Characterization of the Transcriptional Activity of the Cryptic Plasmid pRN1 from *Sulfolobus islandicus* REN1H1 and Regulation of Its Replication Operon^{▽†}

Silvia Berkner and Georg Lipps*

Department of Biochemistry, University of Bayreuth, 95440 Bayreuth, Germany

Received 12 October 2006/Accepted 7 December 2006

The plasmid pRN1 from *Sulfolobus islandicus* REN1H1 belongs to the crenarchaeal plasmid family pRN. The plasmids in this family encode three conserved proteins that participate in plasmid replication and copy number regulation, as suggested by biochemical characterization of the recombinant proteins. In order to deepen our understanding of the molecular biology of these plasmids, we investigated the transcriptional activity of the model plasmid pRN1. We detected five major transcripts present at about 2 to 15 copies per cell. One long transcriptional unit comprises the genes for the plasmid-copy-number control protein Orf56/CopG and the replication protein Orf904. A second transcript with a long 3'-untranslated region codes for the DNA binding protein Orf80. For both transcripts, we identified countertranscripts which could play a regulatory role. The function of the fifth transcript is unclear. For the five transcripts, we determined the start site, the transcript end, the stability, and the abundance in different growth phases. Reporter gene experiments demonstrated that the copy number control protein Orf56 represses transcription of the *orf56-orf904* cotranscript in vivo.

The plasmid pRN1 (5,350 bp) has been isolated from *Sulfolobus islandicus* strain REN1H1 (47) and occurs natively together with plasmid pRN2 in its host strain but has been shown to replicate independently of pRN2 (36). It is a member of the pRN family of genetic elements, comprising pRN1, pRN2, pDL10, pHEN7, and pSSVx (3, 22, 23, 25, 34). The more recently described plasmids pTIK4, pTAU4, pORA1 (18), and pIT3 (35) also contain open reading frames with sequence similarity to open reading frames from the pRN family plasmids. The sequence of pRN1 has been determined (22), and six open reading frames have been identified. Keeling et al. also tentatively assigned functions to two of the encoded proteins based on sequence similarities to characterized proteins. Knowledge of the encoded proteins was substantially increased by the heterologous overexpression and functional as well as structural characterization of the proteins Orf56 (26), Orf80 (27), and Orf904 (28, 29). A model for the replication of pRN1 has been proposed (30), as follows. Orf56, a putative repressor protein, binds as a tetramer to an inverted repeat upstream of its own gene. If *orf56* and *orf904* are cotranscribed, the tetramer would downregulate the expression of both genes and thus control replication initiation. Orf904, a multifunctional replication protein, could melt the replication origin and then synthesize a primer to start the replication of pRN1. The role of Orf80, a sequence-specific DNA binding protein, is still unclear. Orf80 seems to play an important role in plasmid replication or maintenance, as highly conserved homologs

were found to be encoded by almost all *Sulfolobus* plasmids sequenced so far (17), and also, putative binding sites for Orf80 are present on several plasmids (31).

The regulation of plasmid copy number has been studied in detail for different bacterial plasmid families. Orf56 shows sequence similarity to CopG, a protein involved in copy number control of the bacterial plasmid pLS1 from the pMV158 family. CopG has been shown to bind to a 13-bp inverted repeat in the common promoter region (–35 box) of the *copG-repB* operon and to repress its own synthesis and that of the replication protein RepB (15). Additionally, an antisense RNA is needed as a regulatory element in pLS1. RNAII acts as a translational repressor by binding to the ribosome binding site of the *repB* mRNA (10, 13, 14).

Because of its small size, its stable maintenance in *Sulfolobus* cells without integrating into the host genome, and the fact that it does not cause any growth retardation, the multicopy plasmid pRN1 is well suited as a backbone for the construction of *Sulfolobus*-*Escherichia coli* shuttle vectors. However, up to now, information on this plasmid has been limited to bioinformatic analysis and protein function analysis. There is currently no experimental evidence on the mode of replication of pRN1, on a putative origin, or on a minimal replicon. The analysis of the transcripts of this plasmid will help to understand pRN1 in more detail and will perhaps also be useful for finding suitable interruption sites for the insertion of an *E. coli* replicon for the construction of a *Sulfolobus*-*E. coli* shuttle vector.

MATERIALS AND METHODS

Strains and culture conditions. *Sulfolobus islandicus* REN1H1 with plasmids pRN1 and pRN2 (47) was grown in Brock's medium (20) with 0.1% tryptone (BD Biosciences), 0.2% D-arabinose (Fluka), and a vitamin solution (46) in 50-ml shake flask cultures. For larger culture volumes (1.5-liter fermenter), a rich medium (36) was used. *Sulfolobus solfataricus* PH1-16 (33) as well as the parent

* Corresponding author. Mailing address: Department of Biochemistry, University of Bayreuth, 95440 Bayreuth, Germany. Phone: 49-921-552433. Fax: 49-921-552432. E-mail: georg.lipps@uni-bayreuth.de.

† Supplemental material for this article may be found at <http://jb.asm.org/>.

[▽] Published ahead of print on 15 December 2006.

strain *Sulfolobus solfataricus* P1 was grown in Brock's medium with 0.1% tryptone and 0.2% arabinose, with or without $10 \mu\text{g ml}^{-1}$ uracil.

RNA preparation. RNAs were prepared by a Trizol method using RNeasy (Qiagen). RNA preparations were digested with DNase I (Roche) at room temperature for 30 min and subsequently column purified (RNeasy Mini kit; Qiagen). This combination yielded the best reproducibility of parallel RNA preparations (coefficient of variation of RNA concentrations, 13% [$n = 4$]). A crucial problem in the preparation of RNAs from plasmid-containing *Sulfolobus* cells is the large amount of plasmid copurifying in the first step of purification (Trizol method). This amount is reduced by the DNase treatment, but there is still a very low level of residual DNA contamination that cannot be eliminated even by a second DNase digestion at 37°C . Since the second DNase digestion step increases the variability in parallel preparations and does not yield DNA-free RNA preparations, only one DNase digestion step was performed. The residual DNA does not interfere with quantitative reverse transcriptase PCR (qRT-PCR) measurements because the level of DNA molecules compared to cDNA molecules is lower by several orders of magnitude. Nevertheless, a specific PCR product was sometimes also detected in control reactions containing only RNA without the RT step, but at considerably higher threshold cycle (C_T) values than in the corresponding samples.

Reverse transcription. Reverse transcription was performed using Transcriptor reverse transcriptase (Roche) at 55°C to reduce the influence of secondary structures. Initially, an RNase inhibitor (Roche) was added to the RT reactions, but it was subsequently left out because it showed an inhibitory effect on the qPCR when the RT product was used at a 25-fold dilution in the qPCR reaction mix. The inhibition could be circumvented by using a 250-fold dilution, which was commonly used for all sample quantifications.

