GENETIC AND MOLECULAR CHARACTERIZATION OF THE VIRUS/PLASMID HYBRID pSSVx ISOLATED FROM *S.ISLANDICUS* REY 15/4

Tiziana Aucelli

Dottorato in Scienze Biotecnologiche – XX ciclo
Indirizzo Biotecnologie Molecolari
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XX ciclo
Indirizzo Biotecnologie Molecolari
Università di Napoli Federico II

GENETIC AND MOLECULAR CHARACTERIZATION OF THE VIRUS/PLASMID HYBRID pSSVx ISOLATED FROM *S.ISLANDICUS* REY 15/4

Tiziana Aucelli

Dottoranda:	Tiziana Aucelli
Relatore:	Prof. Simonetta Bartolucci
Coordinatore:	Prof. Giovanni Sannia

Alle mie preziose ed inesauribili
fonti di energia e amore incondizionato
papà, mamma, Andrea e Marco
dedico la presente con
un sentito GRAZIE!

INDEX

ABBREVIATIONS

AbrB	antibiotic resistance protein B
bp	base pairs
BRE	transcription factor B recognition element
BSA	albumina di siero bovino
CPs	conjugative plasmids
CRISPRs	clustered regularly interspaced short palindrome repeats
Da	dalton
dATP	2'-deoxyadenosine 5'-triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DSM	Deutsche Sammlung von Mikroorganismen
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
h	hours
HGT	Horizontal gene transfer
hph	hygromycin B phosphotransferase
HTH	helix-turn-helix
INR	initiator element
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IS	insertion sequence
LB	Luria Bertani medium
Lrp	leucine-responsive regulation protein
min	minutes
MITEs	miniature inverted transposable elements
OD	optical density
orf	open reading frame
PCR	polimerase chain reaction
PVDF	polyvinylidene difluoride
RHH	ribbon-helix-helix
RNA	ribonucleic acid
SDS	Sodium Dodecyl Sulphate

SDS-PAGE	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
TBE	45 mM Tris-borate, 1mM EDTA pH 8.0
TBP	TATA-binding protein
TFB	transcriptional factor B
Tris	tris(hydroxymethyl)aminomethane
TYS	tryptone yeast sucrose (medium)
UTR	untraslated region
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside

SUMMARY

This study is part of a large project aimed at developing innovative genetic systems for hyperthermophiles *Archaea* and in particular for the crenarchaeon *S.solfataricus*. The work was developed through different steps:1) physiological characterization of the *S.islandicus* REY 15/4 host-SSV2 virus interaction, 2) construction of the pSSVx-derived *E.coli-S.solfataricus* shuttle vector, 3) structural-functional analysis of the gene product ORF c68 of the pSSVx. The physiological characterization of the SSV2-pSSVx system revealed that in the natural host there is an induction of the replication of the two genetic elements when the host cells reach the late-stationary phase. Conversely, the cellular content of SSV2 and pSSVx remains fairly low in the foreign host *S.solfataricus.* The accumulation of episomal DNA in the former case cannot be traced to decreased packaging activity because of a simultaneous increase in the virus titre in the medium. In addition, the natural host growth-inhibition concurs with the induction of the SSV2 replication. In *S.islandicus* REY 15A, a strain which is closely related to the natural host, it was found that the SSV2 replication process was induced in the same way as in the natural host REY 15/4. The satellite pSSVx virus was engineered to obtain an *E.coli-S.solfataricus* shuttle vector for gene transfer and expression in *S.solfataricus* by fusing site-specifically the pSSVx chromosome with an *E.coli* plasmid replicon and the ampicillin resistance gene. The pSSVx-based vector, named pMSSV, was proven functional like the parental virus, namely it was able to spread efficiently through infected *S.solfataricus* cells with no rearrangements, recombination or integration into the host chromosome. The high copy number of the artificial genetic element was found comparable with that calculated for the wild-type pSSVx in the new host cells, with no need of genetic markers for vector maintenance in the cells and for transfomant enrichment. A pMSSV-derivative plasmid carrying an expression cassette of the *lacS* gene encoding the β-glycosidase from *S.solfataricus* under the control of the *Sulfolobus* chaperonine (thermosome tf55) heat shock promoter was also able to drive the expression of a functional enzyme. Complementation of the β-galactosidase deficiency in a deletion mutant strain of *S.solfataricus* demonstrated that *lacS* gene was an efficient marker for selection of single transformants on solid minimal lactose medium. *orf c68* is the only reverse oriented pSSVx putative gene and has no homologs in the crenarchaeal family of pRN plasmids. It is highly similar (42/53% identity/similarity) to *orf c56* of pSSVi and also finds its homologs in the integrated elements of the *S.tokodaii* and *S.acidocaldarius* genomes, albeit the biological function is still unknown. Secondary structure prediction revealed the presence in the translated ORF c68 protein of a SpoVT/Abr-B like domain which is typical of a bacterial family of transcription factors. The ORF c68 has been expressed in *E.coli* and the DNA-binding properties of the protein have been investigated by EMSA experiments. The expression level of the ORF c68 protein *in vivo* has been detected by western blotting in different growth phases. We found that the protein ORF c68 binds specifically its own promoter region, this suggesting a role as an auto-regulator of its own expression. Foot printing and EMSA experiments have revealed that ORF c68 binds also a region encompassing the promoter region of ORF 60. This finding suggests an hypothetical role of ORF c68 also in modulating the expression of the contiguous *orf 60* gene. The *orf c68* and *orf 60* belong to the same complex transcriptional network, involving also *orf 91* and *orf 892 RepA* (encoding a putative

replication protein) adjacent genes, probably involved in regulating the *RepA* gene expression.

RIASSUNTO

L'avvento delle biotecnologie nel settore dell'industria ha condotto ad una serie di importanti vantaggi; l'utilizzo degli enzimi in processi industriali di varia natura, in luogo della tradizionale tecnologia chimica, ha consentito infatti una minimizzazione dell'impatto ambientale, dato il ridotto consumo energetico e la biodegradabilità dei composti. Non meno importante è la più elevata specificità delle reazioni enzimatiche rispetto a processi sintetici di tipo chimico, con conseguente riduzione di prodotti collaterali indesiderati e dunque abbattimento del potenziale inquinante. Un limite dell'enzimologia industriale è però rappresentato dall'instabilità dei biocatalizzatori impiegati. Gli enzimi isolati dagli *Archaea* ipertermofili presentano invece, rispetto alle controparti mesofile, un'elevata termoresistenza che unita ad una pronunciata stabilità verso i comuni agenti denaturanti, li rende candidati ideali per le biotecnologie industriali. Le elevate temperature cui si ricorre in determinati processi industriali, soprattutto nel campo alimentare per la trasformazione dei carboidrati, consentono, tra gli altri vantaggi, di ottenere una riduzione della flora microbica contaminante, una maggiore solubilità dei substrati ed una più elevata velocità di reazione.

Questo progetto di ricerca si è fondato principalmente sul tentativo di rendere più concreta la possibilità di utilizzare i microrganismi ipertermofili come "bioproduttori" di enzimi termostabili; in particolare si è proposto la messa a punto di opportuni sistemi genetici tramite i quali studiare e regolare l'espressione di proteine sia omologhe che eterologhe all'interno di cellule del crenarchaeon *Sulfolobus solfataricus* che si è rivelato, tra gli *Archaea* ipertermofili, un microrganismo modello sulla base della versatilità che lo caratterizza.

L'impegno sperimentale è stato dunque mirato a fornire un valido contributo per utilizzare *S.solfataricus* come "*cell factory*" per la produzione di enzimi termostabili. A tal proposito un forte limite è sempre stato costituito dalla scarsa disponibilità di opportuni vettori d'espressione per gli organismi archaeali. Mentre per i Batteri la progettazione di vettori é ormai notevolmente avanzata, per gli *Archaea* sviluppi significativi in tal senso si sono avuti soltanto in tempi relativamente recenti, basandosi sulla manipolazione degli svariati elementi genetici, sia di natura plasmidica che virale, rinvenuti soprattutto nell'ambito dei diversi ceppi di sulfolobali. Le difficoltà alla base della messa a punto di sistemi genetici efficienti possono essere svariate; una bassa efficienza di trasformazione ed una scarsa stabilità del costrutto, che può pertanto andare incontro a riarrangiamenti all'interno della cellula ospite, possono costituire dei validi esempi. Più efficaci rispetto ad altri sono risultati diversi vettori navetta per *S.solfataricus* ottenuti a partire dal virus SSV1 che si è pertanto rivelato un potente strumento genetico; esso può essere agevolmente trasferito in cellule di *S.solfartaricus* mediante elettroporazione (Schleper *et al.*, 1992).

Dal ceppo REY 15/4 di *S.islandicus* sono stati isolati due interessanti elementi genetici denominati pSSVx ed SSV2, un plasmide ed un virus rispettivamente. Trattasi in realtà dell'unico sistema plasmide/virus helper isolato in microrganismi archaeali. Gli elementi extracromosomali in oggetto hanno notevolmente stimolato la nostra curiosità scientifica. In particolare il pSSVx è caratterizzato da un intrigante dualismo: esso è un "ibrido genetico" tra un virus ed un plasmide e, differentemente

dagli altri membri della famiglia pRN cui appartiene, risulta rivestito da un vero e proprio involucro virale. L'ingresso del pSSVx all'interno delle cellule è mediato da un processo di infezione che necessita però della presenza del virus helper SSV2, isolato insieme al pSSVx ed appartenente alla famiglia *Fuselloviridae*. Se si pensa al limite imposto allo sviluppo di sistemi di espressione efficienti dalla bassa efficienza di trasformazione degli *Archaea*, appare chiara l'enorme utilità di servirsi di un meccanismo naturale, quale è l'infezione, per agevolare l'ingresso del materiale genetico nelle cellule ospiti. Peraltro è già stata dimostrata la trasmissibilità inter-specie, da *S.islandicus* a *S.solfataricus*, dei virioni (sia SSV2 che pSSVx) prodotti da cellule parentali naturali del *S.islandicus* REY 15/4 (Arnold H. P. *et al*.,1999). Alla luce delle caratteristiche sopra elencate il pSSVx appariva come una promettente base di partenza per la costruzione di un vettore per *S.solfataricus* dalle caratteristiche innovative.
In una prima fase del progetto l'attenzione è stata incentrata sulla caratterizzazione fisiologica del sistema SSV2/pSSVX all'interno sia dell'ospite parentale sia del meglio caratterizzato ospite *S.solfataricus*. E' emerso che in cellule di REY 15/4 ha luogo una forte induzione della replicazione di SSV2 in corrispondenza della fase di crescita stazionaria; il fenomeno induttivo coincide inoltre con l'inibizione della crescita delle cellule ospiti. Estraendo il DNA totale ad intervalli di tempo regolari dalle cellule in crescita, è stato possibile stabilire che nell'arco delle prime 24 ore di incubazione non sono osservabili le forme episomali di SSV2 e pSSVx laddove l'incremento quantitativo dei due elementi raggiunge il suo massimo tra le 26 e le 30 ore di incubazione. Una valutazione quantitativa del fenomeno mediante Southern blotting ha consentito di stabilire che la forma episomale del virus SSV2 si mantiene a livello di una sola copia per cromosoma durante la fase di crescita esponenziale mentre nelle quattro ore successive all'arresto della crescita essa aumenta di circa 50 volte. Quanto sinora esposto circa il processo di induzione della replicazione di SSV2 è stato riscontrato anche per pSSVx; l'induzione sia del virus che del plasmide, inoltre, dipende unicamente da un fattore endogeno, originato dall'ospite o dal virus, e non da fattori estrinseci. L'induzione del virus SSV2 nella fase stazionaria, peraltro, si accompagna ad un notevole incremento della carica virale, ad indicare che l'incremento intracellulare del numero di copie di SSV2 non deriva da una diminuzione dell'attività di packaging di nuove particelle virali ma da un effettivo aumento della velocità di replicazione dei genomi virali.
Come accennato sopra, le cellule dell'ospite REY 15/4 vanno incontro ad una notevole inibizione della crescita in coincidenza dell'induzione della replicazione virale. Al fine di stabilire se l'inibizione della crescita SSV2-dipendente nelle cellule ospiti fosse un processo reversibile, sono stati allestiti esperimenti in cui le colture cellulari venivano opportunamente diluite dopo che l'induzione della replicazione di SSV2 aveva avuto luogo. È emerso quanto segue:

- le cellule ospiti recuperano la capacità di crescere in maniera esponenziale dopo una fase di latenza di lunghezza variabile
- la quantità episomale del virus SSV2 reverte a circa una copia per cromosoma a dimostrazione della reversibilità del processo di induzione della replicazione.

