A novel rudivirus, ARV1, of the hyperthermophilic archaeal
genus *Acidianus*\(^c\)\(^\star\)

Gisle Vestergaard\(^a\), Monika Häring\(^b,c\), Xu Peng\(^a\), Reinhard Rachel\(^b\),
Roger A. Garrett\(^a\), David Prangishvili\(^b,c\),\(^*\)

\(^a\)Danish Archaea Centre, Institute of Molecular Biology, Copenhagen University, Sølvgade 83H, DK-1307 Copenhagen K, Denmark
\(^b\)Department of Microbiology, Archaea Centre, University of Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany
\(^c\)Institut Pasteur, rue Dr. Roux 25, 75724 Paris Cedex 15, France

Received 31 January 2005; returned to author for revision 17 February 2005; accepted 24 February 2005
Available online 8 April 2005

Abstract

Virus ARV1, the first member of the family *Rudiviridae* infecting hyperthermophilic archaea of the genus *Acidianus*, was isolated from a hot spring in Pozzuoli, Italy. The rod-shaped virions, 610 ± 50 nm long and 22 ± 3 nm wide, are non-enveloped and carry a helical nucleoprotein core, with three tail fibers protruding at each end which appear to be involved in adsorption onto the host cell surface. The virions contain two protein components, a major one of 14.4 kDa, which is glycosylated and a minor of about 124 kDa. The linear double-stranded DNA genome yielded 24,655 bp of sequence, including 1365 bp inverted terminal repeats. Coding is on both strands and about 40% of the predicted genes are homologous to those of other hyperthermophilic crenarchaeal viruses, mainly rudiviruses. They include genes encoding the coat protein, two glycosyl transferases and a Holliday junction resolvase. Other assigned functions include a thymidylate synthase and three DNA-binding proteins. The genome sequence and composition differ strongly from those of the *Sulfolobus* rudiviruses SIRV1 and SIRV2, and the genome stability is very high, with no sequence variants being detected. Although the sequences of the inverted terminal repeats of the three rudiviruses are different, they all carry the motif AATTTAGGAATTTAGGAATTT near the genome ends which may constitute a signal for the Holliday junction resolvase and DNA replication.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Virus; Archaea; Hyperthermophile; Linear DNA genome

Introduction

Recent studies of geothermally heated aquatic ecosystems have revealed the presence of widely divergent viruses with unusual morphotypes, most of which have never been observed before in nature (Geslin et al., 2003; Prangishvili et al., 2001; Rachel et al., 2002a; Rice et al., 2001). All viruses isolated so far from such hot habitats have double-stranded DNA genomes, linear or circular, and infect members of the kingdom Crenarchaeota of the Archaea domain. On the basis of their unusual morphotypes they have been classified into 5 new viral families: (i) the *Fuselloviridae*, including spindle-shaped SSV1, SSV2, SSV3 (reviewed by Prangishvili et al., 2001), SSK1, and SSVRH (Wiedenheft et al., 2004); (ii) the *Lipothrixviridae*, including filamentous TTV1, TTV2, TTV3 (Janekovic et al., 1983), SIFV (Arnold et al., 2000a), AFV1 (Bettstetter et al., 2003), and AFV2 (Haring et al., in press); (iii) the *Guttaviridae*, including droplet-shaped SNDV (Arnold et al., 2000b); (iv) the *Globuloviridae*, including spherical PSV (Haring et al., 2004); and (v) the *Rudiviridae*, including rod-shaped SIRV1 and SIRV2 (reviewed by Prangishvili et al., 2001). The

\(^*\) Sequence data from this article are deposited with the EMBL/GenBank Data Libraries under accession nos. AJ875026 and AJ884675.

\(^\star\) Corresponding author. Institut Pasteur, rue Dr. Roux 25, 75724 Paris Cedex 15, France. Fax: +33 01 45 68 8834.
E-mail address: prangish@pasteur.fr (D. Prangishvili).
unusual icosahedral virus STIV is also likely to be classified into a separate family (Rice et al., 2004).