The influence of the input amount of RNA on the transcribed cDNA amount was tested in the range of $0.5 \mu\text{g}$ to $5 \mu\text{g}$ of RNA and found to be linear in that range, allowing the use of different input amounts of RNA.

qRT-PCR. qRT-PCRs were carried out in 96-well plates ($25\text{-}\mu\text{l}$ volume) with an ABI PRISM 7000 sequence detection system (Applied Biosystems). SYBR green real-time PCR master mix (Epicenter) was used according to the instructions of the manufacturer, using ROX dye for signal normalization. The master mix and primers were premixed and dispensed into a 96-well plate, and appropriate dilutions of standards, controls, and samples were added. Primers (Metabion, Martinsried, Germany) were used at a concentration of 500 nM . An example of qRT-PCR results is shown in Fig. 1.

Primer design. Primers were designed using the program VectorNTI (Invitrogen). Several different primers were tested for their efficiency, and the best results were obtained with a primer length of around 30 nucleotides (nt), producing a product of about 150 bp. Since the GC content and the distribution of repetitive sequences change within the sequence of plasmid pRN1, it was not always possible to optimize every PCR to an efficiency of over 90%. The short open reading frames did not leave much space for alternative primer binding sites. To avoid any cross-hybridization with pRN2 present in the samples, all primers were checked to not have high sequence similarity to pRN2. In addition, PCRs performed on the pRN2 plasmid alone did not yield a product.

Standards and calibration. Standards were prepared from plasmid pUCpRN1, the pRN1 part was cut out, and the fragment was purified by agarose gel electrophoresis and extracted using the Wizard SV gel and PCR cleanup system (Promega). The concentration of the fragment was determined photometrically, and a stock dilution containing 10^9 molecules per μl was prepared and stored at -20°C . Fresh dilutions were prepared for every measuring period at concentrations ranging from $10^{8.5}$ to $10^{3.5}$ molecules per μl , yielding final standard concentrations ranging from $10^{7.5}$ to $10^{2.5}$ molecules per μl . For other target genes, similarly purified fragments were used.

Sample analysis was always accompanied by measurements of standards for calibration purposes. The C_T values for standards were plotted against the logarithms of the numbers of molecules per μl contained in the standards, and the calibration function was derived.

Accuracy, precision, and reproducibility. The accuracy of the developed method is difficult to assess with original samples, as the amount of a target RNA cannot be known. To quantify the recovery of a transcript after the RNA purification and reverse transcription steps, an in vitro-transcribed RNA was spiked in three parallels into three *Sulfolobus* cell pellets, and subsequently, the RNA was prepared, reverse transcribed, and quantified. To obtain the test RNA, a 300-bp fragment of the kanamycin resistance gene from *E. coli* was cloned into pGEM-T (Promega) and transcribed using T7 RNA polymerase (Ampliscribe high-yield transcription kit; Epicenter). The resulting transcript was checked for integrity on an agarose gel, and the concentration was determined photometrically. The recovery of the spiked transcript was determined to be 60%, and the precision for the three parallels was 6% (coefficient of variation).

FIG. 1. Example of a qRT-PCR experiment. Standards, samples, and controls for the primer pair RT3_o56_for/rev (length of PCR product, 158 bp) were each used in three replicates. (A) Standard concentrations ranged from $10^{7.5}$ molecules μl^{-1} to $10^{4.5}$ molecules μl^{-1} , with spacing of $10^{0.5}$ molecules μl^{-1} . (B) Transcript numbers were measured in four different samples (25-fold dilution and 250-fold dilution) and in four control reactions without RT (25-fold dilution). (C) Corresponding agarose gel analysis of the qRT-PCR products from panels A and B (only one replicate). M, marker; N, no-template control.

The precision of the qPCR step alone, as judged from the three replicates that were measured for every standard concentration for every primer pair, was always in the range of 1%, with the maximum coefficient of variation being 3.5% (Table 1).

The transcript from the spiking experiment is probably much simpler to extract from the cell pellet than are real RNAs. Thus, the coefficient of variation in parallel RNA preparations for four parallel preparations was determined to be 13%. The precision of RNA extraction, reverse transcription, and qPCR was derived from the law of error propagation and was in the range of 14%. Taking into account the additional variability introduced by the normalization of the transcript levels to cell numbers (derived by measuring the optical density [OD] and calculating cell numbers by a function derived from counting cells in a Neubauer counting chamber, with an estimated error of approximately 10%), the precision of the final transcript number per cell was below 20%.

To assess the reproducibility of the procedure, several RNA transcript levels (*orf56*, *orf904*, *orf80*, *orf90a*, *orf72*, *orf90b*, and *orf90a* [4]) were quantified using cell pellets from different fermentations on different days (sampled at ODs of 1.3

TABLE 1. Summary of different qRT-PCRs^a

qRT-PCR primer	Length of product (bp)	Efficiency (%)	Precision (%)	Correlation coefficient (r ²)
RT3_o56	158	91	0.1–1.1	0.999
RT3_o904	150	95	0.2–2.0	1.000
RT3_o80	160	92	0.2–1.9	0.999
RT_o90a	204	73	0.2–1.5	0.999
RT3_72	168	86	0.2–1.2	0.999
RT2_90b	202	66	0.1–1.5	0.999
RT4_o90a	160	78	0.2–1.6	0.998
RT_no80	168	90	0.3–2.4	0.998
RT5_o904	143	96	0.1–3.5	0.994
RT_lacS	160	92	0.3–1.4	0.998
RT_o80prn2	130	80	0.2–0.7	1.000
RT_o90aprn2	177	Qualitative only	Qualitative only	
RT_kan	150	88	0.2–1.7	0.998

^a Data are from measurements of standard curves for 11 equidistant concentrations in three replicates.

to 1.4) that were extracted on different days and quantified using different standard preparations. The reproducibility was determined to be 24% (difference of two determinations divided by the mean).

Plasmid copy number determinations. Plasmid copy number was determined by qPCR as described above (with primers RT3_orf56_for and -rev). Plasmids were prepared by alkaline lysis and diluted 10⁵- to 10⁶-fold for the measurement.

Three independent replicates were sampled and analyzed for each OD. In addition, the qPCR results were found to be consistent with copy number determinations carried out by comparing band intensities of plasmid dilutions on agarose gels.

Adaptor RT-PCR. During reverse transcription, an adaptor primer consisting of 20 nt complementary to the 3' end of the target RNA and a 20-nt adaptor sequence was incorporated into the cDNA molecule. In the subsequent PCR step, only a primer matching the upstream region of the RNA and a primer consisting solely of the adaptor sequence were used, circumventing the false-positive results caused by traces of residual plasmid DNA in the RNA preparations. The adaptor primer (0.5 µM in the RT reaction) was removed after the RT step by purifying the cDNAs with the Wizard SV gel and PCR cleanup system (Promega). Control PCRs without the adaptor sequence primer and without RT were always included. Primer sequences are listed in Table 2, and additional primers used to narrow down the transcript ends can be found in Table S1 in the supplemental material.

