	QMYRLRDRLR	LLPVEVRRLD	IFNLILMNMD	QIERASDKIA	QGVIIAYRDM	QLERDEMYGY	720
VP2PorcineCons.	IVEDFLKRLH	I FDVARVPDD	QMYRLRDRLR	LLPVEVRRLD	I FNL I LMNMD	QIERASDKIA	QGVIIAYRDM	QLERDEMYGY	720
Consensus	IVEDFLKRLT	FOVARVPDD	OMYRLEDRLE	LLPVEVRRLD		OLERASDKIA	OGVIIAYRDM	OLERDEMYGY	120
001301303	TTED ERRER	740	GanticeRonen	760		780	dorringing	800	
VP2HumanCons	VNIARNLDGF	QQINLEELMR	TGDYAQITNM	LLNNQPVALV	GALPFITDSS	VISLIAKLDA	TVFAQIVKLR	KVDTLKPILY	800
VP2BovineCons.	VNIARNLDGF	QQINLEELMR	TGDYAQITNM	LLNNQPVALV	GALPFVTDSS	VISLIAKLDA	TVFAQIVKLR	KVDTLKPILY	800
VP2PorcineCons. VP2EquineCons	VNIARNLDGF	QQINLEELMR	TGDYAQI TNM	LLNNQPVALV	GALPETTOSS	VISLIAKLDA	TVFAQIVKLR	KVDTLKPILY	800
Consensus	VNIARNLDGF	QQINLEELMR	TGDYAQITNM	LLNNQPVALV	GALPFITDSS	VISLIAKLDA	TVFAQIVKLR	KVDTLKPILY	000
		820		840		860		880	
VP2HumanCons	KINSDSNDFY	LVANYDWVPT	STTKVYKQVP	QQFDFRNSMH	MLTSNLTFTV	YSDLLAFVSA	DTVEPINAVA	FONMRIMNEL	880
VP2BovineCons.	KINSDSNDFY	LVANYDWVPT	STTKVYKQVP	QQFDFRNSMH	MLTSNLTFTV	YSDLLAFVSA	DTVEPINAVA	FDNMRIMNEL	880
VP2EquineCons.	KINSDSNDFY	LVANYDWVPT	STTKVYKQIP	QQFDFRNSMH	MLTSNLTFTV	YSDLLAFVSA	DTVEPINAVA	FDNMRIMNEL	880
Consensus	KINSDSNDFY	LVANYDWVPT	STTKVYKQVP	QQFDFRNSMH	MLTSNLTFTV	YSDLLAFVSA	DTVEPINAVA	FDNMRIMNEL	

Fig. 1

are conservative, i.e. the same type of residue that usually undergoes substitution and replacement is synonymous, i.e. a polar residue is not replaced by a hydrophobic residue (except in cases where residues are substituted by glycine). Even when examining the sequences chronologically, no distinct pattern of variation can be found. Structural homology models were then constructed using the homology builder in the Accelrys Discovery Studio 3.5 suite of tools for each of the major capsid proteins using the consensus sequence along with PDB co-ordinates 3KZ4, 1QHD and 3FMG which provided the structure template for VP2, VP6 and VP7, respectively. The homology models

