



HAL
open science

Protein-protein interactions leading to recruitment of the host DNA sliding clamp by the hyperthermophilic *Sulfolobus islandicus* rod-shaped virus 2

Andrew F. Gardner, Stephen D. Bell, Malcolm F. White, David Prangishvili,
Mart Krupovic

► **To cite this version:**

Andrew F. Gardner, Stephen D. Bell, Malcolm F. White, David Prangishvili, Mart Krupovic. Protein-protein interactions leading to recruitment of the host DNA sliding clamp by the hyperthermophilic *Sulfolobus islandicus* rod-shaped virus 2. *Journal of Virology*, 2014, 88 (12), pp.7105–7108. 10.1128/JVI.00636-14 . pasteur-01977399

HAL Id: pasteur-01977399

<https://pasteur.hal.science/pasteur-01977399>

Submitted on 10 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

Protein-protein interactions leading to the recruitment of the host DNA sliding clamp by the hyperthermophilic archaeal virus SIRV2

Andrew F. Gardner^a, Stephen D. Bell^b, Malcolm F. White^c, David Prangishvili^{d#}, Mart Krupovic^{d#}

^a New England Biolabs, Ipswich, MA 01938, USA

^b Molecular and Cellular Biochemistry Department, Biology Department, Indiana University, Bloomington, Indiana 47405, USA

^c Biomedical Sciences Research Complex, University of St Andrews, Fife KY16 9ST, UK

^d Institut Pasteur, Unité Biologie Moléculaire du Gène chez les Extrêmophiles, Département de Microbiologie, 75015 Paris, France

Running title: Recruitment of PCNA by SIRV2

Corresponding authors:

Mart Krupovic

E-mail: krupovic@pasteur.fr

David Prangishvili

E-mail: david.prangishvili@pasteur.fr

Word count for the abstract: 71

Word count for the text: 1,513

42

43 **Abstract**

44 Viruses infecting hyperthermophilic archaea typically do not encode DNA polymerases,
45 raising questions regarding their genome replication. Here, using a yeast two-hybrid
46 approach we have assessed interactions between proteins of *Sulfolobus islandicus* rod-
47 shaped virus 2 (SIRV2) and the host-encoded proliferating cell nuclear antigen (PCNA),
48 a key DNA replication protein in archaea. Five SIRV2 proteins were found to interact
49 with PCNA, providing insights into the recruitment of host replisome for viral DNA
50 replication.

51

52

53 Viruses infecting Archaea, the third domain of life, constitute one of the most enigmatic
54 sections of the virosphere. Archaeal viruses, especially those infecting
55 hyperthermophilic hosts thriving at temperatures above 70°C, are extremely diverse
56 both morphologically and genetically (27, 28). The vast majority of them are unique to
57 Archaea and do not resemble viruses infecting Bacteria or Eukaryotes. Furthermore,
58 during the past few years it became apparent that the ways hyperthermophilic archaeal
59 viruses interact with their hosts are also often unprecedented in the viral world (2, 3).
60 Indeed, proteins encoded by these viruses typically do not share similarity with proteins
61 in the sequence databases (28), often display new structural folds (18), and play
62 unexpected roles in the viral life cycles (31). An interesting feature of hyperthermophilic
63 archaeal viruses is the general absence of recognizable DNA polymerase genes in their
64 genomes; among the 41 virus isolates for which genome sequences are available, only
65 one encodes a DNA polymerase (26), raising a question as of how genome replication
66 in these viruses is achieved. To answer this puzzling question, in the present study we
67 have investigated how *Sulfolobus islandicus* rod-shaped virus 2 (SIRV2) recruits the
68 cellular machinery for the replication of its genome.

69

70 SIRV2 is one of the most extensively studied archaeal viruses (29) and is the type
71 species of the family *Rudiviridae* within the order *Ligamenvirales* (30). The rod-shaped
72 SIRV2 virions are non-enveloped and are formed from linear double-stranded DNA
73 genomes coated with one major and three minor structural proteins. SIRV2 infects
74 acidophilic hyperthermophiles of the archaeal order *Sulfolobales*, namely *S. islandicus*
75 (12) and *S. solfataricus* (23). The infection cycle starts with rapid virus adsorption to the
76 pili-like structures on the host cell surface, which eventually leads to genome uncoating
77 and internalization (33). Transcription of the viral genes commences within the first few
78 minutes of infection (32) and is followed by the efficient replication of the viral DNA (2).
79 SIRV2 virions, assembled in the cytoplasm, are released from the host ~8–10 hours
80 post infection through large virus-encoded pyramidal structures (2). One of the least
81 understood aspects of SIRV2 infection cycle is genome replication in the absence of
82 virus-encoded DNA polymerase. The only proteins, which might be involved in the

83 replication of SIRV2 genome, are the Holliday junction resolvase P121 (9) and the
84 endonuclease P119c (22) related to enzymes involved in the initiation of the rolling-
85 circle replication of various plasmids and single-stranded DNA viruses (11, 15).
86 However, the exact steps of SIRV2 genome replication in which the two enzymes
87 participate as well as the involvement of cellular players in this process remain unclear.