Determination of the stability of transcripts. Actinomycin D at 10 µg ml⁻¹ was used to inhibit transcription (8). Aliquots of 10 ml were collected 0 min, 15 min, and 30 min after the addition of actinomycin D from the treated and untreated control cultures.

Primer extension. A primer (0.1 pmol) 5' end labeled with T4 polynucleotide kinase (approximately 150 nCi) and 20 to 50 µg of RNA were used in a reverse transcription reaction according to the protocol of the manufacturer (Transcriptor reverse transcriptase; Roche), using a 10-µl reaction mix. The enzyme was heat inactivated at 85°C for 5 min, and 10 µl of denaturing loading buffer was added. The sequencing reactions run alongside the primer extension reactions were prepared using a T7 sequencing kit (GE Healthcare). The influence of secondary structures, such as stem-loops, on the primer extension reaction was tested using a perfect 8+3 stem-loop structure situated directly downstream of the stop codon of *orf904*. A primer extension reaction in that region did not show

TABLE 2. Primers used for qRT-PCR and adaptor RT-PCR experiments

Primer (forward/reverse) and target	Forward sequence (5'–3') ^a	Reverse sequence (5'–3') ^a
qRT-PCR primers		
<i>orf56</i>		
RT3_o56	CCACAAATGGGTAGACCATACAACTATTAAA	TGATTGTCTAAATACTCTTTAATTGCCAGCCT
<i>orf904</i>		
RT3_o904	CGCCACACAAAATCAATCCTCTCTTTG	TGTAGTTCTGTTCTTTCCACGGGCATT
<i>orf80</i>		
RT3_o80	GGGATCCCATATGAGTGATCTGAAGGA	GCCTTTTTTATGGAGTCGTGTCAAGGT
<i>orf90a</i>		
RT_o90a	TATCGTTTTACCTGATGCGTATAC	TTAAGCAAAAGCCTGAGCAA
<i>orf72</i>		
RT3_72	TACTTTCATTATCTCTTCTTAGCTTTTGCTCTC	ATTGTCTAAGATAATCAAAGATTGGGTTAGCA
<i>orf90b</i>		
RT2_90b	ATGCCGACTTCTAACAACATCTTAC	GTTGAGAGTAAGAAACAGAAGCAAA
<i>orf90a</i> plus upstream sequence		
RT4_o90a	CACATGATAAGAGTATAAAATAGGCGTAAAAG	GACGTATACAGTAAAAATTTAGATGCATTGAA
Downstream of <i>orf80</i>		
RT_no80	ACTCTGTGAAACTCCCCAGTTTACC GTTATTA	ATAGCTGAAAGCCGTACGCTTTTACGCCATT
<i>orf904</i>		
RT5_o904	GAAAAAACAAGAAGAACGGGGGTGTACATGAT	GGTCGATTGGGATAATATTAAGCCGTGTCT
RT_lacS	GGGAACACCAGGGTCAGAAGATCCAAATCTG	TCCCATTTTTTGTGCATTATCGTGAAATGTCT
RT_o80prn2	GAATGCATAAGAAGGGGCTGATATATCGTAAA	ATATTTTCAAATGAAAATGAACTCAACAATCA
RT_o90aprn2	TTGTAGGCTGATTGTTGAGTTCAATTTTCATTG	TTTCTATGATATCCATATGACTAAGCTAGCGTG
RT_kan	ACAATCTATCGATTGTATGGGAAGCCCGATGC	GGATAAAATGCTTGATGGTCGGAAGAGGCATA
Adaptor RT-PCR primers		
Adaptor sequence primer		TAATACGACTCACTATAGGG
<i>orf56/904</i>		
RT3o56-for/AA_rev	CCACAAATGGGTAGACCATACAACTATTAAA	TAATACGACTCACTATAGGG TCAATCAAACCGCACGTGCT
<i>orf80</i>		
RT3o80_for/O_rev	GGGATCCCATATGAGTGATCTGAAGGA	TAATACGACTCACTATAGGG AAGGGAGCGTGAAAAATAG
<i>cto904</i>		
U_for/RT3_o904_rev	TAATACGACTCACTATAGGG GACAGATTAATGGAAATCGC	TGTAGTTCTCTTCTTTCCACGGGCATT
<i>cto90a</i>		
V_for/RT_o90a_rev	TAATACGACTCACTATAGGG ATCTTGTCAAACCTGGAAG	TTAAGCAAAAGCCTGAGCAA
<i>cto90b</i>		
J_for/RT2_o90b_rev	TAATACGACTCACTATAGGG CTTTGATTATCTTAGACAAT	GTTGAGAGTAAGAAACAGAAGCAAA

^a For the adaptor RT-PCR primers, the adaptor primer sequence is separated by a vertical line from the plasmid complementary sequence.

FIG. 2. (A) Map of pRN1 with open reading frames. (B) Adaptor RT-PCR (see Materials and Methods) using the last primer yielding a PCR product and a primer next to the mapped transcription start site. Lanes M, marker; lanes 1, adaptor RT-PCR; lanes 2, control without the adaptor primer; lanes 3, control without RT. The lengths of the PCR products and the exact locations of the transcript ends are listed in Table 3.

any sign of impairment by this stem-loop (results not shown); thus, secondary structures are unlikely to cause false termination bands. Another test was conducted to exclude the possibility that residual DNA could cause additional bands. A primer extension reaction with RNA without DNase I treatment showed exactly the same bands as those obtained with DNase I-digested RNA.

Plasmid construction. The promoter region of *orf56* and an upstream sequence (positions 1892 to 2162 in pRN1 [accession number NC_01771]) were PCR amplified using primers 5'-TCCTAGGCTAAGCCGCCCTGTCTAAC and 5'-CCATGGGTGGATCAAAATTGTATCCGC, introducing an *AvrII* site with the forward primer and an *NcoI* site with the reverse primer (underlined). The same promoter region plus the *orf56* coding sequence (positions 1892 to 2382 in pRN1) was PCR amplified using the same forward primer and 5'-CCA TGGACACCCCGTCTCTCTTG as the reverse primer. Fragments were cloned into pGEM-T (Promega) and checked by sequencing. The inserts were cut out using the introduced *AvrII* and *NcoI* sites and then ligated into the preassembling vector pSVA10 (S. V. Albers and M. Jönuscheit, unpublished data) cut with the same enzymes, thus replacing the *tf55a* promoter region of pSVA10 with the *orf56* promoter region or the promoter region plus the *orf56* coding sequence. From these preassembly constructs, the promoter regions, together with the *lacS* reporter gene sequence, were cut out using *BlnI* and *EagI* and ligated into pMJ05 (1), yielding pProm and pProm*orf56*.