This morphological diversity is reflected in the genomes of members of different families. Sequence similarities between their genes are generally very limited, and moreover most predicted genes yield no other matches in public sequence databases (Prangishvili and Garrett, 2004). In addition, some genomes code almost exclusively on one DNA strand while others code on both DNA strands (Häring et al., 2004). Many linear genomes terminate in repeat structures of different types. For example, the rudiviruses carry large inverted terminal repeats (Blum et al., 2001; Peng et al., 2001), the globulovirus PSV exhibits short terminal inverted repeat structures (Häring et al., 2004), and the lipothrixivirus AFV1 has multiple short direct repeats near the termini, reminiscent of eukaryal telomeric structures, with 11 consecutive G-C base pairs at each end (Bettstetter et al., 2003). Some genomic termini are also covalently modified (Arnold et al., 2000a; Blum et al., 2001; Peng et al., 2004). Major differences are also observed in the stability of the crenarchaeal genomes. While most are relatively stable, showing only local regions with minor sequence variations, mostly point mutations, variants of the rudivirus SIRV1 have been detected which exhibit differences in gene content, gene size, and gene order, and they carry putative 12 bp introns which may facilitate changes in the variant genomes (Prangishvili et al., 1999; Peng et al., 2004).

Little is known about genetic diversity or geographical distribution of members of crenarchaeal viral families. Although several members of the Fuselloviridae and Lipothrixviridae have been isolated and compared from geographically diverse locations, the Guttaviridae and Globuloviridae are each represented by a single viral species, and the Rudiviridae contain two highly similar species from Iceland which infect the same strains of the genus Sulfolobus. Here we report on the isolation and characterization of a further novel member of the family Rudiviridae, which infects the genus Acididius and was discovered in a hot acidic spring at Pozzuoli, Italy.

Results

Virus hosts and virus purification

A sample was collected from a hot, acidic spring (87–93 °C, pH 1.5–2) submerged in the major water reservoir in the crater of the Solfatara volcano at Pozzuoli, Italy, and an enrichment culture was established under conditions which are favorable for growth of members of the Sulfolobales that predominate in such environments (Zillig et al., 1994). Rod-shaped particles, resembling rudiviruses, were observed in the enrichment culture by transmission electron microscopy (TEM). After serial dilution of the original enrichment culture by 1:1000, they were still present which indicated that the infected host strain was growing. In order to isolate it, 200 single strains were colony-purified from the enrichment on Gelrite plates. The purified strains could be assigned to the hyper-thermophilic archaeal genus Acididius on the basis of its 16S rDNA sequence. All strains were grown to late exponential phase before analyzing the cell-free supernatant for the presence of virus particles. This led to the identification of a virus-producing strain Acididius sp. Acii26. The virus was isolated from the supernatant of the cell-free culture of this strain by precipitation with polyethylene glycol and it was purified by isopycnic gradient centrifugation in CsCl. This yielded a sharp bluish-white opalescent band with a buoyant density of about 1.33 g ml⁻¹.

The host range was examined by adding purified virus particles to growing cultures of novel Acididius isolates from Pozzuoli and known members of the genera Acididius and Sulfolobus: Acididius ambivalens, Acididius brierleyi, Acididius infernus, Sulfolobus solfataricus, and Sulfolobus islandicus. Propagation of the virus was followed by TEM and was only observed in two strains isolated from the enrichment culture, Acididius pozzuolensis and Acididius sp. AD1. The two strains were closely related. Their 16S rRNA sequences are 99% identical, and both showed 99% identity to the 16S rRNA sequence of the closest known relative, A. brierleyi. The new virus was named Acididius rod-shaped virus 1, ARV1.

Virus structure

Virions of ARV1 are flexible rod-shaped particles measuring 610 ± 50 nm by 22 ± 3 nm. At each end of the virion, three terminal fibers can be resolved, each about 10 nm in length and about 3 nm in width (Figs. 1A and B). Particle staining was not uniform along the length and often a dark broken line was observed along the center. Moreover, the distribution pattern of these dark segments was different for each virus particle (Fig. 1B). In order to gain more insight into the virion structure, electron tomography was performed (Figs. 2A–C). In the central section of the tomographic data set (Fig. 2B), the terminal fibers can be seen to be anchored in the virus particle in plug-like termini, which are approximately 50 nm in length. The middle part of the virion shows two different stain distributions, most likely reflecting different conformations, one with a strongly stained central area, while in the other the stain is distributed more evenly. Careful inspection of the 3D data set led us to conclude that, despite variations in stain distribution, the subunits or building blocks of the virion are arranged helically throughout most or all of the particle.