88

89 Viruses that do not encode their own DNA polymerases rely of the replication machinery
90 of the host. However, viral proteins often play an important role in directing the
91 replisome to the viral replication origins. Indeed, a considerable number of euryarchaeal
92 viruses encode proteins involved in the initiation of DNA replication, including replicative
93 MCM helicases, Cdc6/Orc1 or PCNA (proliferating cell nuclear antigen) homologues
94 (10, 14, 16, 17, 24, 34, 36). Since rudiviruses encode neither a DNA polymerase nor any
95 of the above mentioned proteins involved in the initiation of DNA replication, we
96 hypothesized that their genome replication should rely on the physical recruitment of the
97 host replisome. To verify this possibility, we set out to investigate the interactions
98 between SIRV2 proteins and the heterotrimeric *S. solfataricus* sliding clamp
99 (SsoPCNA1–3) (7, 38). The latter was selected as a likely target because it is a key
100 protein of DNA replication and repair in archaea and eukaryotes, allowing non-
101 sequence-specific enzymes, such as replicative DNA polymerase PolB1, DNA ligase
102 Lig1 and flap endonuclease FEN1, to associate with their DNA substrates (1, 20, 25).
103 To explore which of the SIRV2 proteins might be involved in the recruitment of PCNA
104 and, by extension, of the entire replisome, we employed yeast two-hybrid (Y2H)
105 analysis. SsoPCNA subunits 1 (NP_341936), 2 (NP_342519), and 3 (NP_341944) were
106 cloned into the 'bait' vector (pGBKT7) encoding GAL4 DNA-binding domain, while the
107 viral 'prey' library was created by cloning the PCR-amplified SIRV2 genes into the
108 NdeI/XmaI site of the pGADT7 vector encoding the GAL4 activation domain. All clones
109 were verified by DNA sequencing. *Saccharomyces cerevisiae* AH109 (Clontech) was
110 sequentially transformed with prey and bait plasmids and selected on nutritional media
111 lacking either leucine or tryptophan, respectively. Controls with empty bait or prey
112 vectors were performed to ensure that the fusion protein could not induce expression of
113 selection gene in the absence of a protein partner. In the case of interaction between
114 the bait and prey, the fusion protein is expected to activate the HIS3 expression and
115 complement histidine auxotrophy, allowing growth on yeast minimal media lacking
116 histidine.

117

118 Our Y2H screen has revealed five SIRV2-encoded proteins that interacted with the
119 SsoPCNA: proteins P105a (NP_666544) and P84c (NP_666565) were found to interact
120 with SsoPCNA1, whereas proteins P83a/b (NP_666535/NP_666588), P84c, P119a
121 (NP_666536) and P121 (NP_666569) with SsoPCNA3 (Fig. 1). Notably, no interaction
122 between viral proteins and SsoPCNA2 could be detected. Two of the PCNA-interacting
123 SIRV2 proteins (P84c and P121) are conserved in all rudiviruses, two (P83a/b and

124 P105a) are encoded in a smaller subset of more closely related viruses, while P119a is
125 restricted to SIRV2 (Table 1).

126

127 PCNA is known as a “molecular tool-belt” (38), which interacts with multiple proteins
128 involved in DNA replication and repair (1, 20, 25). These interactions are often mediated
129 via the PCNA-interacting protein (PIP) box, with the consensus sequence of
130 Qxxhxx@@, where ‘h’ represents hydrophobic amino acid residues, ‘@’ corresponds to
131 bulky aromatic residues, and ‘x’ is any residue (37). Sequence analysis of the viral
132 proteins, which were found to interact with SsoPCNA1 and SsoPCNA3, revealed the
133 presence of potential PIP-boxes in all proteins except for P84c (Fig. 2). Interestingly,
134 P119a was found to contain three PIP-boxes (Fig. 2D). Notably, whereas the presence
135 of PIP boxes is suggestive of interaction with PCNA, the absence of the identifiable
136 motif does not necessarily signify the reverse (i.e., lack of interaction), because (i) the
137 exact sequence of PIP boxes is known to vary (37), and (ii) interactions might be
138 mediated by motifs other than PIP (39). As has been previously observed for PCNA-
139 interacting proteins from other viruses (13), the PIP-boxes of SIRV2 displayed variable
140 correspondence to the consensus sequence, which might be important for orchestrating
141 the sequential binding of different protein partners to PCNA.