Il fenomeno di induzione della replicazione del virus SSV2 è stato inoltre dimostrato essere indipendente dalla presenza del pSSVx; all'interno dell'ospite REY31A, derivante dal REY15/4 curato del pSSVx, il fenomeno induttivo in assenza del pSSVx avviene esattamente con le medesime modalità osservate nell'ospite naturale. Dunque la presenza del virus satellite pSSVx non "contribuisce" ad indurre la replicazione del proprio helper.

Per approfondire la nostra indagine fisiologica abbiamo voluto verificare se il fenomeno di induzione della replicazione del pSSVx ed SSV2 si verificasse anche all'interno di cellule del ceppo P2 *S.solfataricus.* Dagli esperimenti condotti è risultato che nell'ospite "non parentale" la replicazione dei due elementi genetici è correlata alla fase di crescita ma l' induzione della replicazione non ha luogo. In *S.solfataricus* la produzione di particelle virali inizia durante la fase di crescita esponenziale, raggiunge il suo massimo prima che le cellule raggiungano la fase stazionaria ed eventualmente decrementa. Un'altra caratteristica interessante relativa all'infezione di *S.solfataricus* da parte di SSV2 risiede nel fatto che la presenza stabile del virus non sortisce alcun effetto sulla crescita delle cellule ospiti laddove in cellule appena infettate la crescita è evidentemente inibita. Ciò suggerisce che nel corso dell'infezione si possano stabilire delle interazioni tra il virus SSV2 ed il suo ospite tali da garantire un'armonica coesistenza.
La caratterizzazione fisiologica del sistema genetico pSSVx-SSV2 ha permesso innanzitutto di stabilire che il plasmide è in grado di seguire fedelmente il "comportamento" del suo helper relativamente a *spreading* e replicazione nell'ospite naturale così come in *S.solfataricus*. Per il virus SSV2 è emerso un probabile ruolo fisiologico nel garantire la sopravvivenza dell'ospite mediante l'arresto della crescita, coincidente proprio con l'induzione della replicazione virale. Il virus potrebbe dunque avere il rulo di "congelare" transitoriamente il metabolismo dell'ospite in uno stato inattivo, qualora le condizioni esterne risultassero troppo sfavorevoli, o alternativamente impedire all'ospite stesso di reagire troppo rapidamente ai cambiamenti che si verificano nelle nicchie ecologiche. E' ovvio come questi fenomeni siano governati da un'affascinante biologia molecolare e che studi futuri potranno rivelare più nel dettaglio i meccanismi alla base delle interazioni ospite-virus così come della co-evoluzione dei virus e dei rispettivi ospiti negli *Archaea*.
Il pSSVx ha dunque costituto la base di partenza per la realizzazione di un costrutto *shuttle E.coli/S.solfataricus.* Vista l'incapacità del pSSVx di propagare autonomamente nelle cellule, è stato innanzitutto necessario ottenere un ceppo ospite di *S.solfataricus* stabilmente infettato con il virus helper SSV2. Si è pertanto scelto di infettare con l'SSV2 il ceppo *S.solfataricus* Gθ white (Bartolucci *et al*., 2003), un mutante spontaneo privo, rispetto al parentale, di attività β-galattosidasica e pertanto utile a sfruttare il meccanismo genetico della complementazione per l'identificazione selettiva dei ricombinanti. Il gene *lacS*, opportunamente clonato all'interno del vettore navetta, sarebbe divenuto perciò un utile e pratico marcatore genetico. Un'accurata analisi della sequenza del pSSVx è stata necessaria al fine di individuare una regione che potesse essere interrotta per linearizzare il pSSVx e ligarlo al vettore pUC19 di *E.coli*, fondamentale per garantire la propagazione del vettore anche nell'ospite batterico. Lo studio della sequenza del pSSVx aveva evidenziato una regione a valle dell'*orf 288* dove immediatamente adiacente ad un' evidente sequenza segnale di terminazione della traduzione veniva riscontrato un sito *Afl*II (posizione 2182). L' "interruzione" della sequenza del pSSVx a tale livello è stata scelta in quanto garantiva l'integrità di tutte le *orfs* presenti sul genoma, necessaria per preservare il costrutto risultante da eventuali alterazioni fisiologiche (stabilità, numero di copie) a carico della porzione archaeale. Il vettore denominato pSSVrt è stato dunque ottenuto mediante linearizzazione del pSSVx in corrispondenza del sito *Afl*II, condotta mediante digestione enzimatica parziale in presenza di etidio bromuro (Stedman *et al*., 1999), e successivo clonaggio nel sito *Sma*I del pUC19. Il pSSVrt presentava un buon livello di replicazione e propagazione in cellule di *S.solfataricus,* all'interno delle quali veniva trasferito per elettroporazione.

Al fine di testare l'efficacia del costrutto pSSVrt come veicolo per il trasferimento e l'espressione di geni omologhi in *S.solfataricus* è stata costruita un'apposita cassetta d'espressione: la sequenza codificante per una β-glicosidasi (*lacS*) con una sua estesa regione 3'-fiancheggiante (3'-UTR) è stata fusa a valle di un frammento di 448 bp contenente la sequenza promotrice del gene codificante per la subunità a della chaperonina TF55 ed i primi cinque codoni di tale *orf*. Tale fusione genica era già stata dimostrata in grado di complementare efficientemente l'assenza di attività β-galattosidasica in *Sullfolobus* quando clonata all'interno di un vettore shuttle derivante dal virus SSV1 (Jonuscheit *et al*., 2003). Il costrutto d'espressione pSSVrt-*lacS* è stato utilizzato per trasformare cellule del ceppo GθW/SSV2 di *S. solfataricus*; data la delezione della sequenza genica *lacS* dell'ospite, l'attività β-galattosidasica, conformemente al meccanismo genetico di complementazione, sarebbe stata attribuibile esclusivamente al gene trasferito. Si è proceduto con un test per rivelare attività β-galattosidasica sui trasformanti, utilizzando come controllo positivo cellule di *S. solfataricus* P2; l'attività enzimatica è risultata nulla confermando l'assenza del vettore d'espressione non rilevabile nemmeno a livello del DNA extracromosomale. Dal momento che le dimensioni del costrutto pSSVrt-*lacS* avevano raggiunto ~11Kb, era lecito pensare che il *packaging* del DNA fosse inficiato dalle sue eccessive dimensioni; ciò si rifletteva ovviamente in un ostacolo alla propagazione del vettore all'interno delle cellule elettroporate. Dunque, la necessità di agevolare i meccanismi di replicazione e diffusione del vettore shuttle in colture di *S.solfataricus*, è stata fronteggiata riducendo le dimensioni dell'intero costrutto attraverso la costruzione di un mini-plasmide derivante dal pUC19, contenente esclusivamente le sequenze responsabili per la replicazione (ColE1) e resistenza ad ampicillina (gene codificante per la β-lattamasi). La sequenza del pSSVx veniva escissa dal pSSVrt e fusa al mini-plasmide di *E.coli* così da ottenere il nuovo vettore denominato pMSSV, nel quale era inoltre inserito il polylinker del pUC28 in sostituzione del precedente. All'interno del costrutto pMSSV abbiamo dunque clonato una nuova cassetta di espressione del gene *lacS* (gentilmente fornita dalla Dott.ssa C.Schleper) nella quale la regione 3'-UTR risultava ridotta in dimensioni rispetto alla precedente. Il vettore d'espressione shuttle pMSSV-*lacS*, le cui dimensioni risultavano pari a 9616 bp, presentava un buon livello di propagazione sia nell'ospite batterico che in quello archaeale. Il costrutto era in grado di diffondere in cellule di *S.solfataricus* analogamente al plasmide parentale pSSVx, nonché di complementare con successo l'assenza di attività β-galattosidasica all'interno del ceppo ospite GθW. La realizzazione del costrutto pMSSV era dunque risultata fruttuosa nell'agevolare la propagazione del derivativo d'espressione a conferma del fatto che le dimensioni del precedente costrutto pSSVrt-*lacS* risultavano compromettenti per il suo efficiente assemblaggio all'interno di particelle virali.

Il costrutto definitivo pMSSV presenta diversi vantaggi che lo rendono particolarmente versatile. Esso presenta innanzitutto un numero di copie per cellula elevato, tale da garantire un buon livello di espressione dei geni clonati al suo interno. In secondo luogo, la capacità del vettore di mantenersi in forma episomale evita la possibilità, verificatasi per altri vettori, che l'integrazione nel genoma della cellula ospite, con conseguente decremento quantitativo della forma libera, si traduca in una riduzione dell'espressione del gene clonato con gli ovvi svantaggi che ne derivano. L'assenza nel costrutto di marcatori selettivi per *S.solfataricus*, inoltre, è senz'altro importante nell'assicurare una più rapida ed efficace rivitalizzazione dei trasformanti; l'individuazione dei ricombinanti risulta meno dispendiosa anche in termini meramente economici, dati i costi elevati legati alla preparazione di terreni