Since no pili- or flagella-like appendages were observed on the host cells, we infer that the virus interacts directly with the cellular surface. When virions were mixed separately with the S-layer and membrane vesicles of the host (prepared as described in Materials and methods), the
ends of the virions were observed attached to the membrane vesicles (Fig. 3). Therefore, we infer that the virion tail fibers participate in adsorption.

**Virus–host relationships**

None of the hosts for ARV1 produced lawns on Gelrite plates under conditions similar to those that were successful for *Sulfolobus* species (Zillig et al., 1994). Moreover, varying the composition of the growth medium or the density of the lawn inoculum did not help. Therefore, the virus titer could only be estimated by comparing by TEM the count of ARV1 virions with the count of virions of the nidovirus SIRV1 of *S. islandicus* for which the titer was determined by plaque assay (Häring et al., 2004).

During a course of productive infection of *A. pozzuoliensis* with ARV1, at a ratio of virions to cells of about 10, no significant amount of cell debris was formed nor was there any decrease in cell density (Fig. 4). This implied that replication of the virus did not induce cell lysis. However, the growth rate of the host cells was retarded on infection and infected cells did not reach the same density as the non-infected cells (Fig. 4).

The estimated ARV1 titer in growing host cell cultures was invariably lower ($\leq 10^5$ particles per ml) than that of the
**Sulfolobus** rudiviruses SIRV1 and SIRV2 (≤10⁹ particles per ml). This result was reinforced by TEM visualization of extruding virus particles. In the late logarithmic growth phase, one extruding ARV1 virion was observed, on average, for each cell of *Acidianus pozzuoliensis* (Figs. 5A and B), whereas under similar growth conditions, at least 10 virions were seen extruding from each SIRV1-infected cell of *S. islandicus* REN2H1 (Fig. 5C).

**Viral proteins**

The protein pattern of ARV1 on an SDS–polyacrylamide gel (SDS–PAGE) after Coomassie brilliant blue (CBB) staining showed one major band with an estimated molecular mass of 20 kDa (Fig. 6A). Although the band was broad, N-terminal sequencing of protein eluted from three horizontal sections of the band yielded the same N-terminal sequence, AKGHTPRSYSQQYS, indicating the presence of a single protein component. The end sequence corresponds to a 14.4-kDa protein encoded in the genome (ORF134b, see below), and we infer that the discrepancy with the gel size estimate (Fig. 6A) reflects that the protein is both very basic and glycosylated (see below). A similar discrepancy was observed for the coat proteins of the *Sulfolobus* rudiviruses (Prangishvili et al., 1999; Peng et al., 2001). A weak upper band visible on the gel at about 116 kDa (Fig. 6A) yielded no N-terminal sequence. However, we assign it to ORF1115, because this is the only large ORF in the genome of ARV1 (see below).

In order to investigate whether the diffuseness of the major band was caused by post-translational modification, the gel was stained with periodic acid Schiff’s reagent (see Materials and methods). The differences in staining methods and solvents used produced differences in the size of the two parts of the same gel. The size markers (M) are indicated.

**Sulfolobus** rudiviruses (Prangishvili et al., 1999; Peng et al., 2001). A weak upper band visible on the gel at about 116 kDa (Fig. 6A) yielded no N-terminal sequence. However, we assign it to ORF1115, because this is the only large ORF in the genome of ARV1 (see below).