142

143 To more critically scrutinize the predicted PIP boxes, we have investigated their exact
144 location within the four PCNA-interacting viral proteins. For this purpose we have
145 constructed structural homology models using Modeller v9.11 (19) for three of the
146 SIRV2 proteins for which structural counterparts could be identified (Fig. 2A-C). This
147 analysis suggested that PIP box of protein P121 is unlikely to be functional, since it is
148 located within the core of the protein fold (Fig. 2C), and thus a different protein region
149 must be responsible for the observed PCNA binding. Similarly, two of the three PIP
150 boxes initially predicted in P119a are found within the secondary structure elements,
151 which are likely to be integral for the tertiary protein structure. The PIP box of P83a/b is
152 located in the loop region between the two hairpin-forming β -strands (Fig. 2A). By
153 contrast, PIP box of P105a and PIP1 box of P119a are located at the extremities of the
154 proteins, lending credence to their functionality.

155

156 Out of five PCNA-interacting SIRV2 proteins only one has been biochemically
157 characterized; protein P121 is a Holliday junction resolvase (Hjr), which is implicated in
158 the resolution of viral genome concatamers, a key step in SIRV2 genome replication (9).
159 Interestingly, SsoPCNA has been previously shown to bind the Hjr of *S. solfataricus* and
160 stimulate its enzymatic activity (8), suggesting that a similar effect might be exerted
161 during the interaction with the viral Hjr. Another interesting protein, which has been
162 identified as an SsoPCNA partner in our Y2H analysis, is P83a/b (genes *ORF83a* and
163 *ORF83b* are located in inverted terminal repeats and encode identical proteins). RNA-
164 seq analysis has shown that *ORF83a/b* transcripts are overwhelmingly dominant

165 starting with the first minutes of infection and remain abundant throughout the infection
 166 (32), suggesting an important role for P83a/b during the replication cycle of SIRV2. In
 167 the course of a structural genomics initiative the X-ray structure of P83a/b homolog from
 168 a closely related virus SIRV1 has been solved (21). The protein was found to form a
 169 hexameric ring and contain the helix-turn-helix DNA-binding domain, suggesting that it
 170 might be involved in DNA metabolism (18). Notably, previous Y2H screen for
 171 interactions among SIRV2 proteins revealed that P83a/b interacts with the viral Hjr,
 172 P121 (32), suggesting that P83a/b, P121 and SsoPCNA form a tripartite complex (Fig.
 173 1B). Whereas *ORF83a/b* transcripts are abundant from the beginning of infection, those
 174 of *ORF121* appear later in infection, suggesting that the intricate interplay between viral
 175 proteins and the cellular PCNA might be regulated by both hierarchical strengths of PIP-
 176 box-mediated protein-protein interactions and transcriptional control of viral gene
 177 expression. Future studies will focus on revealing the molecular and structural details of
 178 the interactions between the host PCNA subunits and the five SIRV2 proteins identified
 179 here, which will clarify how SIRV2 hijacks the replisome of its host.