(b)		20		40		60		80	
VP6HumanCons.	MEVLYSLSKT	LKDARDKIVE	GTLYSNVSDL	IQQFNQMIVT	MNGNDFQTGG	IGNLPIRNWT	FDFGLLGTTL	LNLDANYVET	80
VP6PorcineCons.	MEVLYSLSKT	LKDARDKIVE	GTLYSNVSDL	IQQFNQMIVT	MNGNDFQTGG	IGNLPIRNWT	FDFGLLGTTL	LNLDANYVEN	80
VP6EquineCons. VP6BowineCons	MEVLYSISKT	LKDARDKIVE	GTLYSNVSDI		MNGNEFQTGG		FDFGLLGTTL	LNLDANYVET	80
Consensus	MEVLYSLSKT	LKDARDKIVE	GTLYSNVSDL	LOOFNOMIVE	MNGNXFOTGG	IGNLPIRNWT	FDFGLLGTTL	LNLDANYVET	00
00113011303		100	0121001002	120		140		160	
VP6HumanCons.	ARTTIEYFID	FIDNVCMDEM	ARESQRNGVA	PQSEALRKLA	GIKFKRINFD	NSSEYIENWN	LQNRRQRTGF	VEHKPNIEPY	160
VP6PorcineCons.	ARTTIEYFID	FIDNVCMDEM	ARESQRNGIA	PQSEALRKLS	GIKFKRINFD	NSSEYIENWN	LQNRRQRTGF	VFHKPNILPY	160
VP6EquineCons.	ARTTIEYFID	FIDNVCMDEM	TRESORNGIA	POSDALRKLS	GIKEKRINED	NSSEYIENWN	LONRRORTGE	VEHKPNIEPY	160
Consensus	ARTTIEVELD	FIDNVCMDEM	ARESORNOLA	POSYALPKIS	GIKEKPINED	NESEVIENWN	LONPROPTOF	VEHKONLEDY	100
Consensus	ARTTETFTD	180	ARESQUINTA	200	GIRFRAINFD	220	EQUINIQUIOF	240	
VP6HumanCons.	SASFTLNRSQ	PMHDNLMGTM	WLNAGSEIQV	AGEDYSCALN	APANIQQFEH	IVQLRRALTT	ATITLPDAE	RFSFPRVINS	240
VP6PorcineCons.	SASFTLNRSQ	PAHDNLMGTM	WXNAGSEIQV	AGFDYSCALN	APANIQQFEH	IVPLRRALTT	ATITLLPDAE	RFSFPRVINS	240
VP6EquineCons.	SASFTLNRSQ	PLHNDLMGTM	WLNAGSEIQV	AGFDYSCAFN	APANTQQFEH	IVQLRRALTT	ATITILPDAE	RFSFPRVINS	240
VP6BovineCons.	SASFTLNRSQ	PAHDNLMGTM	WLNAGSEIQV	AGFDYSCALN	APANTQQFEH	IVQLRRVLTT	ATITLLPDAE	RESEPRVINS	240
Consensus	SASFTLNRSQ	PAHDNLMGTM	WLNAGSEIQV	AGFDYSCAIN	APANXQQFEH	IVQLRRALTT	ATITLLPDAE	RFSFPRVINS	
		I I		I I		Ĩ		I	
VP6HumanCons.	ADGATTWFFN	PVILRPNNVE	VEFLLNGQII	NTYQARFGTI	XARNFDTIRL	SEQUMRPPNM	TPAVNALEPO	ADDELEMATY	320
VP6EquineCons.	ADGATTWYFN	PVILRPNNVE	VEFLLNGQII	NTYQARFGTI	ARNEDTIRL	SFOLMRPPNM	TPAVNALEPO	AOPFOHHATY	320
VP6BovineCons.	ADGATTWYFN	PVILRPNNVE	VEFLLNGQII	NTYQARFGTI	ARNFDTIRL	SFQLMRPPNM	TPAVAALFPN	AQPFEHHATV	320
Consensus	ADGATTWYFN	PVILRPNNVE	VEFLLNGQII	NTYQARFGTI	IARNFDTIRL	SFQLMRPPNM	TPAVXALFPQ	AQPFQHHATV	
		340		360		380 I			
VP6HumanCons.	GLTLRIESAV	CESVLADANE	TLLANVTAVR	QEYAIPVGPV	FPPGMNWTEL	ITNYSPSRED	NLQRVFTVAS	IRSMLIK 397	
VP6PorcineCons.	GLTLRIESAV	CESVLADASE	TLLANVTSVR	QEYATPVGPV	FPPGMNWTEL	ITNYSPSRED	NLORVETVAS	IRSMLIK 397	
VP6EquineCons. VP6BovineCons.	GLTLRIESAV	CESVLADASE	TMLANVTSVR	OEYALPVGPV	FPPGMNWTEL	ITNYSPSRED	NLORVETVAS	IRSMLVK 397	
Consensus	GLTLRIESAV	CESVLADAXE	TXLANVTXVR	QEYAIPVGPV	FPPGMNWTEL	ITNYSPSRED	NLQRVFTVAS	IRSMLIK	
(c)		20		40		60		80	
(C)	MYGIEYTTVI		VIIKSVTRTM				GRMDTPYANS		80
(C) VP7HumanCons. VP7PorcineCons.	MYGIEYTTVL Mygieyttvl	20 1 TFLISLVLLN TFLISLVLLN	YILKSVTRTM YILKSVTRTM	40 I DFIIYRFLLV DFIIYRFLLV	IVILAPLIKA IVILAPLIKA	ONYGINLPIT QNYGINLPIT	GSMDTPYANS GSMDTPYANS	TOSETFLTST	80 80
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons.	MYGIEYTTVL Mygieyttvl Mygieyttvl	20 I TFLISLVLLN TFLISLVLLN TFLISII	YILKSVTRTM YILKSVTRTM YILKSLTRMM	DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI	IVILAPLIKA IVILAPLIKA IVILSPLLKA	ONYGINLPIT ONYGINLPIT ONYGINLPIT	GSMDTPYANS GSMDTPYANS GSMDTAYANS	TQSETFLTST TQSETFLTST TQEETFLTST	80 80 80
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL	20 I TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISIILLN	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM	DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST	80 80 80 80
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. Consensus	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL	20 1 TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISXXLLN	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSXTRXM	DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLXX	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TQXETFLTST	80 80 80 80
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. Consensus	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL	20 TFLISLVLIN TFLISLVLIN TFLISIILIN TFLISIILIN TFLISXXLLN 100	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM	40 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLXX 120 I	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLXKA	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT 1400 	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOEETFLTST TOXETFLTST	80 80 80 80
(C) VP7HumanCons. VP7PorcineCons. VP7BovineCons. VP7BovineCons. Consensus VP7HumanCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA	20 I TFLISLVLLN TFLISLLLN TFLISILLN TFLISXXLLN 10 I TEIADTEWKD	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG	40 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLXX 120 WPTGSVYFKE	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT QNYGINLPIT POLYCDYN	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLOL	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOXETFLTST 100 1 DMSELADLIL	80 80 80 80
(C) VP7PorcineCons. VP7EquineCons. VP7BovineCons. Consensus VP7HumanCons. VP7PorcineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA	TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISIILLN TFLISXXLLN TFLIADTEWKD TEIADTEWKD	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG	40 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLXX WPTGSVYFKE WPTGSVYFKE	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YADIASFSVD	QNYGINLPIT ONYGINLPIT QNYGINLPIT QNYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNIV	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLOL LMKYDSNLOL	TOSETFLTST TOSETFLTST TOEETFLTST TQXETFLTST TQXETFLTST DMSELADLIL DMSELADLIL	80 80 80 80 160
(C) VP7PorcineCons. VP7PorcineCons. VP7BovineCons. Consensus VP7HumanCons. VP7PorcineCons. VP7PorcineCons. VP7BovineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA	20 TFLISLVLLN TFLISILLN TFLISILLN TFLISILLN TFLISXLLN I TEIADTEWKD TEINDNSWKD	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSXTRXM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG	40 I V DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLXX 120 WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YADIASFSVD YTDIASFSVD	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT POLYCDYN POLYCDYN POLYCDYNYY POLYCDYNYY	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL	TOSETFLTST TOSETFLTST TOEETFLTST TOZETFLTST TOXETFLTST DMSELADLIL DMSELADLIL DMSELADLIL	80 80 80 160 160 160
(C) VP7PorcineCons. VP7EquineCons. VP7BovineCons. Consensus VP7HumanCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. Consensus	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA	TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISIILLN TFLISXILLN TFLIADTEWKD TEIADTEWKD TEINDNSWKD TEIXDXXWKD	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG	40 1 DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYN POLYCDYNVY POLYCDYNVY POLYCDYNVY	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLOL LMKYDSNLOL LMKYNETLOL LMKYNETLOL LMKYXXLQL	TOSETFLTST TOSETFLTST TOEETFLTST TOXETFLTST I DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL	80 80 80 160 160 160
(C) VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7BovineCons. Consensus VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. Consensus	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA	20 TFLISLVLLN TFLISIULLN TFLISIILLN TFLISILLN TFLISXLLN 10 10 10 10 10 10 10 10 10 10	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG	40 DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE 200	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD	CONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYN POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOEETFLTST DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL 240	80 80 80 160 160
