Correlation between recombination junctions and RNA secondary structure elements in poliovirus Sabin strains

Evangelos Dedepsidis · Zaharoula Kyriakopoulou ·
Vaia Pliaka · Panayotis Markoulatos

Received: 25 November 2009 / Accepted: 28 June 2010 / Published online: 17 July 2010
© Springer Science+Business Media, LLC 2010

Abstract In order to test the hypothesis that RNA structural elements promote the distribution of certain types of recombination junctions in each one of the 2C and 3D poliovirus genomic regions (Sabin 3/Sabin 2 or Sabin 1 in 2C and Sabin 2/Sabin 1 or Sabin 3 in 3D), we searched in 2C and 3D regions of reference Sabin strains for high probability RNA structural elements that could promote recombination. Recombination junctions that were identified in clinical strains of this study, as well as in clinical strains of previous studies, were superimposed on RNA secondary structure models of 2C and 3D genomic regions. Furthermore, we created an in vitro model, based on double infection of cell-culture with two poliovirus strains, for the production and identification of recombinant Sabin strains in 2C and 3D regions. Our intention was to compare the results that refer to the correlation of recombination junctions and RNA secondary structures in 2C and 3D regions of clinical strains, with the respective results of the in vitro model. Most of the recombination junctions of the clinical strains were correlated with RNA secondary structure elements, which were identical between recombining Sabin strains, and also presented high predictive value. In consensus were, the respective results originated from the in vitro model. We propose that the distribution of specific types of recombination junctions in certain regions of

Sabin strains is not fortuitous and is correlated with RNA secondary structure elements identical to both recombination partners. Furthermore, results of this study highlight an important role for the stem region of the RNA structure elements in promoting recombination.

Keywords Poliovirus · Recombination · RNA · Secondary structure

Introduction

Polioviruses, the causal agent of acute paralytic poliomyelitis, belong to the genus *Enterovirus* of the *Picornaviridae* family and are classified into three serotypes poliovirus 1 (PV1), poliovirus 2 (PV2), and poliovirus 3 (PV3). Poliovirus is a nonenveloped positive single-stranded RNA virus, whose genome is approximately 7,500 bases in length. Poliovirus genome contains a 5'-untranslated region (5' UTR), a coding region, a 3'-untranslated region (3' UTR), and a poly(A) tail. The structural (VP4, VP2, VP3, and VP1) and nonstructural (2A, 2B, 2C, 3A, 3B, 3C, and 3D) viral proteins are produced after translation of the coding region and proteolytic cleavage of the polyprotein [1].

There are two vaccines against poliovirus, the inactivated poliovirus vaccine (IPV) and the oral poliovirus vaccine (OPV). OPV consists of live attenuated poliovirus Sabin strains of the three serotypes (Sabin 1, Sabin 2, and Sabin 3). OPV is the vaccine of choice since it has several advantages over IPV, such as: simplicity of administration, induction of mucosal immunity and neutralizing antibodies, and low cost. However, genetic instability of OPV is associated with adverse effects such as vaccine-associated paralytic poliomyelitis (VAPP) and vaccine-derived polioviruses [2].

Electronic supplementary material The online version of this article (doi:10.1007/s11262-010-0512-5) contains supplementary material, which is available to authorized users.

E. Dedepsidis · Z. Kyriakopoulou · V. Pliaka ·
P. Markoulatos (✉)
Department of Biochemistry and Biotechnology, Microbiology–
Virology Laboratory, University of Thessaly, 26 Ploutonos
and Aiolou Str., Larisa 41221, Greece
e-mail: Markoulatos@bio.uth.gr

Poliovirus is considered a rapidly evolving RNA virus [3, 4]. The mechanisms responsible for the high-evolutionary rates of poliovirus are mutations and recombination.

Three different mechanisms of recombination have been proposed for poliovirus. The “primer alignment-and-extension,” the “breakage and ligation,” and the “template switch” mechanism [5–7]. The most widely accepted mechanism for RNA recombination in poliovirus is the template switch mechanism proposed by Kirkegaard and Baltimore [6]. In this mechanism, poliovirus RNA dependant RNA polymerase (RdRp) switches templates during negative strand synthesis.

Intertypic recombination between Sabin strains is a frequent event due to the trivalent nature of OPV. Co infection of the intestinal cells of a vaccinated person with Sabin 1, Sabin 2, and Sabin 3 strains, and the mixed replication complexes locating at few preferential sites in the perinuclear region of the target cells, are facts that promote intertypic recombination [8]. Intertypic recombination between Sabin strains usually takes place at the nonstructural part of the genome and a high proportion of types 2 and 3 strains isolated from healthy vaccinated individuals and from VAPP patients, have been found to be recombinants [9–12]. According to previous studies, specific types of recombination junctions are located in certain genomic regions of clinical strains. Recombination junctions of Sabin 3/Sabin x type (x: Sabin 2 or Sabin 1) are mostly located in 2C genomic region, while recombination junctions of Sabin 2/Sabin x type (x: Sabin 1 or Sabin 3) are mostly located in 3D genomic region [10–14]. However, the underlying cause for the localization of specific recombination types in certain genomic regions remains unsolved. It has been proposed that RNA structural elements are responsible for the distribution of specific types of recombination junctions in each one of the 2C and 3D genomic regions [10–12, 14].