180
 181
 182
 183

184 References

- 185 1. **Beattie, T. R., and S. D. Bell.** 2011. Molecular machines in archaeal DNA replication. *Curr Opin*
 186 *Chem Biol* **15**:614-619.
- 187 2. **Bize, A., E. A. Karlsson, K. Ekefjard, T. E. Quax, M. Pina, M. C. Prevost, P. Forterre, O. Tenaillon,**
 188 **R. Bernander, and D. Prangishvili.** 2009. A unique virus release mechanism in the Archaea. *Proc*
 189 *Natl Acad Sci U S A* **106**:11306-11311.
- 190 3. **Brumfield, S. K., A. C. Ortmann, V. Ruigrok, P. Suci, T. Douglas, and M. J. Young.** 2009. Particle
 191 assembly and ultrastructural features associated with replication of the lytic archaeal virus
 192 *Sulfolobus* turreted icosahedral virus. *J Virol* **83**:5964-5970.
- 193 4. **Cuff, J. A., M. E. Clamp, A. S. Siddiqui, M. Finlay, and G. J. Barton.** 1998. JPred: a consensus
 194 secondary structure prediction server. *Bioinformatics* **14**:892-893.
- 195 5. **Dionne, I., and S. D. Bell.** 2005. Characterization of an archaeal family 4 uracil DNA glycosylase
 196 and its interaction with PCNA and chromatin proteins. *Biochem J* **387**:859-863.
- 197 6. **Dionne, I., N. J. Brown, R. Woodgate, and S. D. Bell.** 2008. On the mechanism of loading the
 198 PCNA sliding clamp by RFC. *Mol Microbiol* **68**:216-222.
- 199 7. **Dionne, I., R. K. Nookala, S. P. Jackson, A. J. Doherty, and S. D. Bell.** 2003. A heterotrimeric
 200 PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Cell* **11**:275-282.
- 201 8. **Dorazi, R., J. L. Parker, and M. F. White.** 2006. PCNA activates the Holliday junction
 202 endonuclease Hjc. *Journal of molecular biology* **364**:243-247.
- 203 9. **Gardner, A. F., C. Guan, and W. E. Jack.** 2011. Biochemical characterization of a structure-
 204 specific resolving enzyme from *Sulfolobus islandicus* rod-shaped virus 2. *PLoS One* **6**:e23668.
- 205 10. **Gorlas, A., E. V. Koonin, N. Bienvenu, D. Prieur, and C. Geslin.** 2012. TPV1, the first virus
 206 isolated from the hyperthermophilic genus *Thermococcus*. *Environ Microbiol* **14**:503-516.
- 207 11. **Ilyina, T. V., and E. V. Koonin.** 1992. Conserved sequence motifs in the initiator proteins for
 208 rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and
 209 archaeobacteria. *Nucleic acids research* **20**:3279-3285.