(C) VP7PorcineCons. VP7PorcineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7BovineCons. VP7HumanCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA	20 TFLISLVLLN TFLISIILLN TFLISIILLN TFLISIILLN TFLISXLLN 10 10 11 11 10 11 10 11 11 10 11 10 11 11	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS	40 DFIIYRFLLY DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE 200 1 CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD OTLGIGCSTT	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYN POLYCDYN POLYCDYN POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY POLYCDYNVY	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL LMKYXXXLQL	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST IOKSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL 240 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	80 80 80 160 160 160 240
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI	TFLISLVLLN TFLISIILLN TFLISIILN TFLISIILN TFLISXXLLN TFLIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXXWKD TEIXDXXWKD TEIXDXXWKD TEIXDXXWKD	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGXS	DFIIYRFLLY DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI DFIIYRFLX 120 WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD QTLGIGCSTT GTLGIGCSTT	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV DINTFETVAT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLOL LMKYDSNLOL LMKYNETLOL LMKYNETLOL LMKYNETLOL LMKYXXXLQL AEKLAITDVY AEKLAITDVY	TQSETFLTST TQEETFLTST TQEETFLTST TQEETFLTST DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL	80 80 80 160 160 160 240 240
(C) VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7HumanCons. VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI	20 TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISIILLN TFLISXXLLN I TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXXWKD I TLYYYQQTDE TLYYYQQTDE	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGXS ANKWISMGSS	40 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD GTLGIGCSTT GTLGIGCLTT	ANYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV DTNTFETVAT DTNTFETVAT DVATFETVAT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYXXXLQL AEKLAITDVV AEKLAITDVV	TOSETFLTST TOSETFLTST TOEETFLTST TQXETFLTST I DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL	80 80 80 160 160 160 240 240 240
(C) VP7PorcineCons. VP7PorcineCons. VP7BovineCons. Consensus VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7BovineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI	TFLISLVLLN TFLISLVLLN TFLISILLN TFLISILLN TFLISILLN TFLISXLLN TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXXWKD TEIXDXXWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS	40 1 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE 2000 1 CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD QTLGIGCSTT QTLGIGCLTT QTLGIGCLTT	CONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV DTNTFETVAT DTNTFETVAT DVATFEEVAT DVATFEEVAT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL LMKYXXXLQL AEKLAITDVV AEKLAITDVV AEKLVITDVV	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOXETFLTST DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DGVNHKLDVT DGVNHKLDVT	80 80 80 160 160 160 240 240 240
(C) VP7PorcineCons. VP7PorcineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7BovineCons. VP7BovineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI	20 TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISIILLN TFLISXXLLN 100 TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXSWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGXS ANKWISMGSS ANKWISMGSS	40 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT 200	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD GTLGIGCSTT GTLGIGCLTT GTLGIGCLTT GTLGIGCXTT	0 0 0 0 0 0 0 0 0 0 0 0 0 0	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS LMKYDSNLOL LMKYDSNLOL LMKYNETLOL LMKYNETLOL LMKYNETLOL LMKYXXXLQL AEKLAITDVV AEKLAITDVV AEKLVITDVV AEKLXITDVV	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOXETFLTST IDMSELADLIL DMSELADLIL	80 80 80 160 160 160 240 240 240
(C) VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI	TFLISLVLLN TFLISLVLLN TFLISILLN TFLISILLN TFLISILLN TFLISXXLLN TEIADTEWKD TEIADTEWKD TEINDNSWKD TEIXDNSWKD TEIXDNSWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS ANKWISMGSS	40 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YTDIASFSVD OTLGIGCSTT GTLGIGCLTT GTLGIGCLTT GTLGIGCXTT	ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV DTNTFETVAT DYNTFETVAT DVATFEEVAT DVATFEEVAT DXXTFEXVAT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL LMKYXXXLQL AEKLAITDVV AEKLAITDVV AEKLVITDVV AEKLVITDVV	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOXETFLTST DMSELADLIL DGVNHKLDVT DGVNHKLDVT DGVNHKLDVT	80 80 80 160 160 160 240 240 240
(C) VP7PorcineCons. VP7PorcineCons. VP7BovineCons. Consensus VP7HumanCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI	20 TFLISLVLLN TFLISILLN TFLISILLN TFLISILLN TFLISILLN I TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEINDNSWKD TEINDNSWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS	40 1 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI 1 WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YTDIASFSVD OTLGIGCSTT GTLGIGCLTT GTLGIGCLTT GTLGIGCLTT TERMMRINWK	CNYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYNIV	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTXYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYXXLQL AEKLAITDVV AEKLAITDVV AEKLAITDVV AEKLYITDVV AEKLYITDVV AEKLXITDVV DYVNQIVQVM	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOXETFLTST DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DGVNHKLDVT DGVNHKLDVT DGVNHKLDVT DGVNHKLDVT 330 SKRSSLDSA	80 80 80 160 160 160 240 240 240 240 240 320
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI	20 TFLISLVLLN TFLISLVLLN TFLISILLN TFLISTILLN TFLISTILN TFLISTILN TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXXWKD 10 TLYYYQQTDE	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS	40 1 DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE 200 201 CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YTDIASFSVD YTDIASFSVD GTLGIGCSTT GTLGIGCLTT GTLGIGCLTT GTLGIGCLTT GTLGIGCLTT TERMMRINWK TERMMRINWK	CNYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV DTNTFETVAT DYATFETVAT DYATFEEVAT DYATFEEVAT DYATFEEVAT DYATFEEVAT CYTFEEVAT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLOL LMKYDSNLOL LMKYDSNLOL LMKYNETLOL LMKYNETLOL LMKYXXLQL AEKLAITDVV AEKLVITDVV AEKLVITDVV AEKLXITDVV AEKLXITDVV DYVNOIVOVM DYVNOIVOVM	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOXETFLTST IDMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL SKRSSLDSA SKRSSLDSA SKRSSLDSA	80 80 80 80 160 160 160 240 240 240 240 320 320
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7Equi	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI TNTCTIRNCK TATCTIRNCK TATCTIRNCK	20 I TFLISLVLLN TFLISILLN TFLISILLN TFLISILLN I I TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD I TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE I	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS IQVGGSNILD IQVGGSDILD	40 1 DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YTDIASFSVD GTLGIGCSTT GTLGIGCSTT GTLGIGCLTT GTLGIGCLTT GTLGIGCLTT TERMMRINWK TERMMRINWK	CNYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV CONTFETVAT DYATFETVAT DYATFEEVAT DYATFEEVAT DYATFEEVAT DYATFEEVAT DYATFEEVAT CANFEEVAT DYATFEEVAT CANFEEVAT DYATFEEVAT DYATFEEVAT CANFEEVAT CANFEEVAT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLQL LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYXXLQL AEKLAITDVV AEKLVITDVV AEKLVITDVV AEKLXITDVV AEKLXITDVV DYVNQIVQVM DYVNQIQAM	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST IDMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL SKRSKSLDSA SKRSKSLDSA SKRSKSLNSA SKRSKSLNSA	80 80 80 80 160 160 160 240 240 240 240 320 320 320 320