In order to test the abovementioned hypothesis, that RNA structural elements are responsible for the distribution of specific types of recombination junctions in each one of the 2C and 3D poliovirus genomic regions, we searched for high probability RNA secondary structure elements in 2C and 3D genomic regions of Sabin 1, Sabin 2, and Sabin 3 reference strains that could promote recombination. Recombination junctions which were identified in 2C and 3D regions of clinical strains of this study, as well as recombination junctions which had been identified in 2C and 3D regions of clinical strains of previous studies were superimposed on the highly probable RNA secondary structure elements of 2C and 3D regions. Furthermore, we created a cell-culture based in vitro model, targeted in the production and isolation of Sabin strains whose recombination junctions were located in either 2C or 3D genomic regions. We created the in vitro

model for two reasons, first, in order to check if the distribution pattern of recombination types which is observed in 2C and 3D genomic regions of clinical strains is also observed in cell-culture strains, and second, in order to gather more evidence about the role of the RNA structural elements in promoting recombination. Finally, we compared the results that refer to the correlation between recombination junctions and RNA secondary structure elements in 2C and 3D regions of clinical strains, with the respective results of the in vitro model.

Materials and methods

Reference strains/clinical strains

The clinical strains of this study were isolated from faecal samples during the time period 1978–1985, and the available data of these strains are described in Table 1. The data for clinical strains of previous studies [10–15], whose recombination junctions were superimposed in RNA secondary structure elements of 2C and 3D regions, are shown in Online Resource 1. The reference strains Sabin 1, Sabin 2, and Sabin 3 were provided by the National Institute for Biological Standards and Control (UK). The reference sequences of the Sabin 1, Sabin 2, and Sabin 3 strains, which were used for the detection of recombination junctions and for the identification of RNA secondary structure elements are cited in GenBank under the accession numbers AY184219, AY184220, and AY184221, respectively.

Virus isolation and identification/extraction of viral RNA

For the clinical strains of this study, virus isolation was performed from faecal samples of VAPP patients and healthy vaccinees, in Hep-2 cells. Initial isolation at that time occurred at the Hellenic Pasteur Institute and seroneutralization with rabbit polyclonal antibodies (National Institute for Public Health and Environment, RIVM, The Netherlands) was performed for serotyping.

In order to avoid viral mixtures during virus isolation, serial tenfold dilutions were performed; the last dilution presenting a complete cytopathic effect (CPE) was passaged once again in Hep-2 cells, and the serotype was retested as described above. Genotyping was performed on the basis of 5' UTR according to the method of Georgopoulou et al. [16]. In brief, primer pair UG52–UC53 was used for amplification of the genomic region nt 162–595 in all clinical isolates. Restriction fragment length polymorphism (RFLP) analysis using HaeIII, HpaII, NcoI, and AvaI restriction enzymes revealed the genotype of each clinical isolate.

Table 1 Available data of clinical strains that were used for identification of recombination junctions

Strain	Serotype	Genotype (5' UTR)	Case type	Isolation period	GenBank accession number
C1	Polio 2	Sabin 2	V.A.P.P	1978–1985	–
C2	Polio 3	Sabin 3	Healthy vaccinee	1978–1985	–
C3	Polio 1	Sabin 1	Healthy vaccinee	1978–1985	–
C4	Polio 1	Sabin 1	Healthy vaccinee	1978–1985	–
C5	Polio 3	Sabin 3	Healthy vaccinee	1978–1985	–
C6	Polio 3	Sabin 3	V.A.P.P	1978–1985	EU598487
C7	Polio 1	Sabin 1	V.A.P.P	1978–1985	EU598488
C8	Polio 3	Sabin 3	Healthy vaccinee	1978–1985	–
C9	Polio 3	Sabin 3	V.A.P.P	1978–1985	EU598486

Viral RNA extraction was performed according to the method of Casas et al. [17].

RT-PCR, RFLP, and sequencing analysis for identification of recombinant clinical isolates/primer pairs for sequencing of recombinant clinical isolates

The RT-PCR as well as the RFLP assays used in the protocol for identification of recombination junctions in clinical strains of this study is described in detail in previous studies [11, 12]. For size analysis of the RFLP fragments, Gel Pro Analyzer software package (Media Cybernetics, USA) was used. Primer pairs that were used for sequencing of the genomic region nt 4717–7415 of any putative recombinant strain defined by RFLP analysis, are described in detail in previous studies [10, 20]. Moreover, primer pairs 71935–EUC2 and UG23–UC15 were used for analysis of the genomic region nt 3206–4965 of any putative recombinant strain defined by RFLP analysis, according to the PCR conditions described in the original publications [18–20]. Where necessary PCR products were extracted from the electrophoresis gel using a PCR gel extraction kit (Qiagen GmbH, Hilden, Germany). All PCR products were sequenced at Macrogen, Inc. Clustalw was used for the identification of recombination junctions in the clinical isolates of this study. Simplot (version 2.5) program was used for the depiction of recombination junctions of clinical isolates of this study.

Cell-culture based in vitro model of recombination

For every combination of a double infection, approximately, 2×10^6 Hep2 cells were introduced in every well of a six-well plate (Sigma, Aldrich) and were incubated at 37°C in growth medium [minimum essential medium (MEM) Eagle's supplemented with 10% FCS] (Sigma, Aldrich) until a monolayer was formed. Growth medium was decanted and the first five wells were inoculated with

5×10^5 TCID₅₀ (tissue culture infectious dose) of a specific Sabin strain, and with 5×10^5 TCID₅₀ of a second specific Sabin strain belonging to a different serotype from the first. All possible combinations of a double infection were performed (Sabin 1–Sabin 2, Sabin 1–Sabin 3, and Sabin 2–Sabin 3) in different six-well plates each time. In the sixth well of the plate, which was used as a negative control, maintenance medium (MEM Eagle's supplemented with 2% FCS) was added. After 1 h of incubation, the medium that contained unbound virus was decanted and 1 ml of maintenance medium was added in each well. Plates were incubated at 37°C until complete CPE was observed. Complete CPE appeared 24 h after infection of Hep2 cells, and the content of every well of the six-well plate was inoculated in ten wells of a 96-well plate (100 µl/well) (Sigma, Aldrich). In every well of the 96-well plate, there were, approximately, 10^4 Hep2 cells. Plates were incubated at 37°C until the appearance of a complete CPE (approximately, 24 h).