Transformation of *Sulfolobus* cells. Conditions for the transformation of cells by electroporation followed the protocol of Schleper et al. (40), as specified in the work of Jönuscheit et al. (21).

Southern blots. Single transformants of *S. solfataricus* PH1-16 containing the reporter gene constructs were tested for the presence and copy number of free and integrated forms of the shuttle vector by Southern blotting. Genomic DNA was prepared from 1 ml of a logarithmically growing culture, using a Chemagenic DNA Bacteria kit (Chemagen, Baesweiler, Germany) according to the instructions of the manufacturer. After *PvuII* digestion, the fragments were separated in a 0.8% agarose gel, transferred to a Hybond N membrane (Amersham) by capillary transfer, fixed by UV irradiation for 5 min on a UV transilluminator, and hybridized to two digoxigenin-labeled probes complementary to positions 2580 to 3432 and 7455 to 7893 of *Sulfolobus shibatae* virus 1 (SSV1) (accession number NC_001338.1), indicating either the integrated or the free form of the shuttle vector construct. Labeling and detection were done using a PCR DIG probe synthesis kit and a digoxigenin labeling and detection kit (Roche).

β -Galactosidase assay. β -Galactosidase activity was measured from crude extracts prepared by a freeze-thaw method (21). Cells from a logarithmically growing culture were resuspended in 50 mM sodium phosphate buffer, pH 7, and subjected to five freeze-thaw cycles ($-196^{\circ}\text{C}/+50^{\circ}\text{C}$). After centrifugation for 30 min at 10,000 rpm, the supernatant was stored at -20°C or assayed directly. All β -galactosidase assays were conducted in triplicate in a 75°C bench-top shaker. The reaction mix consisted of 3 μl of crude extract (or water for blanks) and 90 μl of 50 mM sodium phosphate buffer, pH 7, and the reaction was started by the addition of 7 μl of 12-mg ml^{-1} *o*-nitrophenyl- β -D-galactopyranoside solution. Incubation was continued for 5 min before the tubes were cooled rapidly on ice, and 100 μl of 1 M Na_2CO_3 solution was added to stop the reaction. The concentration of *o*-nitrophenol was subsequently determined in a 96-well plate in a plate reader at 410 nm, using a standard curve generated with *o*-nitrophenol. The protein concentrations in the crude extracts were determined by the method of Ehresmann et al. (16).

RESULTS

Transcribed regions of plasmid pRN1. To determine which open reading frames of pRN1 (Fig. 2A) are transcribed, qRT-PCR was used because of its ability to reliably detect and quantify even transcripts of low abundance not reliably detectable in Northern blots.

For all six open reading frames identified on pRN1, primer pairs were designed (Table 2). qRT-PCRs performed on RNA preparations from cells in different growth phases showed that *orf56*, *orf904*, and *orf80* were transcribed at more than one copy per cell, whereas for *orf90a*, *orf72*, and *orf90b*, only very low (2 to 3 orders of magnitude lower) transcript levels were detected. Since RT-PCR is strand specific, we also tested the antisense direction and identified transcripts within the *orf904*, *orf90a*, and *orf90b* region, named *cto904*, *cto90a*, and *cto90b*, respectively.

Initially, RT-PCR was used to deduce information on the lengths of the transcripts but proved to be not suitable because of residual DNA contamination in RNA preparations. Thus, an adaptor RT-PCR approach (see Materials and Methods) was used to circumvent this problem. Using this approach, we could show that *orf56* and *orf904* are cotranscribed, that the sense transcript of *orf80* has a 3'-untranslated region overlapping with the antisense transcript *cto90a*, and that *cto904* overlaps with the putative Shine-Dalgarno (SD) sequence upstream of *orf904* (Fig. 2 and Table 3; also see Fig. 4).

Transcript starting points. The starting points of the identified transcripts were mapped by primer extension to deter-

TABLE 3. Adaptor RT-PCR results^a

Transcript	Start position(s) in pRN1	Ending positions	Length of adaptor RT-PCR product (bp)
<i>orf56/orf904</i>	2156	5031–5096	2,892
<i>orf80</i>	524	957–1005	454
<i>cto904</i>	2938	2231–2190	715
<i>cto90a</i>	1467, 1492	593–530	664
<i>cto90b</i>	1850	1527–1463	314

^a The lengths of the PCR products shown in Fig. 2 are given, as well as the mapped transcription start sites from primer extension analysis (Fig. 3) and the regions of the transcript ends (range between the last primer that yielded a PCR product and the first primer that did not yield a product). For *cto90a*, the PCR did not cover the first 200 nt of the transcript(s). By additional primer extension reactions, no other transcript start sites were detected in that region.

FIG. 3. (A) Primer extension analysis of transcript start sites. The primer extension reaction mix was loaded on both sides of the sequencing ladder. The sequence complementary to the beginning of the transcript is shown on the left side of each sequencing gel. Additional transcription start sites for cto90a are marked with arrowheads. (B) Aligned sequences upstream of the mapped transcription start sites (last nucleotide, shown in bold). BRE and TATA boxes are underlined and were identified using the consensus sequence of Bell et al. (4), namely, A/GNA/TAAA/TT/CTTAT/AT/AT/AT/ANNANN.

mine the exact start site of each transcript, and thus to be able to identify BRE and TATA box elements, and to verify the presence of the transcripts by an independent method.

The transcript comprising *orf56* and *orf904* starts 9 nt upstream of the ATG start codon for Orf56 at the last guanine nucleotide preceding the start codon (Fig. 3). We also searched for a potential transcript originating from a starting point downstream of *orf56*, comprising only *orf904*, but we found no indication of such an additional transcript by primer extension analysis.

The starting point of the *orf80* transcript was mapped to the G of the Orf80 ATG start codon, which had previously been identified (22). Our findings suggest that the real start codon is situated four codons downstream of the original ATG. The transcript start then maps to the first of three consecutive Gs preceding the new ATG start codon by 10 nucleotides. The name of *orf80* was kept instead of *orf76* for reasons of consistency.

The starting points of the antisense transcripts cto904, cto90a, and cto90b were also determined. Within the coding region of *orf904* and directly following the end of the coding region, two more transcription initiation sites in the antisense direction were determined by primer extension, with start sites at positions 3325 and 5088 in pRN1, starting with a C and a G nucleotide, respectively. In adaptor RT-PCR experiments, it was shown that the resulting transcripts do not run through *orf904* but terminate earlier. Upstream of the mapped transcription start site for cto90b, within a region of approximately 100 nt, four more initiation sites in the antisense direction were detected from primer extension experiments. Also, for the cotranscript and the transcript cto90a, more than one transcription initiation site could be detected in the primer extension analysis. The clearest band closest to the primer was marked as the transcription start site in these cases.