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7PorcineCons. VP7EquineCons. VP7E	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI TNTCTIRNCK TATCTIRNCK TATCTIRNCK	20 TFLISLVLLN TFLISILLN TFLISILLN TFLISILLN TFLISILLN TFIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXXWKD TLYYYQQTDE TLYYQQTDE TLYYQQTDE TLYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYYQ	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS IQVGGSNILD IQVGGSDILD IQVGGSDILD IQVGGSXILD	40 1 DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE 200 CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IXII IXII IXII IXII IXII IXII IXII I	CNYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT QNYGINLPIT PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV CONTFETVAT DYATFETVAT DYATFETVAT DYATFETVAT DYATFETVAT DYATFETVAT CNYFETVAT CNYFYTVV KWWQVFYTVV KWWQVFYTVV	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL LMKYXXLQL AEKLAITDVV AEKLVITDVV AEKLVITDVV AEKLXITDVV AEKLXITDVV AEKLXITDVV AEKLXITDVV AUVOIVQM DYVNQIIQAM	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST IQXETFLTST DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL SKRSSLADUT DGVNHKLDVT DGVNHKLDVT SKRSSLDSA SKRSRSLNSA SKRSRSLNSA SKRSRSLNSA	80 80 80 160 160 160 240 240 240 240 240 240 320 320 320
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. Consensus	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI TNTCTIRNCK TNTCTIRNCK TATCTIRNCK	20 TFLISLVLLN TFLISLVLLN TFLISILLN TFLISILLN TFLISTILLN TFLISTILN TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXXWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYQQTDE TLYYQQTDE TLYYQQTDE TLYYYYQQTDE TLYYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYYQQTDE TLYYYYQQTDE TLYYYYQQTDE TLYYYYQQTDE TLYYYYQQTDE TLYYYYQ	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS IQVGGSNILD IQVGGSDILD IQVGGSDILD IQVGGSXILD	40 1 DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE 200 CTIKVCPLNT	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IVILXPLXKA IXII IXII IXII IXII IXII IXII IXII I	CNYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT QNYGINLPIT PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV PQLYCDYNIV CONTFETVAT DYATFETVAT DYATFETVAT DYATFETVAT DYATFETVAT DYATFETVAT XWWQVFYTVV KWWQVFYTVV KWWQVFYTVV	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL LMKYXXLQL AEKLAITDVV AEKLVITDVV AEKLVITDVV AEKLXITDVV AEKLXITDVV DYVNQIVQM DYVNQIQAM DYVNQIQAM	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TOXETFLTST DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL SKRSSLADUT DGVNHKLDVT DGVNHKLDVT SKRSSLDSA SKRSRSLDSA SKRSRSLNSA SKRSRSLNSA	80 80 80 160 160 160 240 240 240 240 320 320 320
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7BovineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7BovineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7EquineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7HumanCons. VP7HumanCons. VP7HumanCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI TNTCTIRNCK TATCTIRNCK TATCTIRNCK TATCTIRNCK	20 TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISIILLN TFLISXXLLN TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE XLGPRENVAV KLGPRENVAV KLGPRENVAV	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG SANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS IQVGGSNILD IQVGGSDILD IQVGGSDILD IQVGGSDILD	40 DFIIYRFLLV DFIIYRFLLY DFIIYRFLFI DFIIYRFLXX WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIADPTTAPQ ITADPTTAPQ ITADPTTAPQ	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD OTLGIGCSTT OTLGIGCSTT OTLGIGCLTT QTLGIGCLTT TERMMRINWK TERMMRINWK	ANYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV DTNTFETVAT DVATFEZVAT DVATFEZVAT DVATFEZVAT SWUQVFYTIVV KWWQVFYTIVV KWWQVFYTIVV	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL LMKYXXXLQL AEKLAITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV DYVNQIVQVM DYVNQIVQM DYVNQIIQAM DYVNQIIQAM	IN TOSETFLTST TOETFLTST TOETFLTST TOXETFLTST TOXETFLTST INSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DGVNHKLDVT DGVNHKLDVT DGVNHKLDVT SKRSRSLDSA SKRSRSLNSA SKRSRSLNSA	80 80 80 160 160 160 240 240 240 240 320 320 320
(C) VP7HumanCons. VP7PorcineCons. VP7EquineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7PorcineCons. VP7EquineCons. VP7EquineCons. VP7PorcineCons. VP7EquineCons. VP7PorcineCons. VP7EquineCons. VP7PorcineCons. VP7EquineCons. VP7PorcineCons. VP7EquineCons.	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI TNTCTIRNCK TATCTIRNCK TATCTIRNCK TATCTIRNCK	20 TFLISLVLLN TFLISLVLLN TFLISIILLN TFLISIILN TFLISXXLLN TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE XLGPRENVAV KLGPRENVAV KLGPRENVAV	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGXS ANKWISMGSS ANKWISMGSS ANKWISMGSS IOVGGSNILD IQVGGSNILD IQVGGSXILD	40 DFIIYRFLLV DFIIYRFLLY DFIIYRFLFI DFIIYRFLXX 120 WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIAUPTTAPQ ITADPTTAPQ ITADPTTAPQ	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD YXDIASFSVD GTLGIGCSTT GTLGIGCSTT GTLGIGCLTT GTLGIGCLTT GTLGIGCLTT TERMMRINWK TERMMRINWK	ANYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV POLYCDYNVV DTNTFETVAT DVATFEVAT DVATFEVAT DVATFEVAT DVATFEVAT XXWQVFYTIV KWWQVFYTVV KWWQVFYTVV	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLOL LMKYDSNLOL LMKYNETLOL LMKYNETLOL LMKYNETLOL LMKYNETLOL LMKYXXXLQL AEKLAITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV DVVNQIVQVM DVVNQIVQVM DVVNQIIQAM	IN I TOSETFLTST TOEETFLTST TOEETFLTST TOEETFLTST TOXETFLTST IMMELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL SKRSRSLDSA SKRSRSLDSA SKRSRSLNSA SKRSRSLNSA	80 80 80 160 160 240 240 240 240 240 320 320 320 320
(C) VP7PorcineCons. VP7BorcineCons. V	MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL MYGIEYTTVL LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA LCLYYPTEAA NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI NEWLCNPMDI TNTCTIRNCK TATCTIRNCK TATCTIRNCK TATCTIRNCK TATCTIRNCK	20 TFLISLVLLN TFLISIILLN TFLISIILLN TFLISIILN TFLISXXLLN TEIADTEWKD TEIADTEWKD TEINDNSWKD TEINDNSWKD TEIXDXXWKD TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYYQQTDE TLYYQT	YILKSVTRTM YILKSVTRTM YILKSLTRMM YILKSLTRMM YILKSLTRMM TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG TLSQLFLTKG ANKWISMGSS ANKWISMGSS ANKWISMGSS ANKWISMGSS IQVGGSNILD IQVGGSNILD IQVGGSXILD	40 DFIIYRFLLV DFIIYRFLLV DFIIYRFLFI DFIIYRFLFI WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE WPTGSVYFKE 1 CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIKVCPLNT CTIADPTTAPQ ITADPTTAPQ ITADPTTAPQ	IVILAPLIKA IVILAPLIKA IVILSPLLKA IVILSPLLKA IVILXPLXKA YADIASFSVD YTDIASFSVD YTDIASFSVD YXDIASFSVD OTLGIGCSTT OTLGIGCLTT OTLGIGCLTT OTLGIGCLTT TERMMRINWK TERMMRINWK TERMMRINWK	ANYGINLPIT ONYGINLPIT ONYGINLPIT ONYGINLPIT QNYGINLPIT POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV POLYCDYNIV CONTFETVAT DYNTFETVAT DYNTFETVAT DYNTFETVAT DYNTFETVAT DYNTFETVAT CANFELVAT	GSMDTPYANS GSMDTPYANS GSMDTAYANS GSMDTAYANS GSMDTAYANS LMKYDSNLQL LMKYDSNLQL LMKYNETLQL LMKYNETLQL LMKYNETLQL LMKYXXXLQL AEKLAITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV AEKLVITDVV DVVNQIVQVM DVVNQIVQVM DVVNQIIQAM DVVNQIIQAM	TOSETFLTST TOSETFLTST TOEETFLTST TOEETFLTST TQXETFLTST DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL DMSELADLIL SKRSRSLDSA SKRSRSLDSA SKRSRSLDSA SKRSRSLSA	80 80 80 160 160 160 240 240 240 240 240 320 320 320