In order to investigate whether the diffuseness of the major band was caused by post-translational modification, the gel was stained with periodic acid Schiff’s reagent (see Materials and methods). The band corresponding to the major CBB-stained protein was strongly stained indicating that it is highly glycosylated (Fig. 6B). Upper bands, also stained, are present at the border of the concentrating and separating gels and probably represent aggregated protein which is not stained by CBB (Fig. 6B).
Genome and its organization

The nucleic acid extracted from purified virus particles was digestible by type II restriction endonucleases and was insensitive to RNase A, indicating that it was double-stranded DNA. Therefore, a shotgun library was prepared in pUC18 from about one microgram of DNA extracted from the virions. It was sonicated and cloned inserts of 2–3 kb were sequenced using Biorobots and MegaBACE Sequenators (see Materials and methods). Sequencing produced an about fivefold genome coverage and yielded a single contig of approximately 22 kb. For linear genomes the terminal 1–2 kb are never represented in the shotgun libraries, in our hands. Therefore, given the very low amounts of DNA available, we amplified the viral DNA using the GenomiPhi procedure (Amersham Biotech). This produced sufficient DNA for primer walking beyond the ends of the 22-kb contig. Although amplification of linear genomes also produces very low molar yields of the terminal regions, we were able to extend each contig end.

Table 1

Properties of the ORFs and operons of ARV1

<table>
<thead>
<tr>
<th>ORF</th>
<th>SIRV1/2 homolog</th>
<th>Promoter motif</th>
<th>S-D motif</th>
<th>Terminator</th>
<th>Operon</th>
<th>Predicted function</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward strand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44*</td>
<td>ATATATGTC(−18)</td>
<td>ATATATGTC(−18)</td>
<td>(8)TTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>ATATTTATATA(−18)</td>
<td>ATATTTATATA(−18)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>TTTAAAAATA(−23)</td>
<td>GGTG(−7)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>TTATA(−32)</td>
<td>GGTG(−7)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83a</td>
<td>ATATTTATATA(−18)</td>
<td>GGTG(−7)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60a</td>
<td>ATATTTATATA(−18)</td>
<td>GGTG(−7)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>TATAATA−29</td>
<td>(+1)TTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>AAATAATA−25</td>
<td>(+9)TTTTAT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>TTTAAAATA−36</td>
<td>GGTG(−9)</td>
<td>(9)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134a</td>
<td>131/131b</td>
<td>GAGG(−10)</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>GGTG(−9)</td>
<td>(10)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>GGTG(−9)</td>
<td>(10)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74a</td>
<td>GGTG(−9)</td>
<td>(10)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>443</td>
<td>440/436</td>
<td>(6)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>207</td>
<td>207</td>
<td>GGCG(−7)</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>124</td>
<td>TATATAATA(−15)</td>
<td>GGTG(−7)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>TATATAATAAA(−23)</td>
<td>GGTG(−7)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>81/84a</td>
<td>GGTG(−6)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1115</td>
<td>1070</td>
<td>GGTG(−6)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>379</td>
<td>417/309</td>
<td>GGTG(−6)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>368</td>
<td>356</td>
<td>(6)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>196</td>
<td>209/176</td>
<td>GTGG(−8)</td>
<td>(5)TTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>TTTATAATA(−24)</td>
<td>GTGG(−8)</td>
<td>(5)TTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>590</td>
<td>562/564</td>
<td>GGTG(−10)</td>
<td>(10)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>472</td>
<td>486/488</td>
<td>GGTG(−10)</td>
<td>(10)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>121</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134b</td>
<td>134</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>336</td>
<td>335</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>318</td>
<td>268/269</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74b</td>
<td>74/69</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>158</td>
<td>158b</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>46</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reverse strand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>ATATATGTC(−18)</td>
<td>ATATATGTC(−18)</td>
<td>(8)TTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>ATATATATATA(−18)</td>
<td>ATATATATATA(−18)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>TTTATATAATTT(−33)</td>
<td>GTGA(−8)</td>
<td>(3)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83b</td>
<td>ATATATATATATTTGTC</td>
<td>GTTA(−8)</td>
<td>(3)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60b</td>
<td>59</td>
<td>ATATATATATATTTGTC</td>
<td>GTTA(−8)</td>
<td>(3)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>196</td>
<td>209/176</td>
<td>TTTATATAATTT(−24)</td>
<td>GTTA(−8)</td>
<td>(3)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>TTTATATAATATTTGTC</td>
<td>GTTA(−8)</td>
<td>(3)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>590</td>
<td>562/564</td>
<td>TTTATATAATATTTGTC</td>
<td>GTTA(−8)</td>
<td>(3)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>472</td>
<td>486/488</td>
<td>TTTATATAATATTTGTC</td>
<td>GTTA(−8)</td>
<td>(3)TTTTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>121</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134b</td>
<td>134</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>336</td>
<td>335</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>318</td>
<td>268/269</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74b</td>
<td>74/69</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>158</td>
<td>158b</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>46</td>
<td>GGTG(−8)</td>
<td>(8)TTTTTTTTT</td>
<td>Operon</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ORFs are listed together with putative TATA-like promoter, Shine-Dalgarno motifs, and T-rich terminator sequences. Numbers in brackets indicate the position of the last nucleotide of the promoter and S-D sequences, and the first nucleotide of the terminator sequence, relative to the predicted start or stop codon, respectively. The trinucleotide GTC, present in some promoters, is shown in bold. All functional assignments were made on the basis of best matches in the GenBank/EMBL database, except for ORF56 and ORF98 which exhibited HTH4 and PadR motifs, respectively, using Pfam. The corresponding scores (e-values) are given. (*) denotes the motif e-value.
by about 1 kb after several rounds of primer walking. Subsequent rounds of walking yielded no additional sequence. The total sequence obtained was then 24,655 bp which included 1365 bp of an inverted terminal repeat (ITR). No sequence heterogeneities were detected in genome and the overall G + C content was 40%. We established that the 1365-bp terminal repeat was perfect by generating and sequencing PCR products extending from a primer annealed outside of one end of each terminal repeat sequence to a primer annealed close to each end of the genome.