Detection of recombinant Sabin strains in the in vitro model by RT-PCR and cloning

From each well of the 96-well plate viral RNA was extracted, and RT-PCR was performed for the detection of recombinant Sabin strains. Extraction of viral RNA and reverse transcription was performed as previously described [12]. In order to design primer pairs which detect only recombinant molecules in 2C and 3D regions, we generate sequence alignments for 2C and 3D regions of Sabin 1, Sabin 2, and Sabin 3 reference strains, using Clustalw. We searched in the alignments for regions of low homology among Sabin 1, Sabin 2, and Sabin 3 reference strains. Primer pairs were designed to anneal to these low-homology regions aiming to the amplification of only the recombinant molecules that would be produced from the in vitro model. Program PRIMER 3 (<http://www.genome.wi.mit.edu/genomesoftware/other/>) was used for Tm calculation of

primers and in order to check if the selected primer pairs accomplish certain criteria (e.g., no dimer formation, no self complementarity, etc). These primer pairs were design to detect only recombinant molecules in 2C and 3D regions of the poliovirus genome and were able to detect all possible types of recombination junctions in these regions (Online Resources 2 and 3). PCR mixture for each reaction was consisted of 3 µl cDNA, 2 µl of each primer pair (25 pmol), 5 µl dNTPs (10 mM), 5 µl 10× reaction buffer (which contained MgCl₂ up to a final concentration of 2 mM), 0.5 µl (2.5 U) Paq5000 DNA polymerase (Stratagene, La Jolla, USA), and double distilled H₂O up to a final volume of 50 µl. The reaction conditions are described in Online Resources 2 and 3 for 2C and 3D regions, respectively. A final extension step was carried out at 74°C for 5 min.

In order to avoid mixtures of recombinant molecules in PCR products of the in vitro model, cloning was performed in every PCR product of the in vitro model. For cloning purposes, PCR products of the in vitro model were reamplified using GoTaq polymerase (Promega, Madison, WI, USA) which leaves a overhangs at the ends of the products. For cloning of the in vitro model PCR products, pGem-TEasy vector system II (Promega) was used. All cloning products of the in vitro model were sequenced at Macro-gen, Inc. Clustalw was used for identification of the recombination junctions of the in vitro isolates of this study.

In order to exclude the possibility of generation of recombination artifacts during RT-PCR (false-positive results), we mixed the RNA extracts of the reference strains, and we performed RT-PCR control experiments for every primer pair of Online Resources 2 and 3. There was no false positive result obtained in any reaction.

RNA secondary structure elements

Partition function calculations and RNA folding for 2C and 3D regions of positive strand Sabin 1, Sabin 2, and Sabin 3 reference strains, were accomplished through program RNAstructure [21]. Furthermore, 2C and 3D regions were separated in consecutively 600 bp fragments with 300 bp overlaps. Partition function calculation and RNA folding for these 600 bp fragments were also accomplished through RNAstructure. Mfold V 3.2 was used for the depiction of RNA secondary structure elements [22].

Nucleotide sequence accession numbers

The partial nucleotide sequences of isolates C6, C7, and C9 have been deposited in GenBank library under the accession numbers EU598486, EU598488, and EU598487, respectively.

Results

Identification of recombinant clinical isolates by RFLP and sequencing

In this study, nine clinical strains were analyzed for recombination events throughout their genome. More specifically, for each strain RFLP analysis was conducted in three distant genomic regions (VP1, 2C, and 3D) aiming to the identification of digestion profile that would belong to a different genotype from the one predicted on the basis of 5' UTR (Table 1). The strains which presented RFLP profile different from the expected one, were sequenced in order to identify the recombination junction(s).

Out of the nine isolates, which were studied for recombination events, three presented RFLP profiles different from the expected ones. Specifically, strains C6, C7, and C9, all of which were isolated from VAPP patients, presented in the non capsid part of their genome RFLP profiles different from the 5' UTR genotype profile (Online Resource 4, Table 2). In order to identify the exact position of the possible recombination junction, genomic sequencing of region nt 4717–7415 was performed.

Genomic sequencing of region nt 4717–7415 of isolates C6, C7, and C9 revealed two recombination junctions in each of these isolates. In isolate C6, the first recombination junction was located at the end of 3C region (nt 5804–5814) and was of S3/S2 type, while the second was of S2/S1 type and located in 3D region (nt 6893–6901). In isolate C7, the first recombination junction was of type S1/S3 and located in 2A region (nt 3461–3465), while the second recombination junction was of S3/S2 type and located in 2C region (nt 4517–4533). Finally, in isolate C9, the first recombination junction located in 2C region (nt 4449–4464) and was of S3/S2 type, while the second

Table 2 RFLP analysis in 5' UTR, VP1, 2C, and 3D regions of clinical strains used in this study

Clinical strain	(Genotype) 5' UTR	VP1	2C	3D
C1	Sabin 2	Sabin 2	Sabin 2	Sabin 2
C2	Sabin 3	Sabin 3	Sabin 3	Sabin 3
C3	Sabin 1	Sabin 1	Sabin 1	Sabin 1
C4	Sabin 1	Sabin 1	Sabin 1	Sabin 1
C5	Sabin 3	Sabin 3	Sabin 3	Sabin 3
C6	Sabin 3	Sabin 3	Sabin 3	Sabin 2
C7	Sabin 1	Sabin 1	Sabin 3	Sabin 2
C8	Sabin 3	Sabin 3	Sabin 3	Sabin 3
C9	Sabin 3	Sabin 3	Sabin 3	Sabin 2

The characters in bold highlight the different RFLP profiles observed in different genomic regions of the recombinant strains and are important

Fig. 1 a Schematic representation of the recombinant genomes of C6, C7, and C9 clinical isolates of this study. Accession numbers EU598486, EU598488, and EU598487, respectively.

b Schematic representation of the recombinant genomes of V1, V2, V3, V4, V5, and V6 isolates, generated from the cell-culture in vitro model used in this study

recombination junction was of S2/S1 type and located at the end of 3D region (nt 7142–7164) (Online Resource 5A, B; Fig. 1a).