arricchiti per esercitare una pressione selettiva. Il sistema potrebbe ovviare anche ai problemi spesso riscontrati nell'espressione di geni da organismi ipertermofili in ospiti mesofili, relativamente a *folding* non corretto o modifiche post-traduzionali incompatibili.
Alla luce di quanto sinora esposto, il vettore shuttle pMSSV si presenta come un promettente strumento per rendere più concreto lo studio di proteine omologhe ipertermofile, nonchè aumentare la termostabilità di proteine eterologhe mesofile mediante evoluzione diretta, con i conseguenti vantaggi applicativi in campo biotecnologico.
Nell'ottica di ottimizzare le caratteristiche del vettore pMSSV, l'ultima fase del presente lavoro è stata incentrata sulla caratterizzazione biochimica di alcune *orfs* presenti sul genoma del pSSVx. Mentre studi su ORF 91, ORF 76, ORF 154 e ORF 288 sono ancora in corso, risultati interessanti sono stati ottenuti per la proteina ORF c68. Il gene *c68* è l'unico con orientamento antiparallelo rispetto a tutti gli altri disposti sul genoma del pSSVx. La proteina corrispondente ORF c68 non trova omologhe nell'ambito della famiglia di plasmidi archaeali pRN ma presenta elevata similarità di sequenza con l' ORFc56 del pSSVi (Wang *et al*. 2007); presenta inoltre omologia con sequenze codificanti presenti negli elementi integrati dei genomi di *S. tokodaii* e *S.acidocaldarius.* Abbiamo dunque condotto una previsione di struttura secondaria dalla quale è emersa la presenza, all'interno della proteina ORF c68, di un dominio strutturale SpoVT/Abr-B *like* tipicamente riscontrato in una famiglia di fattori trascrizionali procariotici di cui Abr-B è capostipite. La proteina ricombinante ORF c68 è stata ottenuta in cellule BL21-*Codon Plus* (DE3) RIL di *Escherichia coli* e purificata all'omogeneità ricorrendo ad una cromatografia a scambio ionico Resource S, seguita da una cromatografia ad esclusione molecolare. Il livello di espressione *in vivo* della proteina ORF c68, saggiato mediante esperimenti di Western blotting, era rilevabile soltanto nella fase di crescita tardo-esponenziale dell'ospite naturale REY 15/4, in corrispondenza dell'induzione della replicazione del pSSVx. Incoraggiati dalle previsioni di struttura secondaria, nonché da caratteristiche intrinseche di ORF c68 quali punto isoelettrico altamente basico e piccole dimensioni, abbiamo investigato sulla capacità della proteina di legare il DNA mediante dei saggi di ritardo della mobilità elettroforetica (EMSA). Dai nostri esperimenti è emerso che la proteina ORF c68 lega specificamente la propria regione promotrice, costituita da un frammento di DNA che si estende 50 nucleotidi a monte del codone di inizio della traduzione. Tale risultato suggerisce un coinvolgimento di ORF c68 nella regolazione della propria espressione genica. Analisi per *foot printing* ed ulteriori saggi EMSA condotti successivamente, hanno evidenziato che la proteina ORF c68 lega anche una regione intergenica di DNA che comprende, oltre al proprio promotore, la regione promotrice della contigua *orf 60*. Abbiamo dunque ipotizzato che la proteina in oggetto possa avere anche un ruolo nella modulazione dell'espressione genica dell'*orf 60.* L'esistenza di una relazione tra la trascrizione di *orf c68* e quella del gene contiguo *orf 60* è convalidata dall'analisi trascrizionale condotta da Contursi *et al.* (2007): i due geni in questione appartengono al medesimo complesso trascrizionale, che coinvolge anche le adiacenti *orf* 91 ed *orf* 892 *RepA* (codificante per una putativa replicasi), ed entrambe le proteine corrispondenti sono probabilmente coinvolte nella regolazione dell'espressione del gene *RepA*.
Alla luce di quanto sinora esposto, la proteina ORF c68 potrebbe avere un ruolo chiave nel controllo della replicazione del pSSVX modulando sia la propria espressione che quella del gene *orf 60*.

Analisi di cristallografia sulla proteina ORFc68 sono attualmente in corso e saranno senz'altro utili a fornire informazioni strutturali per rendere più completa la caratterizzazione della proteina. Ulteriori investigazioni sul ruolo fisiologico di ORF c68 potranno contribuire a delucidare i meccanismi della regolazione trascrizionale nonchè della propagazione di virus e plasmidi in microrganismi archaeali. Da un punto di vista squisitamente evolutivo, inoltre, poichè geni codificanti per fattori trascrizionali SpoVT/Abr-B *like* sono presenti sui genomi del pSSVi e dei virus STSV1 and SIFV, isolati da *S. tengchongensis* e *S. islandicus*, rispettivamente (Prangishvili D. *et al*., 2006), ma assenti da altri plasmidi procariotici, è verosimile ipotizzare che l'acquisizione da parte del pSSVx dell'*orf c68* (così come delle *orfs 154-288*) sia stata cruciale perché esso potesse "esplicare" la sua parziale natura virale e rispondere ai vari stimoli virali.

Introduction

The Kingdom *Crenarchaeota* in the larger Domain of *Archaea*

In 1970 Carl Woese and coworkers have performed an accurate analysis of the rRNA 16S and have identified the third Domain of life, *Archaea*, which was added to *Prokarya* and *Eukarya* already existing. The *Prokarya* Domain was so divided in *Bacteria* and *Archaea,* the latter composed by *Crenarchaea* (hyperthermophiles and thermoacidophiles), *Euryarchaea* (hyperthermophiles, methanogens and halophiles) and *Korarchaea*. *Archaea* are simple and microscopic organisms but they are quite distinct from more commonly encountered *Bacteria*. In fact, *Archaea* are more similar to *Eukarya* than to *Bacteria* in many important aspects. The crenarchaeal microbial species are characterized by the highest known growth temperatures of any organisms; they flourish under conditions which would quickly kill most of the "higher" organisms. As a rule, they grow best between 80° and 100°C (100°C = 212°F, the boiling point of water at sea level), and several species will not grow below 80°C. Numerous species also prefer to live under very acidic conditions in dilute solutions of hot sulfuric acid. Marine or terrestrial volcanic environments, such as hot springs and shallow or deep-sea hydrothermal vents, constitute the typical habitat of the most of the hyperthermophilic species. Recent analyses of genetic sequences obtained directly from environomental samples, however, indicate the existence of *Crenarchaeota* able to live at low temperature, which have not yet been cultivated. Among the crenarchaeal species, cells range in size from cocci <1µm in diameter to filaments over 100 µm in length; a wide range of cell shapes can be found (**Fig. 1**), including regular cocci clustered in grape-like aggregates (*Staphylothermus*), irregular, lobed cells (*Sulfolobus*), discs (*Thermodiscus*), very thin filaments (<0.5µm diameter; *Thermofilum*), and almost rectangular rods (*Thermoproteus, Pyrobaculum*). Most species possess flagella and are motile. Only a few members of the *Crenarchaeota* exhibit strange morphologies: *Pyrodictium* produces disk-shaped cells connected by extensive networks of proteinaceous fibers which may help it to attach to sulfur granules (**Fig. 1**) .

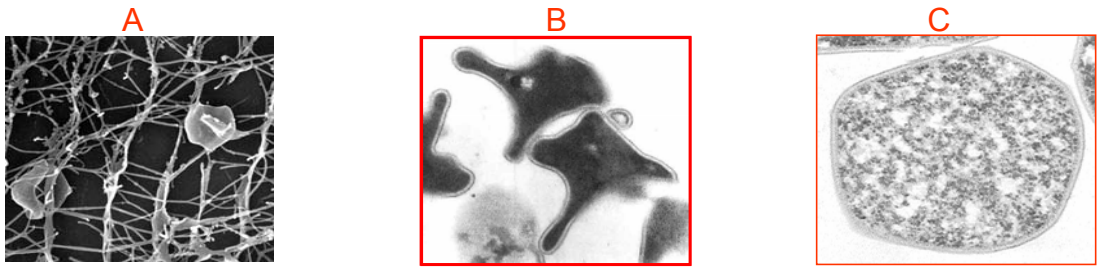

Fig. 1 panel A shows scanning electron micrograph of *Pyrodictium* cells, connected by a network of protein "fibers" (Rieger 1995). The transmission electron micrograph in the **panel B** shows an ultrathin section of several *S.acidocaldarius* cells (diameter of each is about 1 micrometer). The unusual structure of the *Sulfolobus* cell envelope is clearly visible around the margins of the cells. The envelope consists of a layer of a porous glycoprotein layer held out from the cell membrane a distance of about 15 nm. A micrograph of *S. solfataricus* cell is shown in **the panel C**.

Metabolically, *Crenarchaeota* are quite diverse, ranging from chemoorganotrophs to chemolithoautotrophs. They are anaerobes, facultative anaerobes or aerobes, and many utilize sulfur in some way for energy metabolism. Several species are primary producers of organic matter, using carbon dioxide as sole carbon source, and gaining energy by the oxidation of inorganic substances like sulfur and hydrogen, and reduction of sulfur or nitrate. Others grow on organic substrates by aerobic or anaerobic respiration or by fermentation. The most spectacular feature of the *Crenarchaeota*, however, is their tolerance to, and even preference for, extremes of acidity and temperature (**Fig. 2**). While many prefer neutral to slightly acidic pH ranges, members of the crenarchaeal order *Sulfolobales* flourish at pH 1-2 and die above pH 7. Optimum growth temperatures range from 75° to 105°C, and the maximum temperature of growth can be as high as 113°C (*Pyrobolus*). Most species are unable to grow below 70°C, although they can survive for long periods at low temperatures.

Fig. 2 Yellowstone National Park (Wyoming, **USA) and Solfatara (Pozzuoi, Italy).** Hot, sulfur-rich, acidic habitats, like this pool in Yellowstone (on the left) or the fumaroles in the Solfatara area (on the right), are often home to species of *Sulfolobus*. Note yellow deposits of sulfur on banks.

Genetic elements in the crenarchaeon *Sulfolobus*

The kingdom *Crenarchaeota* in the Domain *Archaea* encompasses the two orders of *Thermoproteales* and *Sulfolobales*, both of which comprise hyperthermophiles and extreme thermophiles. *Sulfolobus solfataricus* was one of the first organisms to be recognized as a member of the *Archaea* (Zillig W. *et al*., 1980). It is aerobic and heterotrophic and can easily be grown in liquid culture and also on solid media as colonies and lawns. The *S.solfataricus* genome is one of the first crenarchaeal genomes to be sequenced (Charlebois *et al*., 1998). Due to this early identification, *S. solfataricus* and its relatives have become model organisms for fundamental studies of *Archaea*. Studies of the genus *Sulfolobus* have been instrumental in understanding archaeal mechanisms of transposition (Martusewitsch E. *et al*., 2000), transfection (Schleper C. *et al*., 1992), transformation (Aravalli & Garrett, 1997; Cannio R. *et al*., 1998; Elferink M. *et al*., 1996; Stedman K.M. *et al*., 1999) and conjugation (Reilly & Grogan, 2001; Schleper C. *et al*., 1995). An impressive variety of mobile genetic

elements has recently been discovered in *Archaea* in general, and in *Sulfolobus* in particular: viruses, autonomous insertion sequence (IS) elements, non-autonomous miniature inverted repeat transposable elements (MITEs), small non-conjugative plasmids and large conjugative plasmids (Brugger *et al*., 2002; Prangishvili *et al*., 2001; Rice *et al*., 2001; Zillig *et al*., 1998). Although there have been impressive recent developments in *Sulfolobus* genetics, the tractability of suitable genetic tools represents still a bottleneck.