In order to estimate whether any sequence was lacking at each terminus, we performed a restriction digest with SpeI which cut within the ITR (data not shown). The estimated sizes of the resultant terminal restriction fragments of 1115 (±20) bp was consistent with the actual sequence obtained for this genome region of 1125 bp, and therefore we conclude that the genome sequence is complete. The ARV1 ITRs are shorter than those of the SIRV viruses and show little sequence similarity with them. As for the Sulfolobus rudiviruses, and other crenarchaeal viruses, the terminal regions contain many short, imperfect, direct repeats of 6–7 bp (Bettstetter et al., 2003; Peng et al., 2001). Only a 21-bp sequence near the genomic termini was perfectly conserved in each of the three viral genomes (see Discussion).

Forty-one ORFs were predicted using MUTAGEN as described earlier (Bettstetter et al., 2003). They were distributed on both DNA strands, with AUG start codon usage predominating (~74%). The ORFs are listed together with their putative TATA-like promoter motifs and Shine-Dalgarno motifs in Table 1. The motif patterns resemble those observed earlier for genes of viruses and chromosomes of the Sulfolobus and Acidianus genera where many single genes and first genes of operons generate transcripts which exhibit very short leaders or are leaderless, whereas other genes within operons are generally preceded by Shine-Dalgarno sequences with GGTG predominating (~74%). The ORFs are color-coded in Fig. 7. Five ORFs are shared with rudiviruses SIRV1 and SIRV2 and lipothrixviruses SIFV and AFV1 (blue) while 13 are shared exclusively with the two rudiviruses (red). A further two ORFs only yield matches with the fusellovirus SSV1 (green).

Functions could be assigned to only a few putative genes on the basis of matches with public sequence databases and their properties are summarized in Table 1. They include ORF134b which shows 80–84% sequence identity to the coat proteins of rudiviruses SIRV1 and SIRV2 and corresponds to the major CBB-stained protein in Fig. 6B, as well as two types of glycosyl transferases (ORF336, ORF368), a Holliday junction resolvase (ORF115), a thymidylate synthase (ThyX) (ORF250), and three putative transcriptional regulators (ORF56, ORF60b, ORF98) (Table 1).