Identification of recombinant isolates in cell-culture based in vitro model

From the cell-culture based in vitro model which is targeted in the production and isolation of recombinant Sabin strains in 2C and 3D genomic regions, six isolates (V1–V6) were detected, all of which found to be recombinant in their 3D genomic region (Table 3). All isolates were retraced with the primer pair S12sD3–S21aD3 which is designed to identify S1/S2 recombination events in 3D genomic region. The recombinant isolates of the in vitro model were originated from different wells of the six-well plate and as a consequence every in vitro isolate was result of an independent recombination event.

Isolates V1, V2, and V3 presented recombination junctions of the same type (S1/S2), located exactly in the same spot (nt 7073–7110). In isolates V5 and V6, recombination junctions of S1/S2 type were detected, which were located at the same spot (nt 6837–6841). Finally, in isolate V4 three recombination junctions were detected. The first recombination junction was of S1/S2 type (nt 6837–6841), the second was of S2/S1 type (nt 6867–6895), and the third

recombination junction was of S1/S2 type (nt 6903–6911) (Fig. 1b). Isolate V4 is a unique molecule which contain three different recombination junctions in the 3D region of its genome, because we perform cloning after PCR and prior to sequencing.

Topology of recombination junctions of clinical strains in correlation with RNA secondary structure elements of 2C and 3D genomic regions

In this study, program RNAstructure was used in order to examine whether highly probable RNA secondary structure elements are situated in the regions in which recombination junctions of clinical strains of this study, as well as of clinical strains of previous studies [10–15], were identified.

In Tables 4 and 5 and in Online Resources 6 and 7, the base pair probabilities as well as the positive predictive values of every RNA structure element are presented, which is in correlation with recombination junction of clinical strains.

More specifically, in Tables 4 and 5 the RNA structure elements are presented, which are associated with recombination junctions in 2C and 3D region, respectively, and are also identical between the recombining molecules. In Online Resources 6 and 7, the RNA structure elements are presented, which are associated with recombination

Table 3 Analysis of the recombination junctions of the in vitro model isolates

In vitro model isolate	Recombination type	Region	Recombination junction
V1	S1/S2	3D	7073–7110
V2	S1/S2	3D	7073–7110
V3	S1/S2	3D	7073–7110
V4	S1/S2, S2/S1, S1/S2	3D	6837–6841, 6867–6895, 6903–6911
V5	S1/S2	3D	6837–6841
V6	S1/S2	3D	6837–6841

Table 4 RNA structural elements which correlated with recombination junctions of clinical strains in 2C region, and are identical between recombining partners

Strain	Recombination type	Region	Base pairs	Base pair probability (mean value)	Positive predictive value (%)	Recombination junction location
C9	Sabin 3/Sabin 2	2C	S3: (4449–4463/4489–4475)	$3.3 \times e^{-0.03}$	91 ± 5.9	Stem
			S2: (4449–4463/4489–4475)	$2.28 \times e^{-0.03}$	91 ± 5.9	
D1	Sabin 3/Sabin 2	2C	S3: (4505–4506/4515–4514)	0.91	<73	Stem–loop
			S2: (4504–4506/4516–4514)	0.42	<73	
D2, D3, C7, IM, D4	Sabin 3/Sabin 2	2C	S3: (4523–4526/4534–4531)	0.92	<73	Stem–loop
			S2: (4523–4526/4534–4531)	0.34	73 ± 10.9	
D5	Sabin 3/Sabin 2	2C	S3: (4562–4564/4572–4570)	0.15	73 ± 10.9	Stem–loop
			S2: (4562–4564/4572–4570)	$7.56 \times e^{-0.02}$	76.6 ± 10.3	
EPA, EPB, EPC	Sabin 3/Sabin 2	2C	S3: (4625–4636/4652–4640)	0.42	<73	Stem–single stranded
			S2: (4625–4636/4652–4640)	$4.28 \times e^{-0.02}$	76.6 ± 10.3	
D6	Sabin 3/Sabin 2	2C	S3: (4703–4706/4719–4715)	0.22	<73	Stem
			S2: (4703–4706/4719–4715)	$8.3 \times e^{-0.02}$	91 ± 5.9	
D8	Sabin 3/Sabin 2	2C	S3: (4833–4852/4880–4859)	0.63	<73	Stem
			S2: (4833–4852/4880–4859)	$3.73 \times e^{-0.02}$	91 ± 5.9	
LK6, LK10	Sabin 2/Sabin 1	2C	S2: (4959–4983/5011–4987)	0.3	73 ± 10.9	Single stranded next to RNA structural element
			S1: (4959–4983/5011–4987)	$6 \times e^{-0.02}$	76.6 ± 10.3	
I34	Sabin 2/Sabin 1	2C	S2: (4959–4983/5011–4987)	0.3	73 ± 10.9	Loop–stem
			S1: (4959–4983/5011–4987)	$6 \times e^{-0.02}$	76.6 ± 10.3	
31043	Sabin 2/Sabin 1	2C	S2: (4959–4983/5011–4987)	0.3	73 ± 10.9	Stem
			S1: (4959–4983/5011–4987)	$6 \times e^{-0.02}$	76.6 ± 10.3	
D24	Sabin 2/Sabin 1	2C	S2: (5037–5041/5052–5048)	$3.34 \times e^{-0.02}$	83.2 ± 8.3	Stem–loop–stem
			S1: (5037–5041/5052–5048)	$5.23 \times e^{-0.02}$	76.6 ± 10.3	
8029	Sabin 3/Sabin 1	2C	S3: (4862–4880/4901–4885)	0.14	76.6 ± 10.3	Stem–loop–stem
			S1: (4861–4880/4901–4885)	0.16	76.6 ± 10.3	
D26	Sabin 3/Sabin 1	2C	S3: (4862–4880/4901–4885)	0.14	76.6 ± 10.3	Stem
			S1: (4861–4880/4901–4885)	0.16	76.6 ± 10.3	
D25	Sabin 3/Sabin 1	2C	S3: (4862–4880/4901–4885)	0.14	76.6 ± 10.3	Single stranded next to RNA structural element
			S1: (4861–4880/4901–4885)	0.16	76.6 ± 10.3	