In the upstream regions of the transcript start sites, BRE

and TATA box-like elements were identified (Fig. 3B). BRE and TATA boxes of the cotranscript, the *orf80* transcript, and cto904 show the highest similarity to the consensus sequence derived from different archaeal promoters (4). The identified TATA box is centered around nucleotides –26 and –27 preceding the transcription start site, as predicted by bioinformatic analysis (43). The BRE/TATA box for cto90a is only weakly similar to the consensus sequence. Comparable levels of transcripts are obtained in this case by multiple initiation sites.

Transcript endpoints. Transcript ends were estimated from the adaptor RT-PCR by using different primers spaced approximately 60 nucleotides around the supposed ends. The transcript end was assumed to be situated between the last primer that yielded a PCR product and the first primer that did not yield a product. Figure 4 and Table 3 show all compiled infor-

FIG. 4. Map of pRN1 showing the open reading frames (gray, conserved within the plasmid family; white, not conserved) as well as the characterized transcripts.

FIG. 5. Transcript ends. The region between the last primer yielding a PCR product and the first primer yielding no PCR product is shown (primer binding sites are shaded in gray). Stem-loop structures predicted using Mfold are shown, as well as U-rich sequence stretches (underlined). The ΔG values for the stem-loop structures are shown as predicted by MFold (www.bioinfo.rpi.edu/applications/mfold/) under standard conditions, i.e., 37°C and 1 M NaCl.

mation on starts and ends of the transcripts. We crosschecked the results of the adaptor PCR with the primer extension results and with additional qRT-PCR experiments performed, e.g., for the 3'-untranslated region of *orf80* (results not shown), and found all results to be consistent.

The regions of the supposed transcript ends were searched for terminator sequences. There is only very limited information on transcription termination in *Sulfolobus*. A pyrimidine-rich sequence has been identified in SSV1 as an important feature preceding transcript ends (37). Within the last 60 nt of each transcript, U-rich stretches can be identified. We used MFold (48) to predict the secondary structure of each transcript end. Stem-loop structures were found for every tran-

script (7- to 9-nt stem, 3- to 8-nt loop) (Fig. 5) but showed only limited thermodynamic stability, especially at 75°C.

Transcript levels. Having identified five major transcripts of pRN1 (Fig. 4), we wanted to determine their abundance per cell and if the transcript levels change during growth. For that reason, we first examined the time course of the plasmid copy number during a batch fermentation in a 1.5-liter fermenter with rich medium. The copy number was low directly after inoculation (2 to 4 copies per cell), increased during log phase to 23 copies per cell, and then rapidly decreased during stationary/death phase to the initial level (Fig. 6). For that reason, we chose to sample the culture in early log phase (OD, 0.3), mid-log phase (OD, 1.2), and stationary/death phase (OD, 2.2)

FIG. 6. Growth curve (triangles connected by line, left abscissa) and plasmid copy numbers (crosses, with standard deviations for three different preparations and measurements shown as error bars, right abscissa) of *S. islandicus* REN1H1 cells containing pRN1 and pRN2 during a batch fermentation (1.5 liters, rich medium).

because the plasmid copy number changes were largest at those times. The numbers of transcripts per cell are shown in Fig. 7. The cotranscript was probed at two sites, within *orf56* and within *orf904*, yielding comparable levels of three to eight transcripts per cell and displaying a decrease in stationary/death phase. The transcript of *orf80* showed constant levels during cell growth, at six copies per cell, and also decreased in stationary/death phase. The antisense transcripts *ctorf904* and *ctorf90a* reached the highest transcript levels and increased during logarithmic growth phase (OD 0.3 to OD 1.2). In the case of *ctorf90a*, the increase was fourfold, showing the most dynamic behavior of all transcripts. The transcript *ctorf90b* showed similar levels to those of the cotranscript.

Stability of the transcripts. Transcript stability was examined by using actinomycin D as a transcription inhibitor (8). We observed balanced growth of the treated cultures for a period of only 30 min after the addition of actinomycin D. The resulting transcript levels normalized to the control culture

FIG. 8. Stability of transcripts. Levels are expressed as the ratio of transcript levels in actinomycin D-treated cells to transcript levels in untreated control cells for each time point.

levels (Fig. 8) show rapid decay for *orf80*, with a half-life of 15 to 30 min, intermediate decay for *ctorf904*, *ctorf90a*, and *ctorf90b* (half-life of around 1 h), and slow decay for the cotranscript (half-life of approximately 2 h).

Regulation of the *orf56/orf904* cotranscript in vivo. We showed that *orf56* and *orf904* are cotranscribed and thus wanted to test if Orf56 represses the transcription of the cotranscript and accordingly down regulates the expression of the replication protein Orf904, using two shuttle vector constructs. The construct pProm contains the *lacS* reporter gene under the control of the *orf56* promoter and thus represents the number of transcripts that originate from this promoter in the absence of a repressor. The construct pPromOrf56 additionally contains the Orf56 coding sequence, consequently representing the situation found in the native plasmid, where Orf56 can bind to its operator within the *orf56* promoter region and down regulate the number of transcripts originating from the promoter. The numbers of transcripts originating in both constructs from the *orf56* promoter were determined by qRT-PCR, and additionally, the β -galactosidase activity was measured. The *Sul*-

FIG. 7. Transcript levels during a batch fermentation (1.5 liters, rich medium). The error bars represent the combined uncertainties from the RNA preparation, RT, and qPCR steps and cell number determinations. The overall variation is estimated to be 20%.

FIG. 9. (A) Southern blot showing *S. solfataricus* PH1-16 transformed with pPromOrf56, with pProm, and without a virus vector. (B) β -Galactosidase activities and transcript levels in cells transformed with the reporter gene shuttle vectors pPromOrf56 and pProm. For comparison purposes, the wild-type β -galactosidase activity of the strain *S. solfataricus* P1 (the parent strain of PH1-16) is also shown. The β -galactosidase activity/transcript number was normalized to the number of copies of the *lacS* gene present in the transformants because free vector copies were observed in addition to the integrated form shown in panel A (gene dose effect described by Jonuscheit et al. [21]). (C) Schematic view of reporter gene constructs. (D) Promoter region of the cotranscript. BRE/TATA boxes are underlined, the transcription start site is marked by an asterisk, and the inverted repeat binding site of Orf56 is indicated by arrows. (E) The first 24 amino acids of the Orf904 sequence are shown. Putative start codons are shown in bold, and putative Shine-Dalgarno sequences are underlined. Since the true start codon is not known, the last of the putative start codons was fused to the *lacS* reporter gene.

folobus cells containing the construct expressing Orf56 showed a 23-fold repression of β -galactosidase activity compared to cells carrying the construct containing only the promoter region. At the level of transcription, the repression was found to be 11-fold (Fig. 9). In a Southern blot, episomal copies of the reporter gene constructs pProm and pPromOrf56 were detected in addition to the site-specifically integrated form. Since a gene dosage effect has been described for this reporter gene system (21), the exact copy number of the reporter gene constructs was determined by qPCR to be two for pProm and three for pPromOrf56 for the same *Sulfolobus* culture that was used for determination of the number of transcripts. The transcript numbers as well as β -galactosidase activities were normalized to the different copy numbers.