Plasmids

Among *Sulfolobus* strains conjugative and non-conjugative plasmids have been isolated. The first archaeal conjugative plasmid (CP), pNOB8, was isolated from a Japanese *Sulfolobus* isolate (Schleper *et al*., 1995). Since then, several other CPs have been isolated from colony-cloned strains of *S.islandicus,* and subsequently characterized. Apart from few insights gained from genome analysis, and from the observations that specific cell pairing precedes plasmid transfer (Shleper *et al*., 1995) and that transfer is probably selective and unidirectional (Prangishvili D. *et al*., 1998), little is known about the mechanism of archaeal conjugation. There are three subfamilies of these CPs: pNOB8, pING (Stedman K.M. *et al*., 2000) and pSOG2/4 (Erauso G. *et al.,* 2006). In the case of the pNOB8 and pING plasmids, derived variant plasmids were detected upon propagation. These occur as a result of deletion and recombination (She *et al*., 1998; Stedman *et al*., 2000). In fact, the genomes of *Sulfolobus* CPs are rather stable during propagation by conjugative transfer and replication in the recipients. However, they vary dramatically when they are spread after electroporation and subsequently propagated in the recipients (Prangishvili D. *et al*., 1998). The pING family from *Sulfolobus* is composed by the active variants pING1, -4, and -6 and the functionally defective variants pING2 and -3 (**Fig. 3**), which require the help of an active variant for spreading. Plasmid DNA isolated from clones of the sulfolobicin-producing *S. islandicus* isolate HEN2P2 (Zillig *et al*., 1994) was a complex mixture of variants after electroporation into, and propagation in, *S. solfataricus* P1. However, one strain from a single colony of the P1 transcipient culture yielded predominantly, and in high copy number, a single cccDNA. It produced a characteristic *Eco*RI restriction pattern and was designated pING1.The observation of the *in vivo* generation of complete (pING 2-4) and functionally defective (pING 3-6) pING variants and sequence comparison of these variants revealed novel and efficient mechanisms of genetic rearrangement in *Crenarchaeota*. The formation of pING4 from pING1, of pING2 from both of these, and of pING3 from pING6 in the transfer experiments was deduced from the changes of the *Eco*RI restriction patterns of the plasmid preparations. Comparing the conserved sequences of CPs with some non-conjugative derivatives has provided insight into proteins and DNA sequence motifs putatively involved in conjugation in *Archaea*. Sequence comparison of all *Sulfolobus* CPs revealed three distinct sequence domains. One well-conserved cluster of genes covering approximately 12 kbp of the plasmids' genomes apparently contains the conjugative functions. A second is the putative origin of replication. Finally there is a region proposed to encode replication proteins (Greve B. *et al.*, 2004). Only a few distant homologues to bacterial proteins involved in conjugative transfer (TraG, TrbE) and partitioning (ParA, ParB) have been found.

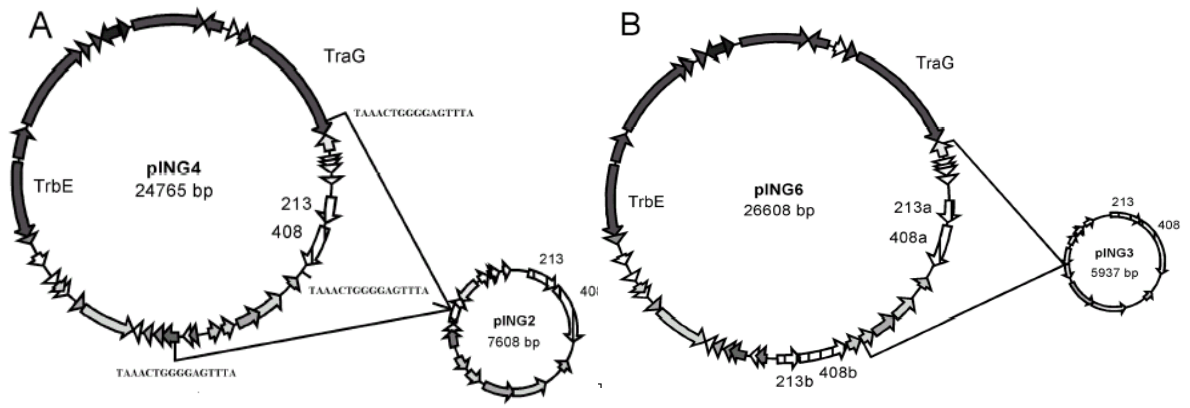

Fig. 3 ORF maps showing, schematically, the generation of the small CP pING2 from the large CP pING4 (A) and pING3 from the large CP pING6 (B). The sequences probably responsible for the generation of pING2 are shown in panel A.

A single strain of *S. islandicus* SOG2/4 was found to harbour two very different but related plasmids named pSOG1 and pSOG2 (Erauso G. *et al.*, 2006). Comparison of these novel CPs with the available counterparts has been used to further identify plasmid features that play key roles in conjugative transfer. The plasmids were separated from each other and transferred into *S. solfataricus*. pSOG1 has a high copy number and is not stable whereas the other conjugative plasmid, pSOG2, has a low copy number and is stably maintained. Plasmid pSOG2 is the first *Sulfolobus* CP found to have these characteristics. The genomes of both pSOG plasmids have been sequenced and were compared to each other and the available *Sulfolobus* CPs. Interestingly, comparison of the restriction endonuclease digestion patterns and Southern blotting of pSOG1 and pSOG2 showed that the two plasmids are very different (Prangishvili *et al*., 1998). Only 1/3 of the sequence of pSOG2 is conserved in pSOG1. Hence, 70 % of the pSOG1 and pSOG2 genomes is largely different and composed of a mixture of genes that often resemble counterparts in previously described *Sulfolobus* CPs. Unlike pNOB8, both pSOG plasmids lack the so-called clustered regularly interspaced short palindrome repeats (CRISPRs). CRISPRs are direct repeats found in the DNA of many *Bacteria* and *Archaea* which range in size from 24 to 48 base pairs. They usually show some dyad symmetry but are not truly palindromic and are separated by spacers of similar length. The mechanism of action of CRISPR systems is unknown, but may represent a prokaryotic analog of eukaryotic RNA interference systems. Hence, the CRISPRs may be used as a defense system of the host cell against plasmids or viruses by silencing the expression of their genes. Like other *Sulfolobus* CPs, the pSOG plasmids possess a gene encoding an integrase of the tyrosine recombinase family. This integrase probably mediates plasmid site-specific integration into the host chromosome at the highly conserved tRNA(Glu) loci.

Two non-conjugative cryptic plasmids named pRN1 and pRN2 from the strain REN1H1 of *S.islandicus* are name-giving for a small family of archaeal plasmids (**Fig. 4**). Both plasmids have been supposed to be dependent on each other because they are always found together in their natural host. Werner *et al*. 2001 have demonstrated that each of the plasmids can stably propagate and replicate on its own independently of the other plasmid (Werner G. *et al*., 2001). pRN plasmids share common, homologous molecular structures (Arnold H.P. *et al*., 1999). Beside the pRN plasmids, pHEN7, a 7.8-Kb plasmid from *S. islandicus* HEN7H2 (Zillig W. *et al*., 1998; Peng X. *et al*., 2000) and pDL10, a 7.6 Kb plasmid from the chemolithoautotrophic cenarchaeon *Acidianus ambivalens*, belong to this plasmid family. pDL10 was also found in several isolates from different geographical locations on Iceland. Another member of the family is the genetic element pSSVx (5.7Kb) from the *S.islandicus* strain REY15/4 which is a hybrid between a plasmid and a virus and spreading of the plasmid required co-transfection with either SSV2 or the related SSV1 fuselloviruses as helpers (Arnold H.P. *et al*., 1999). A full description of pSSVx will be given in a specific paragraph, since the development of the present work was based on this genetic element. The plasmids belonging to pRN family share a large *open reading frame* (2676 bp in pSSVx to 2970 bp in pRN2) encoding a protein with homology to initiator proteins of replication (Rep), an *orf* (CopG homologue) located directly 5' of the *rep* gene, encoding a protein which may control the copy number of the plasmids and the homologous and most highly conserved *orfs* 80, 81, 76 and 73 in pRN1, pRN2, pSSVx and pDL10 respectively. These conserved ORFs encode a DNA-binding protein, PlrA, which probably has a DNA-binding property and is the most highly conserved of the *Sulfolobus* plasmid-encoded proteins (Garrett *et al*., 2004). Lipps *et al*. 2001 demonstrated that the ORF 80 protein of pRN1 represents a novel type of basic leucine-zipper protein.

An integrative non-conjugative extrachromosomal genetic element, denoted as pSSVi, has been isolated from a *S.solfataricus* P2 strain and it was characterized (Wang Y. *et al*., 2007). This genetic element is a double-stranded DNA of 5740 bp in size and contains eight ORFs. It resembles members of the pRN plasmid family in genome organization but shows only weak similarity to the latter in conserved regions. pSSVi has a *copG* gene similar to that of a pRN plasmid, encodes a large replication protein which, unlike a typical pRN RepA, contains no polymerase/primase domain, and lacks the *plrA* gene. Interestingly, pSSVi encodes an SSV-type integrase which probably catalyzes the integration of its genome into a specific site (an Arg-tRNA gene) in the *S. solfataricus* P2 genome. Like pSSVx, pSSVi can be packaged into a spindle-like viral particle and spread with the help of SSV1 or SSV2. pSSVi genome is stably integrated into the host chromosome and it is excised in presence of SSV2 virus. In addition, both SSV1 and SSV2 appeared to replicate more efficiently in the presence of pSSVi. Given the versatile genetic abilities, pSSVi appears to be well suited for a role in horizontal gene transfer.

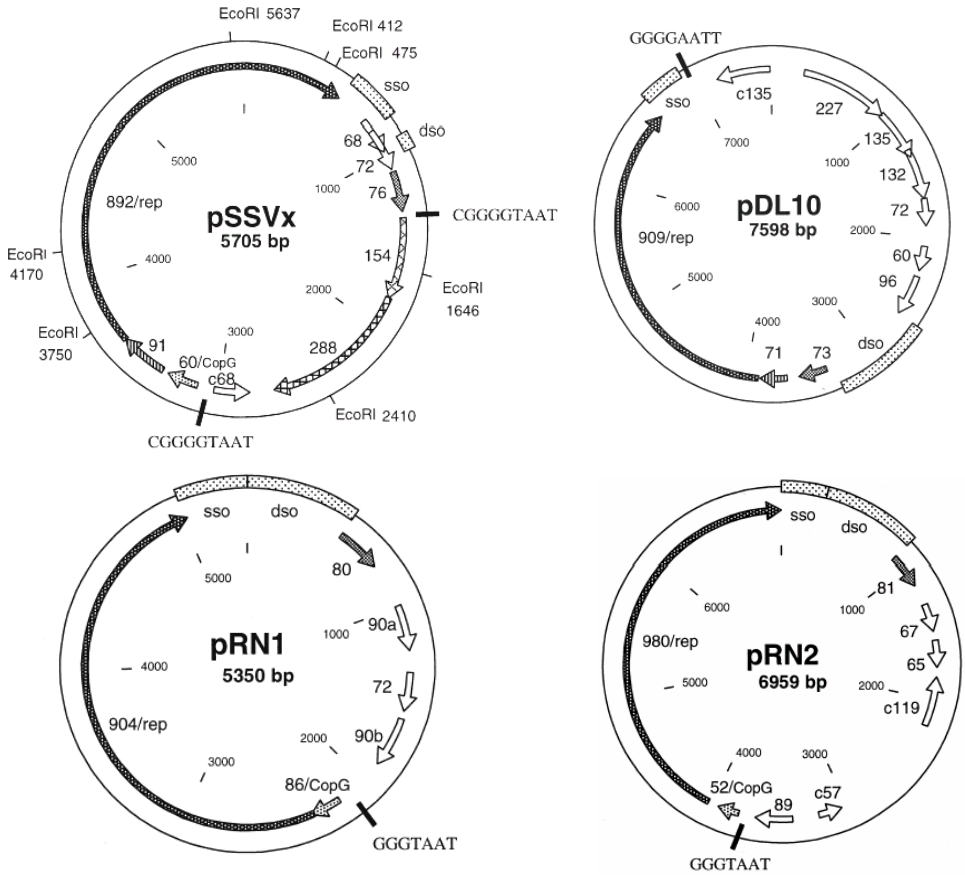

Fig. 4 Maps of the genomes of the plasmids pSSVx, pRN2, pRN1 and pDL10 forming the pRN family. ORFs were named by the number of codons. "c" (complementary) indicates opposite orientation. ORFs in the "minimal replicon" (conserved region) are shown as filled arrows; sequence elements in this region as filled boxes. Homologous ORFs are indicated by the same fill pattern. Consensus motifs located between conserved and variable regions are indicated.