- 210 12. **Jaubert, C., C. Danioux, J. Oberto, D. Cortez, A. Bize, M. Krupovic, Q. She, P. Forterre, D.**
 211 **Prangishvili, and G. Sezonov.** 2013. Genomics and genetics of *Sulfolobus islandicus* LAL14/1, a
 212 model hyperthermophilic archaeon. *Open Biol* **3**:130010.
- 213 13. **Kazlauskas, D., and C. Venclovas.** 2011. Computational analysis of DNA replicases in double-
 214 stranded DNA viruses: relationship with the genome size. *Nucleic acids research* **39**:8291-8305.
- 215 14. **Klein, R., U. Baranyi, N. Rossler, B. Greineder, H. Scholz, and A. Witte.** 2002. *Natrialba magadii*
 216 virus phiCh1: first complete nucleotide sequence and functional organization of a virus infecting
 217 a haloalkaliphilic archaeon. *Mol Microbiol* **45**:851-863.
- 218 15. **Krupovic, M.** 2013. Networks of evolutionary interactions underlying the polyphyletic origin of
 219 ssDNA viruses. *Current opinion in virology* **3**:578-586.
- 220 16. **Krupovic, M., P. Forterre, and D. H. Bamford.** 2010. Comparative analysis of the mosaics
 221 genomes of tailed archaeal viruses and proviruses suggests common themes for virion
 222 architecture and assembly with tailed viruses of bacteria. *Journal of molecular biology* **397**:144-
 223 160.
- 224 17. **Krupovic, M., S. Gribaldo, D. H. Bamford, and P. Forterre.** 2010. The evolutionary history of
 225 archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile
 226 genetic elements. *Mol Biol Evol* **27**:2716-2732.
- 227 18. **Krupovic, M., M. F. White, P. Forterre, and D. Prangishvili.** 2012. Postcards from the edge:
 228 structural genomics of archaeal viruses. *Adv Virus Res* **82**:33-62.
- 229 19. **Marti-Renom, M. A., A. C. Stuart, A. Fiser, R. Sanchez, F. Melo, and A. Sali.** 2000. Comparative
 230 protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **29**:291-325.
- 231 20. **Moldovan, G. L., B. Pfander, and S. Jentsch.** 2007. PCNA, the maestro of the replication fork.
 232 *Cell* **129**:665-679.
- 233 21. **Oke, M., L. G. Carter, K. A. Johnson, H. Liu, S. A. McMahon, X. Yan, M. Kerou, N. D. Weikart, N.**
 234 **Kadi, M. A. Sheikh, S. Schmelz, M. Dorward, M. Zawadzki, C. Cozens, H. Falconer, H. Powers, I.**
 235 **M. Overton, C. A. van Niekerk, X. Peng, P. Patel, R. A. Garrett, D. Prangishvili, C. H. Botting, P.**
 236 **J. Coote, D. T. Dryden, G. J. Barton, U. Schwarz-Linek, G. L. Challis, G. L. Taylor, M. F. White,**
 237 **and J. H. Naismith.** 2010. The Scottish Structural Proteomics Facility: targets, methods and
 238 outputs. *J Struct Funct Genomics* **11**:167-180.
- 239 22. **Oke, M., M. Kerou, H. Liu, X. Peng, R. A. Garrett, D. Prangishvili, J. H. Naismith, and M. F.**
 240 **White.** 2011. A dimeric Rep protein initiates replication of a linear archaeal virus genome:
 241 implications for the Rep mechanism and viral replication. *J Virol* **85**:925-931.
- 242 23. **Okutan, E., L. Deng, S. Mirlashari, K. Uldahl, M. Halim, C. Liu, R. A. Garrett, Q. She, and X. Peng.**
 243 2013. Novel insights into gene regulation of the rudivirus SIRV2 infecting *Sulfolobus* cells. *RNA*
 244 *Biol* **10**:875-885.
- 245 24. **Pagaling, E., R. D. Haigh, W. D. Grant, D. A. Cowan, B. E. Jones, Y. Ma, A. Ventosa, and S.**
 246 **Heaphy.** 2007. Sequence analysis of an Archaeal virus isolated from a hypersaline lake in Inner
 247 Mongolia, China. *BMC Genomics* **8**:410.
- 248 25. **Pan, M., L. M. Kelman, and Z. Kelman.** 2011. The archaeal PCNA proteins. *Biochem Soc Trans*
 249 **39**:20-24.
- 250 26. **Peng, X., T. Basta, M. Haring, R. A. Garrett, and D. Prangishvili.** 2007. Genome of the *Acidianus*
 251 bottle-shaped virus and insights into the replication and packaging mechanisms. *Virology*
 252 **364**:237-243.
- 253 27. **Prangishvili, D.** 2013. The wonderful world of archaeal viruses. *Ann Rev Microbiol* **67**:565-585.
- 254 28. **Prangishvili, D., R. A. Garrett, and E. V. Koonin.** 2006. Evolutionary genomics of archaeal
 255 viruses: unique viral genomes in the third domain of life. *Virus Res* **117**:52-67.
- 256 29. **Prangishvili, D., E. V. Koonin, and M. Krupovic.** 2013. Genomics and biology of Rudiviruses, a
 257 model for the study of virus-host interactions in Archaea. *Biochem Soc Trans* **41**:443-450.

- 258 30. **Prangishvili, D., and M. Krupovic.** 2012. A new proposed taxon for double-stranded DNA
 259 viruses, the order "Ligamenvirales". *Arch Virol* **157**:791-795.
- 260 31. **Quax, T. E., M. Krupovic, S. Lucas, P. Forterre, and D. Prangishvili.** 2010. The *Sulfolobus* rod-
 261 shaped virus 2 encodes a prominent structural component of the unique virion release system in
 262 Archaea. *Virology* **404**:1-4.
- 263 32. **Quax, T. E., M. Voet, O. Sismeiro, M. A. Dillies, B. Jagla, J. Y. Coppee, G. Sezonov, P. Forterre, J.**
 264 **van der Oost, R. Lavigne, and D. Prangishvili.** 2013. Massive activation of archaeal defense
 265 genes during viral infection. *J Virol* **87**:8419-8428.
- 266 33. **Quemin, E. R., S. Lucas, B. Daum, T. E. Quax, W. Kuhlbrandt, P. Forterre, S. V. Albers, D.**
 267 **Prangishvili, and M. Krupovic.** 2013. First insights into the entry process of hyperthermophilic
 268 archaeal viruses. *J Virol* **87**:13379-13385.
- 269 34. **Raymann, K., P. Forterre, C. Brochier-Armanet, and S. Gribaldo.** 2014. Global phylogenomic
 270 analysis disentangles the complex evolutionary history of DNA replication in archaea. *Genome*
 271 *Biol Evol* **6**:192-212.
- 272 35. **Roberts, J. A., S. D. Bell, and M. F. White.** 2003. An archaeal XPF repair endonuclease
 273 dependent on a heterotrimeric PCNA. *Mol Microbiol* **48**:361-371.
- 274 36. **Senčilo, A., D. Jacobs-Sera, D. A. Russell, C. C. Ko, C. A. Bowman, N. S. Atanasova, E. Osterlund,**
 275 **H. M. Oksanen, D. H. Bamford, G. F. Hatfull, E. Roine, and R. W. Hendrix.** 2013. Snapshot of
 276 haloarchaeal tailed virus genomes. *RNA Biol* **10**:803-816.
- 277 37. **Warbrick, E.** 2000. The puzzle of PCNA's many partners. *Bioessays* **22**:997-1006.
- 278 38. **Williams, G. J., K. Johnson, J. Rudolf, S. A. McMahon, L. Carter, M. Oke, H. Liu, G. L. Taylor, M.**
 279 **F. White, and J. H. Naismith.** 2006. Structure of the heterotrimeric PCNA from *Sulfolobus*
 280 *solfataricus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**:944-948.
- 281 39. **Xu, H., P. Zhang, L. Liu, and M. Y. Lee.** 2001. A novel PCNA-binding motif identified by the
 282 panning of a random peptide display library. *Biochemistry* **40**:4512-4520.
- 283