Base pairs probabilities were obtained through RNAstructure, positive predictive values were obtained from Mathews [21]

junctions in 2C and 3D region, respectively, but are different between the recombining molecules.

The most important feature considering the topology of recombination junctions of clinical strains in 2C and 3D regions is the fact that the majority of the recombination junctions were correlated with RNA structure elements which were identical between the recombining Sabin strains. Specifically, in 2C region 21 out of 29 recombination junctions were located in RNA structural elements identical between the recombining partners (Table 4); while in 3D region, 17 out of 28 recombination junctions belong in the same category (Table 5). Moreover, the base pairs which lead to the formation of these RNA structural elements present high probabilities and are also included in the minimum free energy structure and as a consequence the positive predictive values of these base pairs are high [21]. Furthermore, the RNA structure element which is

situated in the respective position of the third Sabin reference molecule, which molecule is not participating in the recombination, is entirely different from the identical RNA structural elements of the two Sabin reference molecules which are participating in the recombination (data not shown).

Most of the recombination junctions were located in the double stranded stem region of the RNA structural elements, while only a small percentage of the recombination junctions was located in the loop region or in single stranded regions downstream or upstream of RNA secondary structure elements.

The depiction of some recombination junctions of clinical strains in RNA secondary structure elements of 2C genomic region is presented in Fig. 2 and Online Resource 8, and in Fig. 3 and Online Resource 9, the depiction of some recombination junctions of clinical strains in RNA

Table 5 RNA structural elements which correlated with recombination junctions of clinical strains in 3D region, and are identical between recombining partners

Strain	Recombination type	Region	Base pairs (bp)	Base pair probability mean value	Positive predictive value (%)	Recombination junction location
EP6	Sabin 2/Sabin 1	3D	S2: (6366–6369/6377–6374) S1: (6366–6369/6377–6374)	0.38 0.39	73 ± 10.9 <73	Stem–loop
EP12, ENP6, D10	Sabin 2/Sabin 1	3D	S2: (6330–6338/6351–6343) S1: (6335–6339/6350–6346)	0.99 $1.57 \times e^{-0.02}$	76.6 ± 10.3 86.7 ± 8.6	Stem–loop
I34, D11	Sabin 2/Sabin 1	3D	S2: (6366–6369/6377–6374) S1: (6366–6369/6377–6374)	0.38 0.39	73 ± 10.9 <73	Single stranded next to RNA structural element
D12	Sabin 2/Sabin 1	3D	S2: (6463–6465/6476–6474) S1: (6462–6465/6477–6474)	0.51 $1.74 \times e^{-0.02}$	<73 86.7 ± 8.6	Stem
IM	Sabin 2/Sabin 1	3D	S2: (6213–6219/6235–6228) S1: (6213–6219/6235–6228)	$4.31 \times e^{-0.02}$ $2.67 \times e^{-0.02}$	91 ± 5.9 86.7 ± 8.6	Stem
ENP8	Sabin 2/Sabin 1	3D	S2: (6779–6782/6791–6788) S1: (6768–6773/6784–6779)	0.28 $3.03 \times e^{-0.03}$	73 ± 10.9 86.7 ± 8.6	Single stranded next to RNA structural element
D18	Sabin 2/Sabin 1	3D	S2: (6880–6887/6899–6892) S1: (6880–6887/6899–6892)	$1.11 \times e^{-0.03}$ 0.17	86.7 ± 8.6 76.6 ± 10.3	Single stranded–Stem
C6	Sabin 2/Sabin 1	3D	S2: (6880–6887/6899–6892) S1: (6880–6887/6899–6892)	$1.11 \times e^{-0.03}$ 0.17	86.7 ± 8.6 76.6 ± 10.3	Stem
C9	Sabin 2/Sabin 1	3D	S2: (7127–7134/7148–7140) (7152–7155/7195–7192) (7157–7160/7189–7186) S1: (7129–7135/7146–7140) (7152–7155/7195–7192) (7157–7160/7189–7186)	$4.13 \times e^{-0.02}$ 0.26 0.24 $4.67 \times e^{-0.02}$ 0.88 1.1	76.6 ± 10.3 73 ± 10.9 73 ± 10.9 86.7 ± 8.6 <73 <73	Stem–single stranded–stem
IF	Sabin 2/Sabin 1	3D	S2: (6160–6173/6281–6268) (6189–6194/6256–6251) (6199–6208/6247–6239) S1: (6160–6173/6281–6268) (6192–6202/6248–6237)	$4.38 \times e^{-0.02}$ $1.06 \times e^{-0.02}$ $2.96 \times e^{-0.04}$ $12 \times e^{-0.02}$ 0.29	76.6 ± 10.3 86.7 ± 8.6 91 ± 5.9 91 ± 5.9 73 ± 10.9	Stem–single stranded–stem
ENP7	Sabin 2/Sabin 3	3D	S2: (6541–6545/6554–6550) S3: (6541–6545/6554–6550)	$1.57 \times e^{-0.03}$ $1.37 \times e^{-0.04}$	91 ± 5.9 91 ± 5.9	Stem
D20	Sabin 2/Sabin 3	3D	S2: (6526–6530/6538–6534) S3: (6523–6526/6535–6532)	0.4 $3.2 \times e^{-0.03}$	73 ± 10.9 86.7 ± 8.6	Loop
591	Sabin 2/Sabin 3	3D	S2: (6526–6530/6538–6534) (6541–6545/6554–6550) S3: (6523–6526/6535–6532) (6541–6545/6554–6550)	0.4 $1.57 \times e^{-0.03}$ $3.2 \times e^{-0.03}$ $1.37 \times e^{-0.04}$	73 ± 10.9 91 ± 5.9 86.7 ± 8.6 91 ± 5.9	Stem–loop–single stranded–stem–loop
D22	Sabin 2/Sabin 3	3D	S2: (6838–6845/6856–6850) S3: (6838–6845/6856–6850)	0.22 0.76	73 ± 10.9 76.6 ± 10.3	Stem–loop