Fig. 1

Deduced consensus sequences

All amino acid sequences for rotavirus proteins VP2 (a) VP6 (b) and VP7 (c) were retrieved from the NCBI database. Amino acids highlighted indicate variation between the consensus sequences. The consensus for each species was computed using the alignment tool of the CLC Genomics workbench suite.

for each of the capsid proteins as a complex with one another are shown in Fig. 2 with the degree of conservation rendered for each protein across species represented by a blue-white spectrum where blue represents variation and white represents no variation (Fig. 3). Numerical data are displayed in Table 2.

Table 1. Number of sequences used for each capsid protein

	Bovine	Equine	Porcine	Human
VP2	78	18	17	114
VP6	166	23	207	540
VP7	327	67	272	570

VP2

The sequence variation of the inner capsid protein, VP2, is low, i.e. 95–97 % of amino acid content has complete conservation within and across species with only 1–17 sites having less than 70% conservation. This correlates with previous findings indicating 84% amino acid similarity (McDonald and Patton, 2008). This inner capsid protein, has few sites of variability on either interface surface (Fig. 3a) owing to the necessary interactions with VP6 on the outside and the viral polymerase (VP1) and the non-specific interactions with the nucleic acid on the inside (Boyle and Holmes, 1986). The core shell domain of VP2 is also said to activate VP1 polymerase activity (McDonald and Patton, 2011), rather than just the N terminus of VP2. The low variability would also

mean that subgroup detection using antibodies to VP2 as proposed and tested by McDonald and Patton (2008) is a more viable option than antibodies to VP6 or VP7. The most variable region on the protein is located on the N-terminus (residues 1-81) as has been reported previously by McDonald and Patton, (2011) but could not be mapped as these residues are not modeled in the original PDB file and has not been modeled or any structure determined (Trask et al., 2012). It is a region that makes up the 'fivefold hub' with other VP2 N-termini (McClain et al., 2010) and is responsible for VP1 and VP3 encapsidation (Zeng et al., 1998), RNA interactions and efficient RNA synthesis (McDonald and Patton, 2011). The amino acid content in this region is also polar and positively charged across all the sequences which correspond to the hydration region of the 'fivefold hub' (McClain et al., 2010).

VP6

The middle capsid protein, VP6, also shows relatively high conservation with at least 87% of the amino acid content of the protein being 100% conserved and with 1–32 sites having less than 70% conservation. This latter figure accounts for less than 8% of the protein. The regions of variability





VP2/6/7 complex

Cartoon representation (a) and surface representation (b) of the complex of the three major capsid proteins in rotavirus. The dimeric complex of VP2 and the trimeric complexes for VP6 and VP7 are shown.



Structural homology models of the major capsid proteins

Existing PDB co-ordinates were used as the template for generating homology models using the deduced consensus sequence for VP2 [PDB:3KZ4] (a) VP6 [PDB:1QHD] (b) and VP7 [PDB:3FMG] (c). Structures are colored according to the degree of conservation where white represents highly conserved sites and blue represents high variation within each species. Homology models were generated using Accelrys Discovery Studio 3.5. Structure renderings according to conversation were performed using ProtSkin (Ritter *et al.*, 2004) and PyMOL (The PyMOL Molecular Graphics System, Version 0.99rc6 Schrödinger, LLC).

for VP6 are scattered across the protein but overall show a lower degree of variability than VP7 (Fig. 3b). It has been reported that a single mutation at position 172 or 305 could alter subgroup antibody specificity (Löpez *et al.*, 1994) but in Fig. 1b it is evident that there is variation at both of these sites across species. It is, therefore, not surprising that currently there are strains of rotaviruses that can no longer be allocated to one of the two subgroups (Desselberger and Iturriza-Gomara, 2001).

The VP6 protein, being the middle capsid layer protein, interacts with both the outer coat protein, VP7 (residues 271 to 342 (Gilbert et al., 2001) and the inner protein, VP2 (Charpilienne et al., 2002) (Fig. 2). Furthermore, VP6 exists as a trimer in the virion (Mathieu et al., 2001) and thus successful formation of the mature virion is dependent on these interactions making them sterically- and orientation-specific. This can only be achieved through well conserved interactions and hence, may maintain the pressure on amino acid sequences to be well conserved. VP6 is also instrumental in replication of the virus (Boudreaux et al., 2015). On cell entry the loss of VP7 initiates a conformational change in the VP6 trimers (Libersou et al., 2008) "switching" the virus from protecting the genome to inducing the viral polymerase VP1, present in the virus core, to synthesize mRNA. A mutational study of VP2-VP6 interacting residues indicated that the conformational change modifies the VP6 to allow the transcripts to exit (Charpilienne et al., 2002). It has been proposed that protein conformational changes are small rearrangements that are initiated through one of many possible pathways or networks within the proteins (Dokland, 2002). The switch mechanism associated with VP6 during the early stages of infection implies the need for relative conservation of the VP6 sequence, protein or multimer. All these factors indicate that there are many pressures on this protein and therefore, it is not surprising that there is a high level of conservation in the amino acid sequence within and across species.

The VP7 protein of blue tongue virus is almost identical in terms of sequence and structure to VP6 (Charpilienne *et al.*, 2002). This would imply that these proteins have evolved to a current optimal structure that is fit for purpose. This generates an opportunity that we can exploit viral proteins to treat viral infection and therefore, it is not surprising that

Table 2. Degree of sequence variation across protein for each species

	Bovine	Equine	Porcine	Human
VP2	3 (27)	5 (44)	5 (40)	5 (41)
VP6	12 (47)	13 (52)	13 (51)	11 (43)
VP7	30 (95)	27 (89)	33 (110)	31 (102)

Variation expressed as a percentage of amino acid content. Actual No. of differing sites is shown in parentheses.

VP6 has also been proposed as a second generation vaccine candidate (Ward and McNeal, 2010). In a mouse model, neutralizing antibodies to VP6 have been shown to protect against infection by stopping replication (Lappalainen *et al.*, 2014) and a significant reduction in viral shedding was observed in faeces of immunized mice. These results suggest a significant role for mucosal rotavirus VP6-specific IgA for the inhibition of viral replication *in vitro* and *in vivo*. The long term efficacy of vaccines using VP6 protein as the main antigen may be assured due the limited tolerance to changes as seen in sequence conservation (Table 2; Fig. 3b). Reassortment between VP6 and VP7 of rotavirus B (Marthaler *et al.*, 2014) occurs in a fairly random pattern and does not appear to affect large segments of the amino acid sequences or sections of the protein structure.

VP7

Regarding VP7, 72% of the amino acid content of this outer capsid protein is always conserved with 14-65 sites having lower than 70% conservation. These latter data account for just 19% of the protein. The epitopes of VP7 identified at regions 87-96 (Green and Kapikian, 1992) correlate with the hypervariable sites on the protein (Fig. 1c). This pattern of variation is also conserved across species (Fig. 3c). The VP7 protein, like VP6, exists as a trimer, forming a co-ordinate complex with Ca²⁺ ions (Dormitzer et al., 2000). However, in contrast to VP6, VP7 has fewer points of contact with neighboring proteins (Fig. 2). The conserved amino acids of VP7 correspond to the regions of subunit interaction of the trimer where negatively charged amino acids containing carboxyl groups are required to co-ordinate around the Ca²⁺ ion as well as residues of the glycosylation site (69 and 71). The glycosylation of VP7 has been said to facilitate the correct disulphide bond formation and folding of rotavirus VP7 (Mirazimi and Svensson, 1998). Disulphide bond rearrangements have been shown in thrombin to allow the protein to undergo an allosteric or conformational switch from an active to inactive form (Huntington and Esmon, 2003). A similar mechanism may function in the VP7 which has 4 disulphide bonds, allowing it to convert from soluble form into the conformation that attaches to VP6 and displaces the lipid layer around the double layered viral particle that then results in the functional triple layered viral particle. Conservation is also seen on the inward facing surface of the VP7 trimer where negatively charged amino acids predominate (Fig. 3c). This is the side of the VP7 protein that interacts with VP6. In contrast, variability exists on the opposite face of the trimer, the face that makes up the exposed surface of the virion. With a selective immune pressure exerted on the outward face, hypervariability in amino acids confers a survival advantage to the virus so that it may escape antibody neutralization.