284 **Figure legends**

285 **Figure 1.** Results of the yeast two-hybrid interactions between SIRV2 proteins and
286 SsoPCNA. (A) An example of yeast two-hybrid analysis demonstrating the interaction
287 between the SIRV2 protein P83a/b and SsoPCNA subunit 3. VEC, empty vector.
288 Interaction between the bait (PCNA3) and prey (P83a/b) activates the *HIS3* expression,
289 complementing histidine auxotrophy and allowing growth on yeast minimal media
290 lacking histidine (-HIS). (B) Schematic representation of the SsoPCNA interactome.
291 SIRV2 proteins found to interact with SsoPCNA1 and 3 are represented by black
292 spheres, while DNA replication and repair proteins of *S. solfataricus* previously reported
293 to bind SsoPCNA are shown as grey spheres. Hjc, Holliday junction endonuclease (8);
294 FEN1, flap structure-specific endonuclease 1 (7); Dpo4, DNA polymerase IV (6); RFC-S
295 and -L, small and large subunits of the replication factor C, respectively (6); PolB1,
296 replicative DNA polymerase (7); UDG1, uracil DNA glycosylase (5); Lig1, DNA ligase
297 (7); XPF, nucleotide excision repair endonuclease (35).

298
299 **Figure 2.** Potential PCNA-interacting protein (PIP) boxes in rudiviral proteins. (A-C)
300 Structural homology models of three PCNA-interacting SIRV2 proteins for which
301 structural homologues could be identified. The locations of the regions corresponding to
302 predicted PIP boxes are encircled. Qxxhxx@@ is the PIP consensus sequence, where
303 'h' represents hydrophobic amino acid residues, '@' corresponds to bulky aromatic
304 residues, and 'x' is any residue (37). (A) P83a/b model was built using as a template the
305 X-ray structure of SIRV1 protein P56a (PDB ID: 2X48). (B) P105a model was built using
306 as a template the X-ray structure of lipothrixvirus SIFV protein ORF14 (PDB ID: 2H36).
307 C. P121 model was built using as a template the X-ray structure of *Sulfolobus*
308 *solfataricus* Holliday junction resolvase (PDB ID: 1HH1). D. The locations of the three
309 PIP boxes in P119a. The secondary structure elements predicted using JPred (4) are
310 indicated above the sequence (H, α -helixes, E, β -strands).

311

312 **Table 1.** Conservation of SIRV2 PCNA-interacting proteins in other viruses.

SIRV2	SIRV1	SRV	ARV1	SMRV1	Characteristics
P83a (NP_666535)	gp01	P57	–	–	Helix-turn-helix protein
P84c (NP_666565)	gp23	P75	gp22	gp22	C-terminal coiled-coil domain
P105a (NP_666544)	gp03	–	–	–	Novel fold; homologues in several lipothrixviruses and fusellovirus SSV6
P119a (NP_666536)	–	–	–	–	No homologues
P121 (NP_666569)	gp27	P116c	gp25	gp19	Holliday junction resolvase; homologues in numerous bacterial and archaeal Caudovirales

313 Genome accession numbers: SIRV2 (NC_004086); SIRV1 (NC_004087); SRV, *Stygiolobus* rod-shaped
 314 virus (FM164764); ARV1, *Acidianus* rod-shaped virus 1 (NC_009965); SMRV1, *Sulfolobales* Mexican
 315 ruidivirus 1 (NC_019413).