Viruses

Evidence supporting an extremely high abundance of viruses in the biosphere has stimulated a strong interest in determining the degree of viral diversity in different ecosystems. Viruses and virus-like particles have been isolated from high-temperature (≥80°C) acidic pH (<3.0) terrestrial environments throughout the world (Rachel *et al*., 2002; Rice *et al*., 2001).

Typically, these environments are found in thermal features such as boiling hot springs, mud pots, fumaroles, and geysers, with a wide range of geochemical compositions. One group of ecosystems which has yielded exciting new results is geothermally heated hot aquatic environments. Screening for viruses in terrestrial hot springs with temperatures above 80°C in Iceland (Zillig *et al.*, 1994) and in Yellowstone National Park (Rice *et al*. 2001) has revealed numerous, different virus morphotypes, many of which have not been previously observed in nature. In order to confirm that hot aquatic environments constitute favorable habitats for viruses with unusual morphotypes, Haring *et al.* 2005a have examined viral diversity from a hot, acid spring (87 to 93°C, pH 1.5) in Pozzuoli, a volcanic area near Naples (Italy), and the results reinforced previous observations. Virus-like particles with five different

morphotypes were isolated from the environmental samples, and they were shown to be infectious virions.
Several of the isolated viruses infect hyperthermophilic *Archaea* from the *Crenarchaeota* kingdom, and they all have double-stranded DNA (dsDNA) genomes.
Very few viruses have been identified from *Archaea* as compared to *Bacteria* and *Eukarya*. The isolation and characterization of viruses often lead to new insights into virus relationships and to a more detailed understanding of the biochemical environment of their host cells.
From a general evolutionary standpoint, viruses of the third Domain of life, *Archaea*, are of major interest. Most cellular characteristics of *Archaea* resemble those of *Bacteria* and, accordingly, one might expect that archaeal viruses would resemble bacteriophages. Moreover, given the existence of a shared gene pool of tailed phages (Casjens, 2005; Pedulla *et al*., 2003) and also the well-established and extensive horizontal gene transfer (HGT) between *Bacteria* and *Archaea*, the archaeal viruses would share many genes with bacteriophages and, effectively, would represent distinct phage varieties. However, the information-processing systems of *Archaea* are distinct from the bacterial counterparts but monophyletic with the eukaryotic ones (Brown and Doolittle, 1997), which suggests the intriguing possibility that *Archaea* might harbour viruses related to eukaryotic viruses, at least with respect to the replication machinery.
Most of the viruses isolated thus far replicate in hyperthermophilic hosts that belong to the *Sulfolobales* family. In terms of virion morphology, the known viruses of *Archaea* fall into two distinct classes: viruses of mesophilic and moderately thermophilic *Euryarchaeota* closely resemble head-and-tail bacteriophages whereas viruses of hyperthermophilic *Crenarchaeota* show a variety of unique morphotypes. Accordingly to this distinction, the sequenced genomes of euryarchaeal viruses encode many proteins homologous to bacteriophage capsid proteins. In contrast, initial analysis of the crenarchaeal viral genomes revealed no relationships with bacteriophages and, generally, very few proteins with detectable homologs.
Prangishvili D. *et al*. (2006a) have described a re-analysis of the proteins encoded by archaeal viruses, with emphasis on comparative genomics of the unique viruses of *Crenarchaeota*. Detailed examination of conserved domains and motifs uncovered a significant number of previously unnoticed homologous relationships among the proteins of crenarchaeal viruses and between viral proteins and those from cellular life forms and allowed functional predictions for some of these conserved genes. The comparative-genomic analysis described by Prangishvili D. *et al*. (2006a), along with previous studies, revealed a peculiar world of crenarchaeal viruses. These viruses share a small pool of genes among themselves and more genes with their hosts. Since a conserved core of orthologous genes is shared by distinct viruses (like the SIRV/ARV-AFV-SIFV cluster), it seems to be justified the existence of a common ancestry of viruses themselves and their morphotypic diversity reflects their fast evolution (**Fig. 5**). Some crenarchaeal viruses like PSV and TTSV1 share no orthologous genes with other viruses; their genomes are almost a *terra incognita* in which even sequence analysis pushed to the limit, as described here, revealed only a few genes with predictable functions but uncertain provenance.

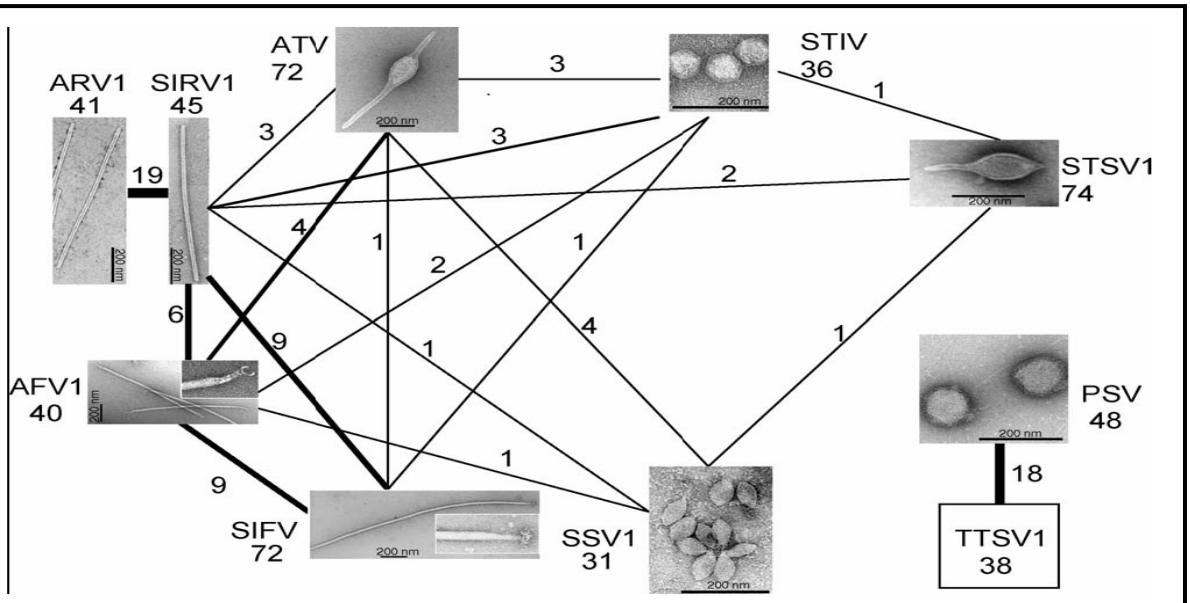

Fig. 5 Crenarchaeal viruses and orthologous relationships among their genes. The orthologous relationships between viral gene sets are depicted in the form of a graph, with electron-microscopic images of the respective virions in the vertices (the bars are 200 nm; the image of TTSV1 was unavailable). The numbers at each edge show the number of inferred orthologous genes between the respective viruses. The thickness of the lines is roughly proportional to the number of orthologs.

Most known dsDNA viruses that replicate in *Crenarchaeota* have morphotypes that have not previously been observed among dsDNA virus of *Euryarchaeota*, *Bacteria* or *Eukarya* (Haring M. *et al*., 2005a; Rachel R. *et al*., 2002; Rice G. *et al*., 2001) although there are exceptions including the spherical PSV (Haring M. *et al*., 2004) and the larger, also spherical STIV (Rice G. *et al*., 2004). Moreover some crenarchaeal viral morphotypes, including the rod-like structures of the rudiviruses, have been observed among eukaryal single-stranded (ss) RNA viruses. Crenarchaeal viruses so far have been classified into eight families: *Ampullaviridae*, *Bicaudaviridae, Guttaviridae, Rudiviridae, Globuloviridae, Paramyxoviridae, Lipothrixviridae* and *Fuselloviridae.* This classification was primarily made on the basis of their unusual or unique morphotypes, and then reinforced by the genomic properties of the viruses (Prangishvili D. *et al*., 2006). The crenarchaeal viruses that have unique virion structures are the bottle-shaped virions of *Ampullaviridae* (Haring *et al*., 2005a), the two-tailed virion of the *Bicaudaviridae* and the droplet-shaped virions of *Guttaviridae*. The virus SNDV has been found in a novel *Sulfolobus* isolate from New Zealand, STH3/3. It differs in many features from all other *Sulfolobus* viruses. In view of its singularity, a new family has, however, not been created. To list it, Zillig *et al*. (1996) proposed to assign it to a “floating genus” *Guttavirus,* because it has the shape of a droplet with a dense beard of thin filaments on its pointed end (**Fig. 6, panel B**) (Zillig W. *et al*., 1996). Its cccDNA is 20 Kbp in size and is extensively modified in an unknown way so that many restriction enzymes cleave it either partially or not at all. The chromosomal DNA of the host is not modified. SNDV can be grown in a laboratory host, STH3/3 from New Zealand, but a plaque test has so far not been developed. Electron photomicrographs indicate that the coat has a helical structure. Almost all of the crenarchaeal viruses are enveloped, in contrast to the large majority of the characterized bacterial and euryarchaeal viruses. Exceptions

are members of the *Rudiviridae* whose virions consists of double stranded DNA complexed with multiple copies of a single 15 KDa DNA-binding protein which is reminiscent of the structure of the ssRNA-containing rod-shaped virions of the plant virions in the family *Tobomaviridae. Rudiviridae* family is represented by two very similar viruses named SIRV1 and SIRV2 (**Fig. 6**, **panel A**) (Arnold H.P. *et al*., 1997). Interestingly, the spherical virion of the *Globuloviridae* consists of an envelope encasing a helical nucleoprotein core that has a structure similar to that of the ribonucleoprotein of another family of eukaryal ssRNA viruses, the *Paramyxoviridae* (Haring M. *et al*., 2004). *S.islandicus* filamentous virus (SIFV) is a member of the family *Lipothrixviridae;* their virions are normally flexible filaments, ranging in length from 0,4 to more than 2μm and in width from 20 to 40 nm, and contain linear double-stranded DNA (**Fig. 6, panel C**). Both termini of SIFV DNA are blocked against digestion by 5'- and 3'-specific exonucleases . The linear dsDNA genome of SIFV is estimated to be 42 kb in length with ≈74 ORFs (Arnold H.P. *et al*., 2000). All lipothrixviruses have envelopes. In TTV1, isolated from *Thermoproteus tenax,* and SIFV1 these envelopes contain hydrophobic proteins and host lipids, the latter in composition differing from those in the host membranes.