DISCUSSION

All open reading frames of pRN1 coding for proteins that are conserved within the pRN family or, as for *orf80*, that are even more widely found in *Sulfolobus* genetic elements (18) have been shown to be transcribed. The three open reading frames that are not conserved show lower levels of transcription (2 to 3 orders of magnitude). A function for these transcripts, or perhaps higher transcription levels under different growth conditions, cannot be ruled out, but since for all three open reading frames the same low levels were measured, these transcript levels could also represent the background level of transcription in pRN1.

The *orf56/orf904* cotranscript and the *orf80* transcript are leaderless transcripts starting only 9 and 10 nucleotides, respectively, upstream of the ATG start codon. *Sulfolobus* has been predicted to have many leaderless transcripts from genome-wide analyses of transcription signals (42, 43). Our findings are also consistent with the report of the *orf57/orf80* transcript of PIT3 (35), which is entirely leaderless and also starts with a G nucleotide.

The cotranscript has no 3'-untranslated region, as it stops within the next 60 nt following the *orf904* stop codon. The transcript of *orf80*, in contrast, has a long (approximately 200 nt) 3'-untranslated region whose transcript levels are comparable to those of the coding region. The 3'-untranslated region and most of the coding region are overlapped by the antisense transcript *ctorf90a*. This transcript might play a regulatory role. The region downstream of *orf80* is not as well conserved as the coding and upstream regions of *orf80* (23). Additional qRT-PCR experiments with the corresponding region of pRN2 showed that a transcript running through *orf81* (equivalent to *orf80* in pRN1) and a transcript in the antisense direction downstream of *orf81* are present. For *Sulfolobus*, a survey of noncoding RNAs found transcripts targeting the 3'-untranslated regions of certain mRNAs (41). It is worth noting that the level of the *orf80* transcript is remarkably stable. It is the only transcript that does not show any changes in transcript level except for declining when cells are dying. This was observed for three different fermentations sampled at different times and

also for different locations within the transcript and might indicate that the number of *orf80* transcripts is tightly regulated.

The transcript *ctorf904*, found in the antisense direction within *orf56/orf904*, spanning about the first one-third of the *orf904* coding sequence and overlapping with the putative SD sequences of *orf904*, could also be involved in copy number control of pRN1. This RNA could act as a translational repressor similar to RNAII of pLS1. In contrast to the short, 50-nt RNAII, *ctorf904* is much longer, but like other regulatory RNAs of bacterial plasmids, it is shorter lived than the mRNA that it overlaps (10).

The transcript found in the antisense direction in the region of *orf90b* does not allow any conclusions to be drawn concerning its putative function.

We checked for the presence of putative open reading frames within the region of the countertranscripts. For *ctorf904*, two open reading frames of 207 bp to 258 bp could be identified (positions 875 to 618 and 864 to 658 in pRN1), and for *ctorf90a*, two open reading frames of 162 bp to 168 bp were found (positions 2781 to 2614 and 2509 to 2348). None of the putative open reading frames showed similarities to those encoding known proteins. We concentrated our transcript analysis on the region of pRN1 that is covered by open reading frames. For that reason, it is possible that further transcripts exist in the region between *orf904* and *orf80*.

The stem-loop structures identified in the region of the transcript ends exhibit some similarity to bacterial rho-independent transcription terminators that consist of a stem of eight or nine, mostly G-C, base pairs and a loop of four to eight bases (45). Reiter et al. (37) suggested that the pyrimidine-rich sequence stretches found in the mRNAs of SSV1 resemble eukaryotic termination signals and could not identify secondary structures near the ends of the mRNAs. In contrast to the viral termination signals, the mapped transcripts of pRN1 clearly show stem-loop structures followed by a U-rich sequence. However, this sequence does not closely resemble the consensus of Reiter et al. (UUUUUUU/CU). Termination has been shown to take place at RNA hairpin structures followed by a U-rich sequence in a *Methanothermobacter methanohydrothermalis* in vitro transcription system, yet with varying degrees of efficiency (38).

In contrast to the quite dynamic behavior of the pRN1 copy number during fermentation, the levels of transcripts did not show large changes over different growth phases, except for declining in stationary/death phase. One possible explanation for this observation could be that the main part of plasmid copy number control is exerted not on the transcriptional level but on the translational or even posttranslational level. It has been shown that some transcripts from *Sulfolobus* have long half-lives compared to those of bacterial transcripts (8), and it was pointed out that this renders control on the transcriptional level difficult. It was proposed that control on the translational level could play an important role in archaea. The half-lives of the pRN1 transcripts (especially the cotranscript) are in the upper range of already determined half-lives for *Sulfolobus* (2, 8), and copy number control might operate on the translational level, perhaps involving the discovered antisense transcripts. For the bacterial plasmid pLS1, the main regulatory element is the antisense transcript RNAII, which responds quickly to

copy number fluctuations, whereas the role of the CopG protein is to keep the synthesis of the *copG-repB* mRNA within narrow limits (14).

These findings are in accordance with a study on the transcription of *Sulfolobus* rod-shaped virus (SIRV1 and SIRV2), where a remarkably constant transcription pattern was observed upon infection of *Sulfolobus* cells (24).

The reporter gene assay proved that Orf56 does indeed repress transcription from its own promoter. The protein Orf56 has been shown in vitro to bind as a tetramer to a 12-bp inverted repeat situated just upstream of the *orf56* start codon. Its binding region does not overlap with BRE or TATA boxes that are found 20 to 33 nt upstream of the ATG, and thus it cannot interfere directly with binding of either transcription factor B (TFB) or TATA box-binding protein (TBP) (32). However, the Orf56 binding site overlaps the transcript start, and thus the contact of this region and the RNA polymerase might be impaired. This bacterial mechanism of RNA polymerase abrogation has been demonstrated for the archaeal species *Archaeoglobus* (MDR repressor) and *Pyrococcus* (LrpA and PhrA) (5, 12, 44). In *Sulfolobus*, transcriptional repression can also occur by blocking of the binding sites for TBP and TFB (6), and it has been suggested that derepression takes longer than in the case of RNA polymerase abrogation, where TBP and TFB can stay bound to the promoter region. To prevent plasmid loss in case of downward copy number fluctuations, it is crucial that the repression of the replication operon can be reversed quickly.