Discussion

Rotavirus RNA polymerase possesses no proofreading ability and thus we would expect to find a random pattern of nucleotide sequence changes. This is supported by the plethora of sequence data available for rotavirus genomes, which show continuously occurring changes (Matthijnssens and van Ranst, 2012). The nucleotide sequence changes should then translate into a random pattern of amino acid changes across the proteins expressed from the RNA segments. However, the results presented here show that changes are limited to a small number of defined locations on the expressed structural proteins, even across species (Fig. 1 and 3). The proteins must, therefore, have a significant number of amino acids that are essential for structural and functional roles to ensure the integrity of the virus and success in the various stages in rotavirus replication. They each represent an advanced multifunctioning system (Zhang et al., 2014). Regions of variability are concentrated on the surfaces that are at some point in the viral life cycle exposed to the host environment. Hence, it could be said that what the proteins sacrifice in structural integrity they gain in evading host defense mechanism.

The conservation of amino acid sequences is deliberate and any introduced variation must remain below a critical threshold - otherwise it almost always leads to removal of the nascent virus from the virus population. The result is the evolution of compact yet effective infectious entities that have produced stable protein capsid structures. These allow us to identify and pursue amino acid sets within these capsid proteins for use as research models to tease out functional relationships not only in the molecular pathogenesis of these viruses but also to provide us with target areas for treatment strategies and vaccine candidates. Recombinant segments of the viral proteins which align to regions of high conservation to produce powerful long lasting vaccines can be expressed. Such drugs and vaccine components should be designed to take full advantage of the regions on each protein which are constant. As seen for influenza vaccine strategies, conserved viral protein segments act as ideal therapeutic targets (Rahn et al., 2015). As an example of a potential rotavirus therapeutic target, VP6 has been shown to interact with heat shock cognate protein hsc70 and this interaction is said to mediate cell entry (Gualtero et al., 2007). The residues on VP6 that are involved in the interaction with hsc70 are amino acids 280–297 and from our data it is clear that this is a highly conserved region (Fig. 1b). Creating a drug that can bind to VP6 region 280-297 would disable possible cell entry and even assembly of the virus. It would also be highly specific and thus should produce fewer side effects.

Furthermore, produced vaccines containing capsid proteins should have long term efficacy. The idea of possible zoonotic reassortment where rotavirus genotypes normally associated

with animal infection may infect humans is not unrealistic based on evidence presented here which illustrates the conservation and symmetry associated with critical proteins. Therefore, one would assume that zoonosis should be taking place more often, producing new more virulent strains which bypass the immune system primed with only limited variability in strain specificity. However, the viral proteins seem to be very similar so even if zoonosis took place it would only affect the human population if the rare sequence variation was present in the zoonotic virus strain and that this virus was able to evade the protective immune response to administered vaccines. This emphasizes the need to not only research those molecular viral features that are critical in rotavirus infection but also to determine host factors that predispose individuals towards rotavirus infection as well as the nature of the immune response following vaccination. The current data also may explain why the vaccines are still relatively effective even if possible zoonosis is taking place because the components of the current vaccines elicit immune responses that may be aimed at epitope regions that are conserved across species.

Nonetheless, assuming that rotavirus capsid proteins are highly evolved, they could be seen as ideal models for the design of drug delivery shells. Mutations in these regions are not represented in the circulating viruses and are, therefore, not tolerated in the capsid system for the full functioning of the virus (Woods, 2015). Some of these functions would be specific to rotavirus but not for production of a basic delivery shell. Full understanding of sequence elements that relate to each functional role will provide insight into how to construct and optimized virus-like particles for use in various applications; aid in diagnostic tools for subtype detection and to design antibodies for non-variable regions on the antigen.

Acknowledgements. This research was supported by High Performance Computing centre of UNISA and by the NRF.

References

- Boudreaux CE, Kelly DF, McDonald SM (2015): Electron microscopic analysis of rotavirus assembly-replication intermediates. Virology 477, 32–41. <u>http://dx.doi.org/10.1016/j.</u> virol.2015.01.003
- Boyle JF, Holmes KV (1986): RNA-binding proteins of bovine rotavirus. J. Virol. 58, 561–568.
- Ciarlet M, Schödel F (2009): Development of a rotavirus vaccine: Clinical safety, immunogenicity, and efficacy of the pentavalent rotavirus vaccine, RotaTeq*. Vaccine, 27 (Suppl. 6), G72–G81. <u>http://dx.doi.org/10.1016/j. vaccine.2009.09.107</u>
- Charpilienne A, Lepault J, Rey F, Cohen J (2002): Identification of rotavirus VP6 residues located at the interface with VP2 that are essential for capsid assembly and transcriptase

activity. J. Virol. 76, 7822–7831. http://dx.doi.org/10.1128/ JVI.76.15.7822-7831.2002