Other two lipothrixviruses have been isolated from *Acidianus, Acidianus* filamentous virus 1 (AFV1) (Bettstetter M. *et al*., 2003) and *Acidianus* filamentous virus 2 (AFV2) (Haring M. *et al*., 2005b). AFV1 is an enveloped 900 × 24-nm filamentous virus with claw-like structures at each end; AFV2 is a flexible filament 1100× 24-nm with terminal collar-like structure with two sets of inserted filaments. Members of the family *Fuselloviridae* are characterized by their 60 × 90-nm spindle-shaped virions that possess an ≈15.5-kb circular double-stranded DNA (dsDNA) genome with ≈35 ORFs. A distinct feature of the known crenarchaeal viruses is their stable relationship with the host; the true lysogeny of *Sulfolobus* harboring *Fuselloviruses* (SSVs) resembles the lysogeny of *Bacteria*. Crenarchaeal viruses with circular genomes from the families *Fuselloviridae* and *Bicaudiviridae*, and the unclassified STSV1 all encode a single integrase gene. The fuselloviruses and the bicaudavirus ATV have been shown to integrate into host chromosomes, whereas STSV1 has not been detected in an integrated state (Haring M. *et al*., 2005c; Schleper C. *et al*., 1992; Xiang X. *et al*., 2005). So far, only two crenarchaeal viruses, TTV1 and ATV, have been shown to lyse their host cells (Haring M. *et al*., 2005c; Janekovic D. *et al*., 1983). Most of the known crenarchaeal viruses, like *Lipothrixviridae* and *Rudiviridae*, carry linear genomes, which persist at a low copy number in the cell and do not integrate into the host chromosome. The host maintains the virus stable while multiplying, suggesting that an equilibrium exists between cell division and virus replication. Moreover, for some crenarcheal viruses like the filamentous virus AFV2, it has been demonstrated that this stability remain unaffected by stress factors, such as UV-irradiation or treatment with mytomicin C (Haring M. *et al*., 2005b).

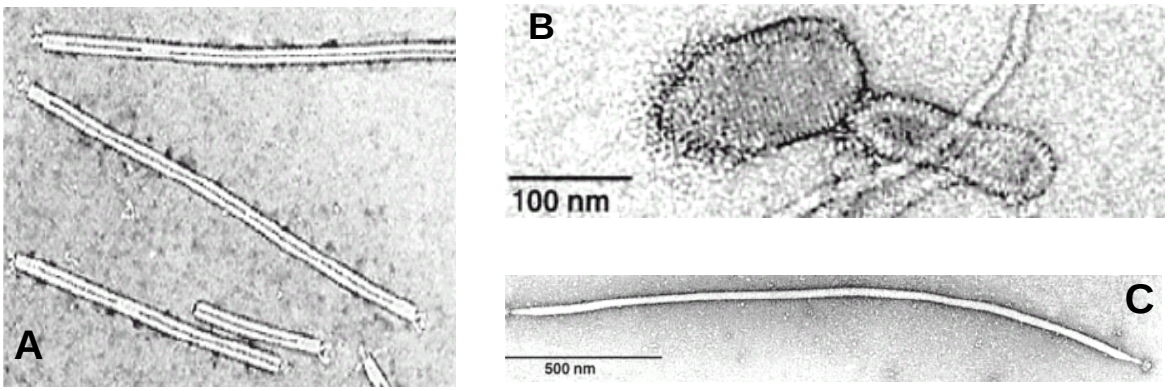

Fig. 6 Electron photomicrographs of a purified preparation of the rudivirus SIRV2 (panel A), the guttavirus SNDV (panel B) and the lipothrixvirus SIFV (panel C).

Rice G. *et al.* (2004) have reported the structural and genetic characterization of an icosahedral virus, named STIV (**Fig. 7**), from a hyperthermo-acidophilic archaeon isolated from a hot spring in Yellowstone National Park (YNP); phylogenetic analysis of the 16S rRNA gene identified the viral host as a close relative of *S. solfataricus*. Icosahedral non-tailed double-stranded DNA (dsDNA) viruses are present in all three domains of life, leading to speculation about a common viral ancestor that predates the divergence of *Eukarya*, *Bacteria*, and *Archaea*. This suggestion is supported by the shared general architecture of this group of viruses and the common fold of their major capsid protein. The previously presented structural models in conjunction with the protein, lipid, and carbohydrate information reported by Rice G. *et al*. (2004) reveal that STIV is strikingly similar to viruses associated with the *Bacteria* and *Eukarya* domains of life, further strengthening the hypothesis for a common ancestor of this group of dsDNA viruses from all domains of life. The structural organization of the STIV capsid is similar to other viruses with large T numbers such as human adenovirus, bacteriophage PRD1, and the algal virus PBCV-1. The predicted secondary structure of the STIV major capsid protein is predominately β-sheet. This is similar to the secondary structure of the major capsid proteins of adenovirus, PRD1, and PBCV-1 determined by x-ray crystallography. Like PRD1 and PBCV-1, STIV may contain an internal lipid envelope.

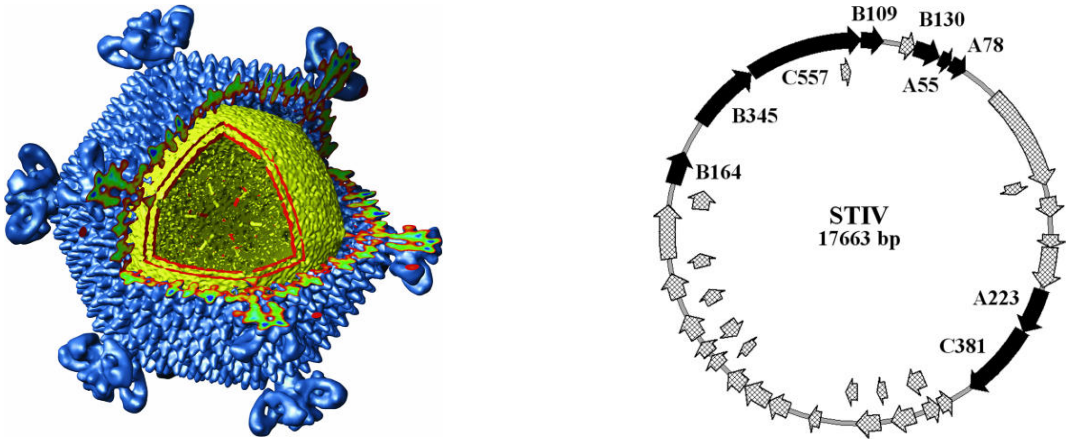

Fig. 7 A model of an STIV particle and the genome map .
On the left the protein shell is blue and has been removed from one quarter of the particle to reveal the inner lipid layers (yellow). The capsid diameter is ~70 nm, and the turrets extend 13 nm above the surface.
The location of capsid proteins on the STIV genome map is represented on the right. ORFs are named according to frame (A to F) and number of predicted amino acids. Proteins identified by protease mass mapping (solid black arrows), in general, cluster together. There are eight proteins from annotated ORFs and one additional protein, A78. This protein lacks a start methionine and was not included in the original annotation.

Like most viruses described to date that replicate in the *Crenarchaeota,* there is not significant similarity between the potential gene products encoded on the viral genome of STIV with those available in public databases. A similar lack of homology to sequences in public databases has also been observed in the genome analysis in the sequenced viral genomes of the other four families of viruses that also replicate in *Sulfolobales* hosts. This suggests that viruses of the *Sulfolobales* may have unique processes for carrying out their replication cycles and/or may reflect unique requirements for replication in hyperthermophilic archaeal hosts. Perhaps the ORF conserved in all except the *Fuselloviruses* is involved in this replication or adaptation to extreme conditions. It is also possible that these viruses can undergo rapid evolution, or that they have been ecologically isolated for a long period due to the unique environments in which they inhabit, which has obscured their relationships at the gene level to other viruses and organisms. It seems probable that a combination of these factors contribute to the genetic diversity of these viruses, but the relative ecological isolation of the *Sulfolobales* and their viruses may be the most important factor that has resulted in their unique genome composition.

Fuselloviridae family

To date, the *Fuselloviridae* are the most widespread on earth in the *Sulfolobus* genus with viruses sharing similar morphology as well as DNA genome size and organization. The first spindle-shaped virus was discovered in *S.shibatae B 12,* a crenarchaeon isolated from a sulphurous hot spring in Beppu, Japan (Grogan D. *et al*., 1990; Martin A. *et al*., 1984; Yeats S. *et al*. 1982). The virus was named SSV1 and has been characterized in greater detail. SSV1 virus production can be induced in both the natural and foreign hosts by UV irradiation or treatment with mitomycin C (Reiter W.D. *et al*., 1987; Schleper C. *et al*., 1992). All transcripts in the viral genome have been mapped and their promoters and terminators have been determined (Palm *et al*., 1991; Reiter *et al*., 1987). The fusellovirus SSV1 is the best studied member of this family and demonstrated to be temperate both in *S. shibatae* and in non-natural but related *Sulfolobus* hosts, such as *S. solfataricus* (Martin A. *et al*., 1984; Schleper C. *et al*., 1992); infection, integration of DNA into the host chromosome and production of virions cause apparently no phenotype change but a significant growth retardation of the host cells which can be visualized as turbid plaques around propagation foci on plated lawns of indicator host cells (She *et al*., 2002). The SSV1 genome is stable in three forms: (1) the viral genome packaged in virus particles which is positively super-coiled; (2) the episomal genome in the *Sulfolobus* cells which includes positively and negatively super-coiled and relaxed double-stranded DNAs (Prangishvili D. *et al*., 2001; Snyder J.C. *et al*., 2003); (3) a provirus is present in the chromosomes of both the natural host, *S.shibatae* B12, and the foreign host, *S.solfataricus* P1 (Reiter W.D. and Palm P. 1990; Schleper C. *et al*., 1992). The functions of four SSV1 open reading frames (ORFs) have so far been identified based on functional analysis. The ORF codings of the three structural proteins VP1, VP2 and VP3 have been identified by sequencing the purified coat proteins (Reiter *et al*., 1987). The fourth ORF encodes a tyrosine integrase (Esposito and Scocca 1997; Nunes-Duby *et al*., 1998) of the SSV-type, which facilitates the recombination between the virus and archaeal host attachment site, *attP* and *attA* whereas the remaining SSV ORFs show little or no significant sequence similarity to any sequence with a known function in public databases. A few more SSV viruses have been isolated from different locations in the world, forming a novel *Fuselloviridae* virus family (Arnold H.P. *et al*., 1999a,b). These include: SSV2 carried by *S.islandicus* REY15/4, which was originally isolated from a solfataric hot spring in Reykjanes, Iceland (Arnold H.P. *et al*., 1999a,b; Stedman K.M *et al*., 2003); SSV RH, which was obtained from an enrichment culture sampled in the Ragged Hills region of the Norris Geyser basin in Yellowstone National Park, United States (Snyder *et al.*, 2003; Wiedenheft *et al.*, 2004) and SSV K1, which was isolated in the Geyer Valley region of the Uzhno-Kamchatsky National Park in the Kamchatka peninsula, Russia (Snyder *et al.*, 2003; Wiedenheft *et al.*, 2004). The fusellovirus SSV2 was isolated, characterized and its complete genomic sequence determined. SSV2 virus shares with SSV1 similar morphology, replication and DNA size (Stedman K.M. *et al*., 2003). The overall genome architecture is conserved but the low similarity in the sequences should be responsible for the higher copy number and the lack of a strong ultraviolet induction of episomal SSV2 DNA and particle production, as well as the different integration of the SSV2 genome which occurs into the host chromosome at the site of a glycyl-tRNA instead of arginyl-tRNA (Wiedenheft B. *et al*., 2004).