Base pairs probabilities were obtained through RNAstructure, positive predictive values were obtained from Mathews [21]

secondary structure elements of 3D genomic region is presented.

The recombination junctions of clinical isolates D2, D3, D4, C7, IM, and D8 in 2C region, were located in regions which are governed by many alternative base pairs of similar probability, and this is the reason for which positive predictive value for these base pairs is low.

Topology of recombination junctions of the in vitro model isolates in correlation with RNA secondary structure elements of 3D region

We were aiming to the comparison between the results that refer to the correlation between recombination junctions and RNA secondary structure elements in clinical strains,

Fig. 2 Representation of recombination junctions of isolates D6, EPA, EPB, EPC, C9, and D5 in RNA secondary structure elements of 2C genomic region of Sabin 2 and Sabin 3 reference strains

with the respective results of the in vitro model isolates. For this reason, RNAstructure was used in order to clarify whether the recombination junctions of the in vitro model isolates are located in spots of 3D genomic region, in which highly probable RNA secondary structure elements are situated.

In Table 6, the base pair probabilities and the positive predictive values of every RNA structure element are shown, which is in correlation with recombination junction of in vitro model isolates.

The results considering the correlation of recombination junctions of the in vitro model isolates with RNA secondary structure elements are in consensus with the respective results which come from the clinical strains. More specifically, all recombination junctions of the in vitro model isolates are correlated with RNA structural elements which are identical between the two recombining partners, as was the case for the majority of recombination junctions of the clinical strains in 2C and 3D regions. Moreover, the positive predictive value of the base pairs that lead to the formation of these RNA structure elements is high. Furthermore, in consensus with the majority of the

Fig. 3 Representation of recombination junctions of isolates IM, D18, and C6 in RNA secondary structure elements of 3D genomic region of Sabin 1 and Sabin 2 reference strains

recombination junctions of clinical strains, the recombination junctions of the in vitro model isolates are located in the stem region of the RNA structural elements. In Online Resources 10–12, the depiction of some recombination junctions of in vitro model isolates in RNA structural elements of 3D region is presented.

The third recombination junction of isolate V4 is located in region which is governed by many alternative base pairs of similar probability, and this is the reason why positive predictive value for the base pairs that lead to the formation of this RNA secondary structure element is low.

Discussion

In this study, in order to test the hypothesis that RNA structural elements are responsible for the distribution of certain types of recombination junctions in each one of the 2C and 3D (Sabin 3/Sabin 2 or Sabin 1 in 2C and Sabin 2/Sabin 1 or Sabin 3 in 3D) genomic regions of clinical isolates of poliovirus Sabin strains, we searched in 2C and 3D regions of reference Sabin strains for high probability RNA structural elements that could promote recombination. Furthermore, in order to gather more evidence about the role of the RNA structural elements in promoting

Table 6 RNA structural elements which are correlated with recombination junctions of the in vitro model isolates, and are identical between recombining partners

In vitro model isolate	Recombination type	Region	Base pairs (bp)	Base pair probability mean value	Positive predictive value (%)	Recombination junction location
V4	Sabin 2/Sabin 1	3D	S2: (6880–6887/6899–6892) S1: (6880–6887/6899–6892)	$1.11 \times e^{-0.03}$ 0.17	86.7 ± 8.6 76.6 ± 10.3	Single stranded–Stem
V4, V5, V6	Sabin 1/Sabin 2	3D	S2: (6827–6828/6837–6836) S1: (6821–6828/6843–6836)	0.52 0.13	<73 76.6 ± 10.3	Stem
V4	Sabin 1/Sabin 2	3D	S2: (6900–6904/6912–6908) S1: (6899–6902/6916–6913)	$2.87 \times e^{-0.02}$ 0.74	76.6 ± 10.3 <73	Stem–loop
V1, V2, V3	Sabin 1/Sabin 2	3D	S2: (7091–7096/7115–7100) S1: (7091–7096/7115–7100)	0.49 $1.79 \times e^{-0.03}$	<73 91 ± 5.9	Stem

Base pairs probabilities were obtained through RNAstructure, positive predictive values were obtained from Mathews [21]

recombination, we created a cell-culture-based in vitro model targeted in the production and identification of recombinant Sabin strains in 2C and 3D genomic regions. Results from clinical strains and in vitro model isolates of this study propose correlation between certain types of recombination junctions in specific regions, and similar RNA structural elements in both recombination partners, highlighting a prominent role for the stem region of the RNA structure elements in promoting recombination.