The 11-fold repression at the transcriptional level is comparable to the 10- to 20-fold repression observed in the presence of the transcriptional repressor CopR for plasmid pIP501 (9), which is also regulated by a repressor protein and an antisense RNA. Similar to the case for CopR of pIP501 and CopG of pLS1 (10), the transcription of the cotranscript is not totally blocked in the presence of Orf56.

The reporter gene assay using β -galactosidase activity yielded about two times more repression than the direct measurement of transcript levels. This could point to another mechanism at the translational level being active in the reporter gene shuttle vector constructs. It either favors the expression of the LacS protein in the pProm (without Orf56) construct or reduces its expression in the pPromOrf56 (with Orf56) construct. A mechanism that would act like the former could be the enhanced translational efficiency of a leaderless transcript (as in pProm) compared to initiation at an SD sequence (as in pPromOrf56). This has been observed for *Halobacterium salinarum* in vivo (39) and has also been described for a *Sulfolobus solfataricus* in vitro translation system (11).

The pRN family plasmids contain all homologs of Orf56/CopG and the replication protein Orf904 in their conserved regions. In the plasmids pTAU4, pTIK4, pORA1, and pIT3 as well as in conjugative plasmids and viruses of *Sulfolobus*, CopG-like proteins have been annotated preceding an open reading frame predicted to code for a replication protein (17). In some cases (e.g., pORA1 and pSSVx), the CopG-like proteins do not directly precede the replication protein but are followed by another short open reading frame encoding a protein also showing some similarity to Cop proteins (18). Inverted repeats as putative binding sites upstream of the

CopG-like proteins have been identified in pKEF9 and pARN4 but not in pORA1 and pTIK4 (18, 19). The transcription of two CopG-like proteins preceding a replication protein has been examined so far only for pIT3. Two distinct transcriptional units consisting of a transcript for the CopG-like proteins and one for the replication protein (35) have been found, unlike the case in pRN1. Therefore, care has to be taken when transferring the results obtained with pRN1 to other annotated CopG-like/replication protein couples from *Sulfolobus* plasmids.

The combined information obtained from the biochemical characterization of the proteins encoded by pRN1, the analysis of its transcription pattern, and the results of first genetic experiments has considerably improved our understanding of plasmid organization and copy number regulation. The mechanism of copy number regulation shows similarities to that of gram-positive bacterial rolling circle plasmids of the pMV158 family, whereas the replication protein encoded in the auto-regulated operon is a specific archaeal protein. Unexpectedly, in a large part of pRN1, transcriptional activity can be detected. This is reminiscent of the finding that transcripts are produced by at least half of the typical eukaryotic genome (7). The transcript levels are low, allowing the plasmid to stay inconspicuous without being a burden for the cell. Whether all of the transcripts have a function remains to be determined. Improvement of genetic systems for *Sulfolobus* will allow us to address still unresolved issues, such as the mode of replication of pRN1 or further mechanisms of regulation.

ACKNOWLEDGMENTS

We are grateful to Christa Schleper for the original *S. islandicus* REN1H1 strain, to Alexandra Kessler, Monika Häring, and David Prangishvili for helpful discussions, to Sonja Albers for her advice and encouragement with the virus vector system, and to Kirsten Beck for helpful comments on the manuscript.

This work was supported by the DFG (grant Li913/3 to G.L.).