**Discussion**

The novel virus ARV1 which infects strains of the hyperthermophilic archaean genus *Acidianus* was assigned to the family *Rudiviridae*. The rod-shaped virions are similar in structure to the known rudiviruses SIRV1 and SIRV2 of *Sulfolobus* in that they all lack an envelope and carry three tail fibers at each end (Prangishvili et al., 1999). The virion carries one major coat protein, which is glycosylated, and a minor protein of about 124 kDa, both of which have homologues in other rudiviruses (Table 1).
Lengths of the virions of rudiviruses appear to be proportional to the lengths of their genomes, as indicated by the following data: SIRV2, 900 ± 50 nm long, 35.5 kb DNA; SIRV1, 830 ± 50 nm long, 32.3 kb DNA; and ARV1, 610 ± 50 long, 24.7 kb DNA (Prangishvili et al., 1999; Peng et al., 2001). This proportionality suggests that the DNA assumes the same conformation within each virion type which is also consistent with the high amino acid sequence identity of the coat proteins of the three viruses.

One striking difference between the TEM images of ARV1, and of SIRV1 and SIRV2, concerns the dark line along the center of the virions. For SIRV1 and SIRV2 it is uniform along the whole length, except for about 50 nm at each terminus and it was inferred to arise from a central cavity in the virions which is blocked at each end by 45 × 6 nm plugs (Prangishvili et al., 1999). However, a broken line is observed in virions of ARV1 (Figs. 1 and 2) which underlines the earlier interpretation and suggests that the dark line may arise also from positive staining of a DNA–protein complex. Heterogeneity of the structure along the whole length of virions, observed in the TEM, could also result from uneven shrinkage caused by uranyl acetate staining.

Although little is known about the adsorption mechanism(s) of the rudiviruses SIRV1 and SIRV2, the present work strongly indicates that ARV1 attaches directly to the cellular membrane prior to infection. The attachment occurs via the ends of the virus particles and it is likely that the terminal fibers are directly involved.

Most crenarchaeal viruses do not kill their hosts on viral replication (reviewed by Prangishvili et al., 2001) and the growth curves of the ARV1-infected host suggest a similar virus–host relationship for ARV1. This inference is reinforced by electron microscopic observations, which revealed no evidence for lysis of host cells on virion release (Fig. 5).

The double-stranded DNA genome was sequenced and yielded 24,655 bp including 1365 bp of an inverted terminal repeat (ITR). Most predicted protein genes of ARV1, as for other crenarchaeal viruses, show several positive matches with genes of other crenarchaeal viruses but few matches with other genes in public sequence databases (summarized in Fig. 7). Sixteen ORFs gave good matches with ORFs of the rudiviruses SIRV1 and SIRV2 and four of these also yielded good matches with ORFs of the lipothrixviruses SIFV and AFV1 (Arnold et al., 2000a; Bettstetter et al., 2003; Peng et al., 2001). In addition, ORFs 108 and 250 were exclusively homologous to genes in the lipothrixvirus SIFV and Sulfolobus chromosomes, respectively (Arnold et al., 2000a). The few functional assignments of these homologous proteins include a viral coat protein, glycosyl transferases, and a Holliday junction resolvase (Table 1). It is worth noting that at least some archaeological viruses do have their own glycosyl transferases (Arnold et al., 2000a; Bettstetter et al., 2003; Peng et al., 2001) rather than using the host cell’s enzymes. Obviously, these viruses require a glycosylation mechanism for specifically modifying viral proteins. This might be important either for the assembly of the virus particle, or for modifying proteins in a unique manner for a specific function, for example, to modify the tail fiber proteins involved in the virus-host interaction.

A further protein encoded in ARV1 is a thymidylate synthase (ThyX). This contrasts with the rudiviruses SIRV1 and SIRV2 both of which produce a dUTPase (Prangishvili et al., 1998; Peng et al., 2001). These enzymes function in adjacent steps of the de novo synthesis pathway of thymidine nucleotides and either enzyme can help to maintain a low dUTP–dTTP ratio and thereby minimize misincorporation of uracil into DNA (Chen et al., 2002). This is important at high temperatures when dCTP deamnates more rapidly to yield dUTP (Hogrefe et al., 2002) and crenarchaeal DNA polymerases are known to be strongly inhibited by dU-containing DNA in contrast to their mesophilic counterparts of Bacteria and Eukarya (Lasken et al., 1996). Thus the ThyX enzyme is probably important for the maintenance and stability of the hyperthermophilic virus.