The results of the in vitro model about the distribution pattern of recombination types in 2C and 3D genomic regions did not correlate with the distribution pattern of recombination types in 2C and 3D genomic regions of the clinical strains of this study, as well of previous studies [10–15]. More specifically, in all strains of the in vitro model, we identified S1/S2 recombination junctions in 3D genomic region, while the majority of clinical strains in previous studies presented recombination junctions of type S2/Sx (x: S1 or S3) in 3D region. Nevertheless, we cannot reach to a safe conclusion about the distribution pattern of recombination junctions in 2C and 3D regions of the in vitro model strains due to the small number of in vitro model strains.

The results of this paper revealed that in clinical strains of this study as well as in clinical strains of previous studies [10–15], there is probably a correlation between recombination junctions of S3/Sx type (x: S2 or S1) and RNA secondary structure elements in 2C region, as well as between recombination junctions of S2/Sx type (x: S1 or S3) and RNA secondary structure elements in 3D region of the poliovirus genome.

More specifically, in 2C and 3D genomic regions, most of the recombination junctions of clinical strains were located in RNA secondary structure elements similar between the recombining partners. Of utmost importance is the fact that in all abovementioned cases, the RNA

structure element which is situated in the respective position of the third Sabin reference molecule, which molecule is not participating in the recombination, is entirely different from the identical RNA structure elements of the two Sabin reference molecules which are participating in the recombination. Moreover, these identical RNA secondary structure elements were consisted of high probability base pairs and also belonged to the minimum free energy structure. Consequently, these base pairs as well as the RNA secondary structure elements present high-positive predictive values [21]. Furthermore, all recombination junctions of the in vitro model isolates were identified in regions where RNA secondary structure elements similar between the two recombining partners were located. In addition, these RNA structure elements presented high-positive predictive values. This fact adds more evidence about the role of RNA secondary structure elements in promoting recombination. For these reasons, we propose that the distribution of certain types of recombination junctions in 2C and 3D genomic regions of poliovirus Sabin strains may not be fortuitous and may be promoted by RNA secondary structure elements which are similar between the recombining partners.

The most widely accepted mechanism of RNA recombination in poliovirus is the template switch mechanism [6] in which polymerase switches templates during negative strand synthesis. Romanova et al. [23] proposed a theoretical model for the polymerase template switching. According to this model, the two positive stranded RNA molecules (recombining partners) which act as templates during negative strand synthesis, form a heteroduplex in regions where self complementary sequences that form potential RNA structures exist. The viral RNA polymerase while copying one of these RNA molecules normally overcomes the heteroduplex region. However, it may pause or stop in regions of heteroduplex which are either single

stranded or partially base paired. The nascent 3' end of the complementary negative strand is detached in conjunction with RdRp from the first strand that is used as a template, and is attached to a homologous site of the second RNA molecule of the heteroduplex, which will now act as template for the continuation of the negative strand synthesis.

According to Romanova et al. [23], recombination junctions are located either in single-stranded regions or in the loops of the RNA secondary structure elements. In our study, the vast majority of the recombination junctions were located in the stem region of the RNA secondary structure elements (Tables 4, 5, 6; Figs. 2, 3; Online Resources 8–12). Furthermore, the RNA secondary structure models of Romanova et al. [23] are less reliable than the models presented here, since there were not algorithms that could reliably predict the secondary structure of RNA molecules at that time.

According to the abovementioned, we could propose a modification to the model of Romanova et al. [23] which is presented in Fig. 4. More specifically, two plus stranded RNA molecules form a heteroduplex, a hypothesis that is reinforced by recent studies which proved that the replication of poliovirus takes place in mixed replication complexes in which participate genomes that belong to different serotypes [8]. During the formation of the heteroduplex, base pairs are formed between the regions that belong to different recombining partners but have the potential to form high probability RNA structure elements which are similar between the recombining molecules. As a result, the heteroduplex in these regions is firmly established. It is probable that when the poliovirus RdRp in conjunction with the nascent negative strand reaches one of these firmly established regions of the heteroduplex, the 3' end of the nascent negative strand is detached from the first molecule which acts as template and is attached to the second molecule of the heteroduplex which will now act as a template. It is possible that the modest activity of 3AB poliovirus protein in destabilising the double stranded helix would be responsible for the detachment of the nascent negative strand from the first template-molecule, while the chaperone activity of poliovirus protein 3AB would be responsible for the attachment of the nascent negative strand to the second template-molecule [24]. Certainly, more experiments have to be done in order to fully clarify the proposed mechanism since there were regions bearing different RNA structural elements between the recombination partners, which also suffered recombination (Online Resources 6 and 7).

However, the abovementioned hypothesis is reinforced by results of previous studies concerning positive polarity RNA viruses. More specifically, Makino et al. [25] proposed that identical RNA secondary structure elements between parental molecules promote RNA recombination in murine coronavirus. Furthermore, Nagy and Bujarski

Fig. 4 Representation of the proposed modified model for the mechanism of RNA recombination in poliovirus. **a** High probability RNA secondary structure elements identical between the recombining molecules. **b** Formation of the heteroduplex in regions which have the potential to form high probability RNA secondary structure elements identical between the two recombining partners. **c** Detachment of the negative nascent strand from the template genome. **d** Attachment of the negative nascent strand to another template genome

[26] have proved that in brome mosaic virus both the incidence of recombination as well as the location of the recombination junctions depend on the stability of the heteroduplex.