REFERENCES

- Albers, S. V., M. Jonuscheit, S. Dinkelaker, T. Urich, A. Kletzin, R. Tampé, A. J. M. Driessen, and C. Schleper. 2006. Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Environ. Microbiol.* **72**:102–111.
- Andersson, A., M. Lundgren, S. Eriksson, M. Rosenlund, R. Bernander, and P. Nilsson. 2006. Global analysis of mRNA stability in the archaeon *Sulfolobus*. *Genome Biol.* **7**:R99.
- Arnold, H. P., Q. She, H. Phan, K. M. Stedman, D. Prangishvili, I. Holz, J. K. Kristiansson, R. A. Garrett, and W. Zillig. 1999. The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. *Mol. Microbiol.* **34**:217–226.
- Bell, S. D., P. L. Kosa, P. B. Sigler, and S. P. Jackson. 1999. Orientation of the transcription preinitiation complex in archaea. *Proc. Natl. Acad. Sci. USA* **96**:13662–13667.
- Bell, S. D., S. S. Cairns, R. L. Robson, and S. P. Jackson. 1999. Transcriptional regulation of an archaeal operon in vivo and in vitro. *Mol. Cell* **4**:971–982.
- Bell, S. D., and S. P. Jackson. 2000. Mechanism of autoregulation by an archaeal transcriptional repressor. *J. Biol. Chem.* **275**:31624–31629.
- Bickel, K. S., and D. R. Morris. 2006. Silencing the transcriptome's dark matter: mechanisms for suppressing translation of intergenic transcripts. *Mol. Cell* **22**:309–316.
- Bini, E., V. Dikshit, K. Dirksen, M. Drozda, and P. Blum. 2002. Stability of mRNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *RNA* **8**:1129–1136.
- Brantl, S. 1994. The copR gene product of plasmid pIP501 acts as a transcriptional repressor at the essential repR promoter. *Mol. Microbiol.* **14**:473–483.
- Brantl, S. 2004. Plasmid replication control by antisense RNAs, p. 47–62. In B. E. Funnell and G. J. Phillips (ed.), *Plasmid biology*. ASM Press, Washington, DC.
- Condo, I., A. Ciammaruconi, D. Benelli, D. Ruggero, and P. Londei. 1999. Cis-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* **34**:377–384.
- Dahlke, I., and M. Thomm. 2002. A Pyrococcus homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. *Nucleic Acids Res.* **30**:701–710.
- del Solar, G., P. Acebo, and M. Espinosa. 1997. Replication control of plasmid pLS1: the antisense RNAII and the compact rnaII region are involved in translational regulation of the initiator RepB synthesis. *Mol. Microbiol.* **23**:95–108.
- del Solar, G., and M. Espinosa. 2000. Plasmid copy number control: an ever growing story. *Mol. Microbiol.* **37**:492–500.
- del Solar, G., and M. Espinosa. 1992. The copy number of plasmid pLS1 is regulated by two trans-acting plasmid products: the antisense RNAII and the repressor protein, RepA. *Mol. Microbiol.* **6**:83–94.
- Ehresmann, B., P. Imbault, and J. H. Weil. 1973. Spectrophotometric determination of protein concentration in cell extracts containing tRNA's and rRNA's. *Anal. Biochem.* **54**:454–463.
- Garrett, R. A., P. Redder, B. Greve, K. Brügger, L. Chen, and Q. She. 2004. Archaeal plasmids, p. 377–392. In B. E. Funnell and G. J. Phillips (ed.), *Plasmid biology*. ASM Press, Washington, DC.
- Greve, B., S. Jensen, H. Phan, K. Brügger, W. Zillig, Q. She, and R. A. Garrett. 2004. Novel RepA-MCM proteins encoded in plasmids pTAU4, pORA1 and pTIK4 from *Sulfolobus neozealandicus*. *Archaea* **1**:319–325.
- Greve, B., S. Jensen, K. Brügger, W. Zillig, and R. A. Garrett. 2004. Genomic comparison of archaeal conjugative plasmids from *Sulfolobus*. *Archaea* **1**:231–239.
- Grogan, D. W. 1989. Phenotypic characterization of the archaeobacterial genus *Sulfolobus*: comparison of five wild-type strains. *J. Bacteriol.* **171**:6710–6719.
- Jonuscheit, M., E. Martusewitsch, K. M. Stedman, and C. Schleper. 2003. A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Mol. Microbiol.* **48**:1241–1252.
- Keeling, P. J., H.-P. Klenk, R. K. Singh, O. Feeley, C. Schleper, W. Zillig, W. F. Doolittle, and C. W. Sensen. 1996. Complete nucleotide sequence of the *Sulfolobus islandicus* multicopy plasmid pRN1. *Plasmid* **35**:141–144.
- Keeling, P. J., H.-P. Klenk, R. K. Singh, M. E. Schenk, C. W. Sensen, W. Zillig, and W. F. Doolittle. 1998. *Sulfolobus islandicus* plasmids pRN1 and pRN2 share distant but common evolutionary ancestry. *Extremophiles* **2**:391–393.
- Kessler, A., A. B. Brinkman, J. van der Oost, and D. Prangishvili. 2004. Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon *Sulfolobus*. *J. Bacteriol.* **186**:7745–7753.
- Kletzin, A., A. Lieke, T. Urich, R. L. Charlebois, and C. W. Sensen. 1999. Molecular analysis of pDL10 from *Acidiphilium ambivalens* reveals a family of related plasmids from extremely thermophilic and acidophilic archaea. *Genetics* **152**:1307–1314.
- Lipps, G., M. Stegert, and G. Krauss. 2001. Thermostable and site-specific DNA binding of the gene product ORF56 from the *Sulfolobus islandicus* plasmid pRN1, a putative archaeal plasmid copy control protein. *Nucleic Acids Res.* **29**:904–913.
- Lipps, G., P. Ibanez, T. Stroessenreuther, K. Hekimian, and G. Krauss. 2001. The protein ORF80 from the acidophilic and thermophilic archaeon *Sulfolobus islandicus* binds highly site-specifically to double-stranded DNA and represents a novel type of basic leucine zipper protein. *Nucleic Acids Res.* **29**:4973–4982.
- Lipps, G., S. Röther, C. Hart, and G. Krauss. 2003. A novel type of replicative enzyme harboring ATPase, primase and DNA polymerase activity. *EMBO J.* **22**:2516–2525.
- Lipps, G., A. O. Weinzierl, G. von Scheven, C. Buchen, and P. Cramer. 2004. Structure of a bifunctional DNA primase-polymerase. *Nat. Struct. Mol. Biol.* **11**:157–162.
- Lipps, G. 2004. The replication protein of the *Sulfolobus islandicus* plasmid pRN1. *Biochem. Soc. Trans.* **32**:240–243.
- Lipps, G. 2006. Plasmids and viruses of the thermoacidophilic crenarchaeote *Sulfolobus*. *Extremophiles* **10**:17–28.
- Littlefield, O., Y. Korkhin, and P. B. Sigler. 1999. The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc. Natl. Acad. Sci. USA* **96**:13668–13673.
- Martusewitsch, E., C. W. Sensen, and C. Schleper. 2000. High spontaneous mutation rate in the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by transposable elements. *J. Bacteriol.* **182**:2574–2581.
- Peng, X., I. Holz, W. Zillig, R. A. Garrett, and Q. She. 2000. Evolution of the family of pRN plasmids and their integrase-mediated insertion into the chromosome of the crenarchaeon *Sulfolobus solfataricus*. *J. Mol. Biol.* **303**:449–454.
- Prato, S., R. Cannio, H.-P. Klenk, P. Contursi, M. Rossi, and S. Bartolucci. 2006. pIT3, a cryptic plasmid isolated from the hyperthermophilic crenarchaeon *Sulfolobus solfataricus* IT3. *Plasmid* **56**:35–45.
- Purschke, W. G., and G. Schäfer. 2001. Independent replication of the

- plasmids pRN1 and pRN2 in the archaeon *Sulfolobus islandicus*. FEMS Microbiol. Lett. **200**:97–102.
37. **Reiter, W.-D., P. Palm, and W. Zillig.** 1988. Transcription termination in the archaeobacterium *Sulfolobus*: signal structures and linkage to transcription initiation. Nucleic Acids Res. **16**:2445–2459.
 38. **Santangelo, T. J., and J. N. Reeve.** 2006. Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. J. Mol. Biol. **355**:196–210.
 39. **Sartorius-Neef, S., and F. Pfeifer.** 2004. In vivo studies on putative Shine-Dalgarno sequences of the halophilic archaeon *Halobacterium salinarum*. Mol. Microbiol. **51**:579–588.
 40. **Schleper, C., K. Kubo, and W. Zillig.** 1992. The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with the viral DNA. Proc. Natl. Acad. Sci. USA **89**:7645–7649.
 41. **Tang, T.-H., N. Polacek, M. Zywicki, H. Huber, K. Brügger, R. A. Garrett, J. P. Bachellerie, and A. Hüttenhofer.** 2005. Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. Mol. Microbiol. **55**:469–481.
 42. **Tolstrup, N., C. W. Sensen, R. A. Garrett, and I. G. Clausen.** 2000. Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*. Extremophiles **4**:175–179.
 43. **Torarinsson, E., H.-P. Klenk, and R. A. Garrett.** 2005. Divergent transcriptional and translational signals in archaea. Environ. Microbiol. **7**:47–54.
 44. **Vierke, G., A. Engelmann, C. Hebbeln, and M. Thomm.** 2003. A novel archaeal transcriptional regulator of heat-shock response. J. Biol. Chem. **278**:18–26.
 45. **Wilson, K. S., and P. H. von Hippel.** 1995. Transcription termination at intrinsic terminators: the role of the RNA hairpin. Proc. Natl. Acad. Sci. USA **92**:8793–8797.
 46. **Wolin, E. A., M. J. Wolin, and R. S. Wolfe.** 1963. Formation of methane by bacterial extracts. J. Biol. Chem. **238**:2882–2886.
 47. **Zillig, W., A. Kletzin, C. Schleper, I. Holz, D. Janekovic, J. Hain, M. Lanzendörfer, and J. K. Kristiansson.** 1994. Screening for Sulfolobales, their plasmids, and their viruses in Icelandic solfataras. Syst. Appl. Microbiol. **16**:606–628.
 48. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244**:48–52.