Transcriptional analyses of the rudiviruses SIRV1 and SIRV2 demonstrated earlier that about two thirds of the TATA-like motifs were followed directly by the sequence GTC, and it was speculated that this could be a signal for transcriptional regulation by a factor encoded in the rudiviral or host genomes (Kessler et al., 2004). The same pattern was discernible for about one third of the ARV1 promoter regions, although for homologous genes in the three viral genomes there was little evidence of the GTC motif being conserved in specific gene promoters (Table 1). Another result of these transcriptional analyses was that no transcripts were observed for eight of the earlier predicted genes (Peng et al., 2001). This is consistent with none of these ORFs being conserved in ARV1, with the possible exception of a CopG protein homolog (ORF59).

The large perfect ITRs of ARV1 are likely to have a similar function of facilitating viral genome replication as those of rudiviruses SIRV1 and SIRV2 (Blum et al., 2001; Peng et al., 2001). The SIRV ITRs were shown earlier to exhibit three tandem direct repeats, TTTTTTTGC, near the genomic termini, and also to carry degenerate direct repeats spaced throughout the ITR each which include one to three copies of the sequence AAATTCC (Peng et al., 2001). In ARV1 the former repeats are absent as are multiple copies of the degenerate repeats. However, one of the latter is present corresponding to the (complementary) sequence AATTAGGAATTTAGGAATTT. This is identical in each of the three viral genomes and is located 132–152 bp from the termini of SIRV1 and SIRV2 and 101–121 bp from the end of ARV1. Since it is the only highly conserved sequence in all three viral ITRs, it may be involved in genome replication or in resolution of replication intermediates (Peng et al., 2001).

The genome of ARV1 appears to be very stable because no evidence was accrued for the occurrence of any sequence variants during the sequencing of most of the genome from
the natural viral DNA with a five-fold sequence coverage. In this respect it closely resembles SIRV2 but it contrasts strongly with SIRV1 which exhibits widely differing sequence variants (Peng et al., 2004) and, to a lesser degree, with the other sequenced crenarchaeal viral genomes all of which show some level of sequence heterogeneity albeit often confined to a few local regions as was reported, for example, for the Pyrobiaculum spherical virus (PSV) (Häring et al., 2004).

In conclusion, the present results demonstrate that rudiviruses are not confined to the genus Sulfolobus but also infect members of another crenarchaeal genus, Acidianus. The results also show that the rudiviruses are widespread geographically in that they have now been isolated from Italy, as well as Iceland, and morphologically similar rod-shaped particles have been reported from hot springs in North America (Rachel et al., 2002a; Rice et al., 2001). The genetic diversity of these rudiviruses, and the unusual mechanism involved in the rearrangement of their genomes (Peng et al., 2004), suggests that they will provide a fruitful research area for studying the early development and evolution of viruses.

Materials and methods

Enrichment culture, isolation of virus hosts, and virus purification

An enrichment culture of a sample from an acidic hot spring (pH 2, 93 °C) in Pozzuoli, Italy, was established as described earlier (Rachel et al., 2002a). Single strains were isolated by plating on colloidal sulfur-containing Gelrite (Kelco, San Diego) (Zillig et al., 1994). They were grown in the same medium as used for the enrichment culture and (Bettstetter et al., 2003) the 16S rDNA genes were performed as described earlier (Rachel et al., 2002b). The genetic diversity of these rudiviruses, and the unusual mechanism involved in the rearrangement of their genomes (Peng et al., 2004), suggests that they will provide a fruitful research area for studying the early development and evolution of viruses.

Preparation and sequencing of viral DNA

Virus DNA was prepared after disrupting virus particles by treatment with 1% SDS for 1 h at room temperature as described (Häring et al., 2004). A shotgun library was then prepared for the ARV1 viral genome. DNA was sonicated to produce fragments of about 2 kb and these were cloned into the Smal I site of the pUC18 vector. DNA was extracted from single colonies using a Model 8000 Robot, (Qiagen, Westburg) and sequenced in a MegaBACE 1000 sequenator (Amersham Biotech). The viral sequence was assembled using a system Procise 492 (Applied Biosystems, Foster City, USA).