It is worth mentioning the fact that all recombination junctions of the in vitro model isolates were located in 3D genomic region while no recombination junction was identified in 2C genomic region. Furthermore, all in vitro model isolates presented S1/S2 recombination junctions, which according to our knowledge is a rare event. This paradox could be explained if we consider that some isolates, which bare certain types of recombination junctions, may present higher fitness in cell-culture than in the body of the vaccinee. This hypothesis is reinforced by a previous study in which a PV1/PV2 capsid recombinant was isolated from cell-culture [8]. This recombination type in the capsid coding region has not been reported so far in clinical strains.

In addition, a question is raised from the fact that in certain regions of clinical and in vitro model isolates, with identical RNA structure elements between the parental strains, certain recombination types were identified (e.g., S3/S2 in C7, S1/S2 in V5). What could prevent recombination of the opposite type in the same region of these isolates (e.g., S2/S3 in C7, S2/S1 in V5)?

We propose that recombination junctions of the opposite type do not occur in the same region due to constraints concerning the communication and “cooperation” between the 5′ and 3′ ends of poliovirus isolates during replication of the genome. In other words, specific combinations of 5′ and 3′ termini might be preferred in recombination isolates, while the opposite combination of termini may reduce the replication efficiency of these isolates leading to lower fitness.

References

1. T. Pfister, E. Wimmer, C. Mirzayan, *Encyclopedia of Virology*, vol. 2e (Academic Press, San Diego, 1999)
2. O.M. Kew, R.W. Sutter, E.M. de Gourville, W.R. Dowdle, M.A. Pallansch, *Annu. Rev. Microbiol.* **59**, 587–635 (2005)
3. G.V. Gavrillin, E.A. Cherkasova, G.Y. Lipskaya, O.M. Kew, V.I. Agol, *J. Virol.* **74**, 7381–7390 (2000)
4. H.M. Liu, D.P. Zheng, L.B. Zhang, M.S. Oberste, M.A. Palansch, O.M. Kew, *J. Virol.* **74**, 11153–11161 (2000)
5. A.P. Gmyl, E.V. Belousov, S.V. Maslova, E.V. Khitrina, A.B. Chetverin, V.I. Agol, *J. Virol.* **73**, 8958–8965 (1999)
6. K. Kirkegaard, D. Baltimore, *Cell* **47**, 433–443 (1986)
7. A. Pierangeli, M. Bucci, M. Forzan, P. Pagnotti, M. Equestre, R. Perez Bercoff, *J. Gen. Virol.* **80**, 1889–1897 (1999)
8. D. Egger, K. Bienz, *J. Virol.* **76**, 10960–10971 (2002)
9. E.A. Cherkasova, E.A. Korotkova, M.L. Yakovenko, O.E. Ivanova, T.P. Ereemeeva, K.M. Chumakov, V.I. Agol, *J. Virol.* **76**, 6791–6799 (2002)
10. N.S. Cuervo, S. Guillot, N. Romanenkova, M. Combiescu, A. Aubert-Combiescu, M. Seghier, V. Caro, R. Crainic, *J. Virol.* **75**, 5740–5751 (2001)
11. I. Karakasiliotis, P. Markoulatos, T. Katsorchis, *Mol. Cell. Probes* **18**, 103–109 (2004)
12. E. Paximadi, I. Karakasiliotis, Z. Mamuris, C. Stathopoulos, V. Krikelis, P. Markoulatos, *Virus Genes* **32**, 203–210 (2006)
13. E. Paximadi, I. Karakasiliotis, E. Bolanaki, A. Krikelis, P. Markoulatos, *Virus Genes* **35**, 541–548 (2008)
14. E. Paximadi, I. Karakasiliotis, D. Papaventsis, G. Papageorgiou, P. Markoulatos, *J. Appl. Microbiol.* **104**, 1153–1162 (2007)
15. J. Martin, E. Samoilovich, G. Dunn, A. Lackenby, E. Feldman, A. Heath, E. Svirchevskaya, G. Cooper, M. Yermalovich, P.D. Minor, *J. Virol.* **76**, 10921–10928 (2002)
16. A. Georgopoulou, P. Markoulatos, N. Spyrou, N.C. Vamvakopoulos, *J. Clin. Microbiol.* **141**, 1047–1054 (2000)
17. I. Casas, L. Powell, P.E. Klapper, G.M. Cleator, *J. Virol. Methods* **53**, 25–36 (1995)
18. M.N. Mulders, G.Y. Lipskaya, H.G. van der Avoort, M.P. Koopmans, O.M. Kew, A.M. van Loom, *J. Infect. Dis.* **171**, 1399–1405 (1995)
19. V. Caro, S. Guillot, A. Candrea, F. Delpeyroux, R. Crainic, *J. Gen. Virol.* **82**, 79–91 (2001)
20. S. Guillot, V. Caro, N. Cuervo, E. Korotkova, M. Combiescu, A. Persu, A. Aubert-Combiescu, F. Delpeyroux, R. Crainic, *J. Virol.* **74**, 8434–8443 (2000)
21. D.H. Mathews, *RNA* **10**, 1178–1190 (2004)
22. M. Zuker, *Nucl. Acids Res.* **31**, 3406–3415 (2003)
23. L.I. Romanova, V.M. Blinov, E.A. Tolskaya, E.G. Viktorova, M.S. Kolesnikova, E.A. Guseva, V.I. Agol, *Virology* **155**, 202–213 (1986)
24. J.J. DeStefano, O. Titilope, *J. Virol.* **80**, 1662–1671 (2006)
25. S. Makino, J.G. Keck, S.A. Stohlman, M.M.C. Lai, *J. Virol.* **57**, 729–737 (1986)
26. P.D. Nagy, J.J. Bujarski, *Proc. Natl. Acad. Sci. USA* **90**, 6390–6394 (1993)