- Collins PJ, Mulherin E, O'Shea H, Cashman O, Lennon G, Pidgeon E, Coughlan S, Hall W, Fanning S (2015): Changing patterns of rotavirus strains circulating in Ireland: Re-emergence of G2P[4] and identification of novel genotypes in Ireland. J. Med. Virol. 87, 764–773. <u>http:// dx.doi.org/10.1002/jmv.24095</u>
- Cunliffe NA, Witte D, Ngwira BM, Todd S, Bostock NJ, Turner AM, Chimpeni P, Victor JC, Steele AD, Bouckenooghe A, Neuzil KM (2012): Efficacy of human rotavirus vaccine against severe gastroenteritis in Malawian children in the first two years of life: A randomized, double-blind, placebo controlled trial. Vaccine 30 (Suppl. 1), A36–A43. http://dx.doi.org/10.1016/j.vaccine.2011.09.120
- Desselberger U, Iturriza-Gómara M, Gray JJ (2001): Rotavirus epidemiology and surveillance. Novartis Foundation Symposium 238, 125–152. <u>http://dx.doi.</u> <u>org/10.1002/0470846534.ch9</u>
- Dokland T (2000): Freedom and restraint: Themes in virus capsid assembly. Structure 8, R157–R162. <u>http://dx.doi.</u> org/10.1016/s0969-2126(00)00181-7
- Dormitzer PR, Greenberg HB, Harrison SC (2000): Purified recombinant rotavirus VP7 forms soluble, calciumdependent trimmers. Virology 277, 420-428. <u>http:// dx.doi.org/10.1006/viro.2000.0625</u>
- Estes MK, Cohen J (1989): Rotavirus gene structure and function. Microbiol. Rev. 53, 410–449.
- Green KY, Kapikian AZ (1992): Identification of VP7 epitopes associated with protection against human rotavirus illness or shedding in volunteers. J. Virol. 66, 548–553.
- Gilbert JM, Feng N, Patton JT, Greenberg HB (2001): Rotavirus assembly – Interaction of surface protein VP7 with middle layer protein VP6. Arch. Virol. 146, 1155–1171. <u>http:// dx.doi.org/10.1007/s007050170112</u>
- Gualtero DF, Guzmán F, Acosta O, Guerrero CA (2007): Amino acid domains 280-297 of VP6 and 531-554 of VP4 are implicated in heat shock cognate protein hsc70-mediated rotavirus infection. Arch. Virol. 152, 2183–2196. <u>http:// dx.doi.org/10.1007/s00705-007-1055-5</u>
- Honeyman MC, Coulson BS, Stone NL, Gellert SA, Goldwater PN, Steele CE, Couper JJ, Tait BD, Colman PG, Harrison LC (2000): Association between rotavirus infection and pancreatic islet autoimmunity in children at risk of developing type 1 diabetes. Diabetes 49, 1319–1324. <u>http:// dx.doi.org/10.2337/diabetes.49.8.1319</u>
- Huntington JA, Esmon CT (2003): The molecular basis of thrombin allostery revealed by a 1.8 Å structure of the "slow" form. Structure 11, 469–479. <u>http://dx.doi.org/10.1016/S0969-2126(03)00049-2</u>
- Jere KC, Mlera L, O'Neill HG, Potgieter AC, Page NA, Seheri ML, van Dijk, AA (2011): Whole genome analyses of African G2, G8, G9, and G12 rotavirus strains using sequence-independent amplification and 454° pyrosequencing. J. Med. Virol. 83, 2018–2042. http://dx.doi.org/10.1002/jmv.22207
- Lappalainen S, Pastor AR, Tamminen K, López-Guerrero V, Esquivel-Guadarrama F, Palomares LA, Vesikari T, Blazevic

V (2014): Immune responses elicited against rotavirus middle layer protein VP6 inhibit viral replication in vitro and in vivo. Human Vaccines and Immunotherapeutics 10, 2039–2047. http://dx.doi.org/10.4161/hv.28858

- Libersou S, Siebert X, Ouldali M, Estrozi LF, Navaza J, Charpilienne A, Garnier P, Poncet D, Lepault J (2008): Geometric mismatches within the concentric layers of rotavirus particles: A potential regulatory switch of viral particle transcription activity. J. Virol. 82, 2844–2852. <u>http:// dx.doi.org/10.1128/JVI.02268-07</u>
- Löpez S, Espinosa R, Greenberg HB, Arias, CF (1994): Mapping the subgroup epitopes of rotavirus protein VP6 Virology 204, 153–162. <u>http://dx.doi.org/10.1006/</u> <u>viro.1994.1519</u>
- Madhi SA, Cunliffe NA, Steele D, Witte D, Kirsten M, Louw C, Ngwira B, Victor JC, Gillard PH, Cheuvart BB, Han HH, Neuzil KM (2010): Effect of human rotavirus vaccine on severe diarrhea in African infants. New England J. Medicine, 362, 289–298. <u>http://dx.doi.org/10.1056/ NEJMoa0904797</u>
- Marthaler D., Suzuki T, Rossow K, Culhane M, Collins J, Goyal S, Tsunemitsu H, Ciarlet M, Matthijnssens, J (2014): VP6 genetic diversity, reassortment, intragenic recombination and classification of rotavirus B in American and Japanese pigs. Vet. Microbiol. 172, 359–366. <u>http://dx.doi. org/10.1016/j.vetmic.2014.05.015</u>
- Mathieu M, Petitpas I, Navaza J, Lepault J, Kohli E., Pothier P, Prasad BVV, Cohen J, Rey FA (2001): Atomic structure of the major capsid protein of rotavirus: Implications for the architecture of the virion. EMBO J. 20, 1485–1497. http://dx.doi.org/10.1093/emboj/20.7.1485
- Matthijnssens J, Ciarlet M, Rahman M, Attoui H, Bányai K, Estes MK, Gentsch JR, Iturriza-Gómara M, Kirkwood CD, Martella V, Mertens PPC, Nakagomi O, Patton JT, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Desselberger U, van Ranst, M (2008): Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments. Arch. Virol. 153, 1621–1629. http://dx.doi.org/10.1007/s00705-008-0155-1
- Matthijnssens J, Ciarlet M, McDonald SM, Attoui H, Bányai K, Brister JR, Buesa J, Esona MD, Estes MK, Gentsch JR, Iturriza-Gómara M, Johne R, Kirkwood CD, Martella V, Mertens PPC, Nakagomi O, Parre-o V, Rahman M, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Patton JT, Desselberger U, van Ranst M (2011): Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). Arch. Virol. 156, 1397–1413. <u>http://dx.doi.org/10.1007/</u> s00705-011-1006-z
- Matthijnssens J, van Ranst M (2012): Genotype constellation and evolution of group A rotaviruses infecting humans. Curr. Opin. Virol. 2, 426–433. <u>http://dx.doi.org/10.1016/j.</u> coviro.2012.04.007
- McClain B, Settembre E, Temple BRS, Bellamy AR, Harrison SC (2010): X-ray crystal structure of the rotavirus inner capsid particle at 3.8 Å resolution. J. Mol. Biol. 397, 587–599. http://dx.doi.org/10.1016/j.jmb.2010.01.055