Preparation of membrane vesicles and S-layer sacculi

Membrane vesicles and S-layer sacculi were prepared from A. pozzuoliensis as described (Grogan, 1996).

Protein analysis

Proteins were extracted and analyzed in 13.5% SDS–polyacrylamide gels (Laemmli, 1970) and stained with Coomassie Brilliant blue R-250 (Serva). PAS-staining (Periodic acid Schiff’s Reaction) was performed using the protocol of the manufacturer (Sigma). Protein sequencing was performed using a system Procise 492 (Applied Biosystems, Foster City, USA).

Transmission electron microscopy

Samples were deposited on a carbon-coated copper grid, negatively stained with 3% uranyl acetate, pH 4.5, and examined in a CM12 transmission electron microscope (FEI, Eindhoven, The Netherlands) operated at 120keV. The magnification was calibrated using catalase crystals negatively stained with uranyl acetate (Reilein, 1998). All images were digitally recorded using a slow-scan CCD-camera connected to a PC running TVIPS software (TVIPS GmbH, Gauting, Germany). For heavy-metal shadowing, cells were treated with 0.2% glutaraldehyde for 30 min, placed on a carbon-coated copper grid, and washed once with double-distilled water. After air-drying they were unidirectionally shadowed with Pt/C (15°). Freeze-etching of samples was performed as described earlier (Rachel et al., 2002b).

Preparation and analysis of cellular DNA

DNA was extracted from cells of Acidianus as described (Bettstetter et al., 2003). The 16S rDNA genes were amplified by the polymerase chain reaction (PCR) using the primers 8aF and 1512 uR (Eder et al., 1999). The PCR products were sequenced and the sequences were analyzed to determine the phylogenetic status of each isolate.

Preparation and sequencing of virus DNA

Virus DNA was prepared after disrupting virus particles by treatment with 1% SDS for 1 h at room temperature as described (Häring et al., 2004). A shotgun library was then prepared for the ARV1 viral genome. DNA was sonicated to produce fragments of about 2 kb and these were cloned into the Smal I site of the pUC18 vector. DNA was extracted from single colonies using a Model 8000 Robot, (Qiagen, Westburg) and sequenced in a MegaBACE 1000 sequenator (Amersham Biotech). The viral sequence was assembled using the Sequencher program 4.2 (Gene Code). Additional primers for gap closing and polishing were created using Primers for Mac V. 1.0. ORFs were identified using Mutagen V. 4.0 (Brügger et al., 2003) and searched against: the
GenBank/EMBL databases (Tatiana and Thomas, 1999), Pfam (http://www.sanger.ac.uk/Software/Pfam) and SMART (http://smart.embl-heidelberg.de/). Genome maps were generated, and archaeal viral genomes were compared, using Mutagen V 4.0 (Brügger et al., 2003). Nucleotide sequences or amino acid sequences were aligned by T-COFFEE (Notredame et al., 2000).

Acknowledgments

We thank Hien Phan for help with the DNA sequencing, Kim Brügger for setting up and maintaining a viral sequence database, Ariane Briegel for visualization using the Amira software package, and Patrick Forterre and Hans-W. Ackermann for helpful discussions. Thanks are extended to Ian J. Molineux and two anonymous reviewers for valuable comments on the manuscript. The work was supported by grants from the Deutsche Forschungsgemeinschaft (PR 663/2-1), the Danish Natural Science Research Council for an Archaea Centre, the Forschungsgemeinschaft (PR 663/2-1), the Danish Natural Science Research Council for an Archaea Centre, the Danish Technical Scientific Research Council for a postdoctoral grant to X. P. and by EU Sulfolobus project grant (QLK3-2000-00649).

References

Blum, H., Zillig, W., Mallok, S., Domdey, H., Prangishvili, D., 2001. The genome of the archaeal virus SIRV1 has features in common with the genomes of eukaryal viruses. Virolology 281, 6–9.
different and highly organised mechanisms of translation initiation in the archaeon Sulfolobus solfataricus. Extremophiles 4, 175–179.