Sequence and orientation similarity between SSV1 and SSV2 has been observed in the ORFs of one half of the genome, but not in those of the other half (Stedman K.M *et al*., 2003; Wiedenheft B. *et al*., 2004). A comparative genomics analysis performed by Prangishvili *et al*. (2006a) as previously described with a selected set of archaeal viruses, revealed that SSV1 encodes ORFs orthologous to those present in the genomes of *Acidianus* ATV (Prangishvili D. *et al*., 2006b) and AFV 1 (Bettstetter M. *et al*., 2003) as well as *Sulfolobus* SIRV 1 (Blum *et al*., 2001; Peng *et al*., 2001) and STSV1 (Xiang *et al*., 2005) viruses, all of which belong to crenarchaeal viruses.

SSV2-pSSVx genetic system

In the SSV2-pSSVx genetic system, harboured by *the S.islandicus* strain REY 15/4, SSV2 acts as an ordinary virus and a helper virus to pSSVx, while pSSVx is a hybrid between a plasmid and a fusellovirus. In fact, pSSVx genetic element is able to be maintained in non integrative form and is assumed to generate virus particles thanks to the packaging mechanisms of SSV2 (Arnold H.P. *et al*., 1999a,b; Stedman K.M. *et al*., 2003). pSSVx is also able of spreading in the cell cultures of *S. solfataricus* but only in the presence of either SSV2 or the SSV1, necessary as helpers (Arnold H.P. *et al*., 1999a). Therefore virus preparations from REY15/4 cultures appear more heterogeneous than those of SSV1, containing normal size virions and identically shaped but smaller particles (**Fig. 8**). The size difference between SSV2 and pSSVx suggests that the particle size is determined by the size of the DNA, either directly or by packaging the smaller DNA into a conditionally available subclass of smaller capsids as in the case of the “defective” bacteriophage P4 (reviewed by Lindquist *et al*., 1993). The composition of the capsids appear to be the same as those of SSV2 virions. The absence of genes in pSSVx that encode the major structural proteins of the virus demonstrates that pSSVx is a hybrid in which the plasmid DNA is packaged into a virus-like particle and can thus spread like the helper virus itself. The relationship of pSSVx to its helper fusellovirus SSV2 resembles that of the parasitic “defective” bacteriophage P4 to its compulsory helper, bacteriophage P2, especially in the packaging of these elements into helper-virus-like but smaller particles assembled from structural components of the helper. However, whereas pSSVx appears to use the structural genes of SSV2 only for spreading, bacteriophage P4 modifies the lifestyle of its helper (reviewed by Lindquist *et al*., 1993). pSSVx can be considered as a hybrid between a cryptic plasmid of the pRN family and a virus, with properties of both but lacking autonomous infectivity. Packaging of other plasmids into virus-like particles of generally transducing phages occur via homologous recombination, leading to integration into the virus genome, or by the mispackaging of products of the plasmidic rolling circle replication. An example of this mechanisms of plasmid-packaging is given by the transduction of the low-copy number plasmid pBR322 by the transducing bacteriophage P22 of *Salmonella typhimurium* (Mann and Slauch, 1997).

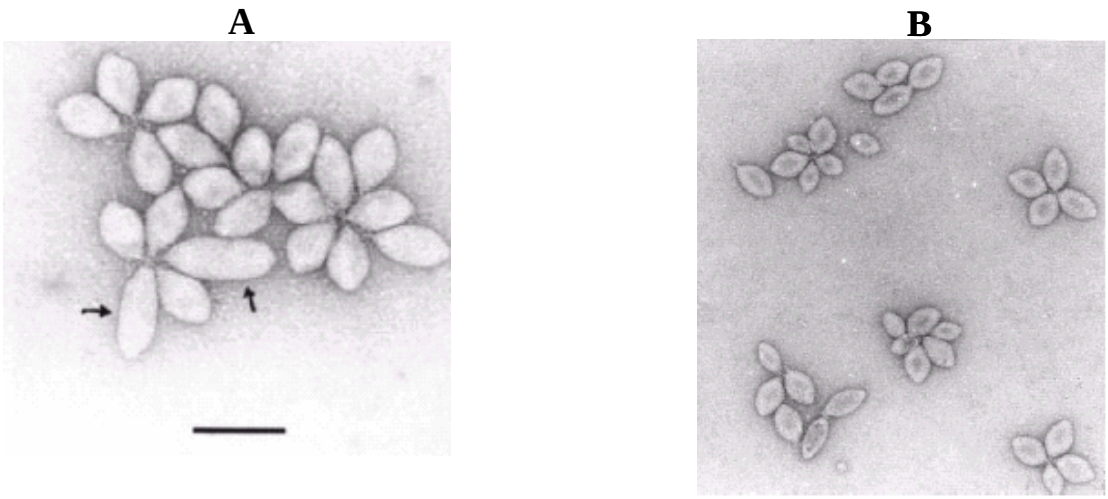

Fig. 8 Electron micrographs showing rosettes of large particles of SSV2 from a culture of *S. solfataricus* strain P1 infected with a single plaque-cloned SSV2 (**panel A**, the bar represents 100nm); the arrows indicate aberrant giant particles (Martin A. *et al*., 1984). **Panel B** shows mixtures of large SSV2 and small pSSVx particles, mostly in rosettes, from cultures of strain REY 15/4. All are negatively stained.

By sequence similarity, the pSSVx plasmid has been assigned to the pRN family (Kletzin *et al*., 1999; Peng *et al*., 2000) of *Sulfolobales* plasmids sharing a highly conserved region previously described, constituted by three ORFs. Two short ORFs, 68 and 72, overlap and are located within sequences that might be origins of replication and are, therefore, not expressed (**Fig. 4**). With the exception of c68, all ORFs of pSSVx are oriented in the same direction. ORF 60 of pSSVx had high sequence similarity to ORF52 of pRN2, and less but significant similarity to a CopG homologue in the *Staphylococcus aureus* plasmid pE194 (Byeon W.H. and Weisblum B., 1990) and to ORF56 of pRN1. All these ORFs show significant similarity to CopG proteins, which control the copy number of bacterial plasmids. Like ORF52 of pRN2, ORF56 of pRN1 and ORF71 of pDL10, the ORFs 60 and 91 of pSSVx are situated immediately upstream of a highly conserved large ORF encoding about 900 amino acid residues (**Fig. 4**). The C-terminal portions of the large ORFs exhibit significant sequence similarity to viral helicases or helicase domains in replication proteins. Lipps *et al*. 2003 have demonstrated that the recombinant N-terminal domain of the large ORF 904 not only harboured a DNA-dependent ATPase activity, but also a DNA polymerase and primase activity although there was no detectable sequence similarity to known DNA polymerases or primases. As already mentioned, differently from the members of the pRN family that are unable to spread even in the presence of helper viruses, pSSVx contains two open reading frames, named 154 and 288, showing high sequence similarity to a tandem of ORFs in both SSV1 and SSV2 genomes. It has been suggested that the proteins encoded by these ORFs can specifically recognize pSSVx DNA but associate with viral helper components necessary for capsid formation and packaging. In pSSVx, the region which shows no sequence similarity with other pRN plasmids contains three open reading frames, two of which, named 254 and 288, are juxtapositioned and show high sequence similarity to a tandem of ORFs in fusellovirus genomes. Neither pRN1 nor pRN2, which lack this tandem, spread in the presence of the fuselloviruses, which implies

that the sequences of these ORFs enable pSSVx to use the packaging system of the viral helpers for spreading. A trascriptional analysis of the pSSVx genetic element was developed by Contursi *et al*. (2007): a combination of Northern blot, primer extension and RT-PCR experiments, revealed the presence of nine major transcripts whose expression was differentially and temporally regulated over the growth cycle of *S. islandicus*.

Archaeal Transcription

Transcription, a core gene expression process, involves different molecular components participating in initiation, elongation, termination and regulation. The basic principles of transcriptional regulation in *Bacteria* and *Eukarya* have been outlined, contrasting with the *Archaea*, where such mechanisms are less well-understood. The study of the archaeal transcriptional machinery is crucial to understand the evolution of the transcription machinery in the three domains of life. Furthermore, these analyses may reveal how *Archaea* respond to environmental challenges, in particular, given their possible association with various aspects of human disease. In fact, *Archaea* share some characteristics with known pathogens that may reflect the potential to cause disease. Such characteristics include ample access to a host and capabilities for long-term colonization and coexistence with endogenous flora in a host. The detection of anaerobic *Archaea* in the human colonic, vaginal and oral microbial flora demonstrates their ability to colonize the human host. (Eckburg *et al*. 2003).
Several previous studies have shown that the RNA polymerase core enzyme exhibits structural similarity between the *Archaea* and the *Eukarya*. Moreover, the minimal set of factors required for *in vitro* transcription initiation in *Archaea* consists of TATA-box binding protein (TBP), TFIIB and RNA polymerase II (Werner and Weinzierl 2002). In the current view, the transcription initiation complex in *Archaea* is a simple, ancient version of the eukaryal one and probably constitutes the starting point for the evolution of the actual, complex eukaryotic transcription machinery.
Evidence from sequence similarity studies between RNA polymerases suggests that archaeal transcription shared certain components with that of the *Eukarya*, a conclusion further supported by the discovery of TFIIB in *Pyrococcus furiosus* and TBP in *P. woesei*. The strong similarity of archaeal transcription initiation factors with their eukaryotic counterparts has reinforced the concept that these domains shared an ancestral transcriptional apparatus. However, a rigorous comparison with *Bacteria* was not performed until entire archaeal genomes became available. It was then shown that *Archaea* contain a significant proportion of bacterial transcriptional regulators, besides their eukaryotic-like transcription initiation factors (Kyrpides and Ouzounis 1999), like Lrp and sigma -70. From the emerging sequences of archaeal genomes, a large number of open reading frames (ORFs) potentially coding for bacterium-type transcription regulation factors has been identified. This finding suggests that bacterium-and eukaryote-type elements cooperate in the regulation of gene expression in *Archaea*. Given the chimeric nature of the archaeal transcription apparatus, neither the bacterial nor the eukaryotic model can be simply applied to the regulation of gene expression in these organisms. Regulatory circuits need to be elucidated to explain the interplay among eukaryote- and bacterium-type elements.