- McDonald SM, Patton JT (2008): Molecular characterization of a subgroup specificity associated with the rotavirus inner capsid protein VP2. J. Virol. 82, 2752–2764. <u>http://dx.doi.</u> org/10.1128/JVI.02492-07
- McDonald SM, Patton JT (2011). Rotavirus VP2 core shell regions critical for viral polymerase activation. J. Virol. 85, 3095–3105. <u>http://dx.doi.org/10.1128/JVI.02360-10</u>
- Mirazimi A, Svensson L (1998): Carbohydrates facilitate correct disulfide bond formation and folding of rotavirus VP7. J. Virol. 72, 3887–3892.
- Mohamed MS, El-Hameed RHA, Sayed AI, Soror, SH (2015): Novel antiviral compounds against gastroenteric viral infections. Archiv der Pharmazie 348, 194–205. <u>http:// dx.doi.org/10.1002/ardp.201400387</u>
- Molbak K, Perch TKF, Mikkelsen, CS (2000): The estimation of mortality due to rotavirus infections in sub-Saharan Africa [1]. Vaccine 19, 393–395. <u>http://dx.doi.org/10.1016/</u> <u>S0264-410X(00)00199-7</u>
- Nagai M, Shimada S, Fujii Y, Moriyama H, Oba M, Katayama Y, Tsuchiaka S, Okazaki S, Omatsu T, Furuya T, Koyama S, Shirai J, Katayama K, Mizutani T (2015): H2 genotypes of G4P[6], G5P[7], and G9[23] porcine rotaviruses show super-short RNA electropherotypes. Vet. Microbiol. 176, 250–256. <u>http://dx.doi.org/10.1016/j.vetmic.2015.02.002</u>
- Nemoto M, Nagai M, Tsunemitsu H, Omatsu T, Furuya T, Shirai J, Kondo T, Fujii Y, Todaka R, Katayama K, Mizutani T (2015): Whole-genome sequence analysis of G3 and G14 equine group A rotaviruses isolated in the late 1990s and 2009-2010. Arch. Virol. 9 (in press). <u>http://dx.doi.</u> org/10.1007/s00705-015-2374-6
- Nyaga MM, Jere KC, Esona MD, Seheri ML, Stucker KM, Halpin RA, Akopov A, Stockwell TB, Peenze I, Diop A, Ndiaye K, Boula A, Maphalala G, Berejena C, Mwenda JM, Steele AD, Wentworth DE, Mphahlele MJ (2015): Whole genome detection of rotavirus mixed infections in human, porcine and bovine samples co-infected with various rotavirus strains collected from sub-Saharan Africa. Infect. Genet. Evol. 31, 321–334. http://dx.doi.org/10.1016/j.meegid.2015.02.011
- Papp H, Marton S, Farkas SL, Jakab F, Martella V, Malik YS, Palya V, Bányai K (2014): Classification and characterization of a laboratory chicken rotavirus strain carrying G7P[35] neutralization antigens on the genotype 4 backbone gene configuration. Biologicals 42, 299–304. <u>http://dx.doi. org/10.1016/j.biologicals.2014.08.004</u>
- Parashar UD, Hummelman EG, Bresee JS, Miller MA, Glass RI (2003): Global illness and deaths caused by rotavirus disease in children. Emerg. Infect. Dis. 9, 565–572. <u>http:// dx.doi.org/10.3201/eid0905.020562</u>
- Parashar UD, Burton A, Lanata C, Boschi-Pinto C, Shibuya K, Steele D, Birmingham M, Glass RI (2009): Global mortality associated with rotavirus disease among children in 2004. J. Infect. Dis. 200 (Suppl. 1), S9–S15. <u>http://dx.doi.</u> <u>org/10.1086/605025</u>
- Plosker GL (2010): Pentavalent rotavirus vaccine (RotaTeq*): A review of its use in the prevention of rotavirus gastroenteritis in Europe. Drugs 70, 1165–1188. <u>http://dx.doi.</u> org/10.2165/11205030-00000000-00000

- Ritter B, Denisov AY, Philie J, Deprez C, Tung EC, Gehring K, McPherson PS (2004): Two WXXF-based motifs in NE-CAPs define the specificity of accessory protein binding to AP-1 and AP-2. EMBO J. 23, 3701–3710. <u>http://dx.doi. org/10.1038/sj.emboj.7600378</u>
- Rahn J, Hoffmann D, Harder TC, Beer M (2015): Vaccines against influenza A viruses in poultry and swine: Status and future developments. Vaccine 33, 2414–2424. <u>http://dx.doi.</u> <u>org/10.1016/j.vaccine.2015.03.052</u>
- Rossignol JF, Abu-Zekry M, Hussein A, Santoro MG (2006): Effect of nitazoxanide for treatment of severe rotavirus diarrhoea: randomised double-blind placebo-controlled trial. Lancet 368, 124–129. <u>http://dx.doi.org/10.1016/ S0140-6736(06)68852-1</u>
- Sarker SA, Casswall TH, Mahalanabis D, Alam NH, Albert MJ, Brüssow H, Fuchs GJ, Hammerström L (1998): Successful treatment of rotavirus diarrhea in children with immunoglobulin from immunized bovine colostrum. Pediatric Infect. Dis. J. 17, 1149–1154. <u>http://dx.doi. org/10.1097/00006454-199812000-00010</u>
- Sow SO, Tapia M, Haidara FC, Ciarlet M, Diallo F, Kodio M, Doumbia M, Dembélé RD, Traoré O, Onwuchekwa UU, Lewis KDC, Victor JC, Steele AD, Neuzil KM, Kotloff KL, Levine MM (2012): Efficacy of the oral pentavalent rotavirus vaccine in Mali. Vaccine 30 (Suppl. 1), A71–A78. <u>http:// dx.doi.org/10.1016/j.vaccine.2011.11.094</u>
- Tate JE, Burton AH, Boschi-Pinto, C, Steele, AD, Duque, J, Parashar, UD (2012): 2008 estimate of worldwide rotavirusassociated mortality in children younger than 5 years before the introduction of universal rotavirus vaccination programmes: A systematic review and meta-analysis (2012) The Lancet Infect. Dis. 12, 136–141. <u>http://dx.doi. org/10.1016/S1473-3099(11)70253-5</u>
- Trask SD, Ogden KM, Patton JT (2012): Interactions among capsid proteins orchestrate rotavirus particle functions. Curr. Opin. Virol. 2, 373–379. <u>http://dx.doi.org/10.1016/j.</u> <u>coviro.2012.04.005</u>
- Wang FT, Mast TC, Glass RJ, Loughlin J, Seeger JD (2010): Effectiveness of the pentavalent rotavirus vaccine in preventing gastroenteritis in the United States. Pediatrics 125, e208-e213. <u>http://dx.doi.org/10.1542/ peds.2009-1246</u>
- Ward RL, Bernstein DI (2009): Rotarix: A rotavirus vaccine for the world. Clin. Infect. Dis. 48, 222–228. <u>http://dx.doi.org/10.1086/595702</u>
- Ward RL, McNeal MM (2010): VP6: A candidate rotavirus vaccine. J. Infect. Dis. 202 (Suppl. 1), S101–S107. <u>http://dx.doi.org/10.1086/653556</u>
- Zaman K, Anh DD, Victor JC, Shin S, Yunus Md, Dallas MJ, Podder G, Thiem VD, Mai LTP, Luby SP, Tho LH, Coia ML, Lewis K, Rivers SB, Sack DA, Schödel F, Steele AD, Neuzil KM, Ciarlet M (2010): Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in Asia: A randomised, double-blind, placebo-controlled trial. The Lancet 376, 615–623. <u>http://dx.doi.org/10.1016/S0140-6736-(10)60755-6</u>

- Zeng CQY, Estes MK, Charpilienne A, Cohen J (1998): The N terminus of rotavirus VP2 is necessary for encapsidation of VP1 and VP3. J. Virol. 72, 201–208.
- Zhang S, McDonald PW, Thompson TA, Dennis AF, Akopov A, Kirkness EF, Patton JT, McDonald SM (2014): Analysis of human rotaviruses from a single location over an 18-year time span suggests that protein coadaption influences

gene constellations. J. Virol. 1, 88, 9842–9863. <u>http://</u> <u>dx.doi.org/10.1128/JVI.01562-14</u>

Woods RJ (2015): Intrasegmental recombination does not contribute to the long-term evolution of group A rotavirus. Infect. Genet. Evol. 32, 354–360. <u>http://dx.doi.org/10.1016/j.</u> <u>meegid.2015.03.035</u>