Archaeal promoters

Archaeal promoter elements are different from the bacterial paradigm. Instead, they share some similarities with eukaryal RNA polymerase II promoters. A variety of studies have underscored that the TATA box is the major basal promoter element throughout the *Archaea*, and that it is important for transcription initiation at stable RNA (rRNA and tRNA) genes as well as at protein-encoding genes. Single nucleotide changes can result in drastic reduction in transcript levels both *in vitro* and *in vivo*. Three elements were detected that are shared by all archaeal groups: (1) an initiator element (INR) around the transcription initiation site; (2) the TATA box centred around -26/ -27,representing a highly conserved 8 bp sequence element (TTTAWAtr, with W=A/T, R=A/G) which binds TBP; and (3) an element located upstream and adjacent to the TATA box that binds TFB which is designated "*transcription factor B recognition element*" (BRE). BRE consists of a 6 bp purine-rich region (RNWAAW R= purine; W= A/T, N= any base), and determines the promoter strength and the orientation of the transcription apparatus.
The consensus sequences of TATA boxes, BREs and INR elements differ for subgroups of *Archaea*; this indicates that the molecular details of, for example, TBP-DNA interactions are variable, although the TBPs are conserved throughout the *Archaea* and the overall binding to DNA will be the same. Furthermore, the variability in BREs, INRs and additional sequence elements obviously led to the prediction that the importance and even the number of basal transcription factors will not be the same throughout the *Archaea*.

Transcriptional regulators

As discussed above, the archaeal transcriptional initiation machinery closely resembles core elements of the eukaryal polymerase II system. The basal transcription machinery of *Archaea* consists of a multi-subunit RNA polymerase, the TATA box-binding protein TBP and the transcription factor TFB (a homologue of the eukaryal TFIIB). With the ongoing genome sequencing projects many transcription regulators could be identified in archaeal genomes. Surprisingly, many of them were homologs to the members of the bacterial Lrp-like regulator family *(leucine responsive regulatory protein*), a family of global and specific regulators with members in both *Bacteria* and *Archaea* but not in *Eukarya* (Brinkman *et al* ., 2003). How regulation of an eukaryotic-like system could occur using bacterial like regulators remains an intriguing question, mainly from an evolutionary point of view.
Lrp is the prototype, and most extensively studied member, of the Lrp/AsnC family in both Gram-positive and Gram-negative *Bacteria*. Lrp is a homodimer containing two identical subunits of 18 KDa, and it acts either as an activator or as a repressor on a number of different genes and operons. Besides its role as a specific transcriptional regulator, Lrp also acts as a chromosomal organizer, inducing conformational changes in DNA and promoting the formation of higher-order DNA-protein complexes. Lrp-like regulators consist of an N-terminal DNA-binding domain of the *helix–turn–helix* (HTH) type that is connected with a hinge to the C-terminal RAM domain (*regulation of amino acid metabolism*) responsible for cofactor binding and oligomerization (Ettema *et al*., 2002). The structure of LrpA, from the hyperthermophilic euryarchaeon *Pyrococcus furiosus*, has been determined (Leonard *et al*., 2001). Almost all Lrp-like regulators bind their own promoter/operator-region,

thereby exerting negative autoregulation independently of effector molecules (Ouhammouch *et al* ., 2003). Without exception, footprinting experiments with archaeal members have revealed rather large zones of protection suggesting binding of more than one protein molecule. Different binding sites for Lrp-like proteins often exhibit rather low sequence conservation, as is the case for *Escherichia coli* Lrp, the archetype of the family. It seems that some Lrp-like proteins may not bind an ordered array of highly conserved base-specific groups, but rather interact with relatively long sequences exhibiting a specific secondary structure, as does LrpC from *Bacillus subtilis*. Binding of Lrp-like regulators often induces bending of the promoter region as was shown by hyper-reactivity for DNase I upon binding of LrpA from *P. furiosus* and Ss-Lrp from *S. solfataricus* (Brinkman *et al* ., 2000; Enoru-Eta *et al*., 2002). How do the archaeal Lrp-like regulators modulate transcription initiation? Both negative and positive effects have been reported.

Except the transcription activators Ptr2 from *Methanocaldococcus jannaschii*, which stimulate the *fdxA* and *rb2* promoters by facilitated recruitment of TBP to the TATA box (Ouhammouch *et al* ., 2003), and the homologous Lrp protein Mth from *Methanothermococcus thermolithotrophicus*, the regulators are exclusively repressors. Lrs14 from *S. solfataricus* exerts negative autoregulation by sterically inhibiting the binding of the transcription factors TATA-box binding protein (TBP) and transcription factor B (TFB) to the TATA box and BRE (transcription factor B responsive element) (Bell and Jackson, 2000). On the other hand, LrpA does not inhibit the binding of the transcription factors, but inhibits the recruitment of the RNA polymerase by physically blocking the access to the transcription start site (Brinkman *et al.,* 2000). Ss-LrpB, a novel Lrp-like DNA-binding protein from the hyperthermophilic crenarchaeon *S.solfataricus*, was shown to bind cooperatively to three regularly spaced targets in its own control region, with as consensus the 15 bp palindrome. Binding to the border sites occurred with high affinity; the target in the middle proved to be a low affinity site which is stably bound only when both flanking sites are occupied.

In the context of gene expression of the rudivirus SIRV1 in *Sulfolobus* host cells, a novel archaeal transcriptional regulator was isolated (Kessler A. *et al*., 2006). The 14 kDa protein, termed *Sulfolobus* transcription activator 1, Sta1, is encoded on the host chromosome. Its activating effect on transcription initiation from viral promoters was demonstrated in *in vitro* transcription experiments using a reconstituted host system containing the RNA polymerase, TATA-binding protein (TBP) and transcription factor B (TFB). Most pronounced activation was observed at low concentrations of either of the two transcription factors, TBP or TFB. Sta1 was able to bind viral promoters independently of any component of the host pre-initiation complex. Two binding sites were revealed by foot-printing, one located in the core promoter region and the second around 30 bp upstream of it. Comparative modelling, NMR and circular dichroism of Sta1 indicated that the protein contained a winged helix–turn–helix motif, most probably involved in DNA binding. This strategy of the archaeal virus to recruit a host cell regulator to promote transcription of its genes resembles eukaryal virus–host relationships.

The most common gene products in crenarchaeal viruses are small proteins containing the *ribbon-helix-helix* (RHH) domain. With the exception of PSV and TTSV1, all sequenced crenarchaeal viral genomes encode at least one RHH-domain protein, and STSV1 has a notable expansion of six paralogs. The RHH-domain may be considered a distinct, highly derived version of the classic helix-turn-helix (HTH) DNA-binding domain. Typically, the RHH-domain proteins are transcription

regulators, the best characterized ones being the methionine repressor MetJ (Somers and Phillips, 1992), the bacteriophage P22 Arc repressor (Cordes *et al*., 1999), and the plasmid-encoded repressor CopG (Gomis-Ruth *et al*., 1998). The RHH-domain consists of a β-strand and two α-helices and typically forms dimers, with the helices involved in dimerization and the strands, which together form a β-ribbon, recognizing the target sequences in DNA by inserting into the major groove (Gomis-Ruth *et al*., 1998). Some of the RHH domains are not easy to recognize by sequence analysis owing to their small size (50 amino acids) and limited sequence conservation; therefore identification of the full complement of these domains among viral proteins required careful inspection of the results of PSI-BLAST and CDD searches. The prevalence of the RHH domains in crenarchaeal viruses may not be particularly surprising given that small RHH-domain proteins are common in *Archaea*, being nearly as abundant as typical HTH domains (Perez-Rueda *et al*., 2004). It is notable, however, that no RHH-domain proteins were detected in the available genomes of euryarchaeal viruses, suggesting that these small and compact proteins are particularly apt for transcription regulation in hyperthermophiles. In addition to the most prevalent RHH domains, crenarchaeal viruses encode a considerable number of other predicted transcriptional regulators. The majority of viruses have at least one protein containing an HTH domain but none of these genes appear to be orthologs, suggesting multiple, independent routes of acquisition.

Two crenarchaeal viruses, STSV1 and SIFV, encode members of another class of prokaryotic transcription regulators, the SpoVT/AbrB-like proteins that have the so-called *looped hinge helix* fold, a variation of the β -barrel (Coles *et al*., 2005). The provenance of the looped-hinge helix proteins in the two viruses is obviously different. The STSV1 protein, which contains a duplication of the looped-hinge helix domain, is closely related to several predicted transcription regulators from *Sulfolobus* and is a relatively recent acquisition from the host. In contrast, the SIFV protein has no closely related homologs and is only remotely similar to several bacterial proteins of this class, leaving the origin of the viral gene unclear.

When switching from exponential growth to stationary phase, *Bacillus subtilis* may induce a number of functions designed to insure survival in a more hostile environment. The phase in which the cell decides on its response to a degrading environment is called the transition state; it is dominated by the activity of a heterogeneous class of transcription factors, called “transition-state regulators”. In fact, a large collection of genes, that are normally silent during rare periods of robust growth, is transcriptionally activated in the transition state as a consequence of metabolic and environmental changes brought about by nutritional deprivation. These genes remain silent during rapid growth, in part, because of small, global negative transcriptional regulatory factors given the name transition state regulators. In *B.subtilis*, these include the small, DNA-binding proteins AbrB, Abh, Hpr, SinR, and SpoVT. SpoVT is a late sporulation factor that modulates forespore-specific, σG-dependent transcription. AbrB and SpoVT share 65% sequence identity in their N domains, but they seem to have clearly distinct DNA binding activities, since AbrB-N cannot substitute for the homologous domain in SpoVT (Dong *et al*., 2004). In contrast to AbrB, which has a short, largely unstructured C-terminal region, SpoVT has a folded C-terminal domain of about 125 residues, which is essential for the structure and function of the N domain (Dong *et al*., 2004).

AbrB (antibiotic resistance protein B) was one of the first identified, but until recently it has remained a mysterious component in the regulatory network controlling late growth processes in *B.subtilis*. AbrB’s influence extends to a variety of genes

encoding diverse, stationary phase-induced functions. The AbrB protein is known to bind to the promoters of many of these genes, but there is little sequence similarity within their 5′ regulatory regions from which a consensus AbrB target sequence could be deduced (Zuber P., 2000). *In vitro* selection of AbrB-binding sequences revealed substrates with little or no similarity to native AbrB targets. The promiscuity of AbrB with respect to its DNA target sequences was suggested to be a function of its preference for a particular three-dimensional DNA architecture (minor groove width, degree of propeller twisting, etc.) and not a specific sequence. While bent DNA is found in some AbrB it is not a characteristic of all AbrB-binding sites nor is it a property of high affinity targets. Another determinant of AbrB affinity for DNA is the multimeric state of AbrB. Vaughn J.L. *et al.* (2000) using mass spectrometry of AbrB DNA complexes and subunit mixing experiments, revealed the tetrameric and dimeric organization of DNA-bound AbrB. Then they showed that the multimeric state of AbrB is determined, in part, by the structure of the target itself. The higher order structure of AbrB-DNA complexes can vary between one target gene and another, suggesting that AbrB undergoes a change in its tertiary and quaternary structure to conform to the site of interaction offered by the DNA target.

The structural flexibility of AbrB explains why the protein is able to specifically bind to DNA targets that have little or no sequence similarity. The AbrB protein can be divided into domains; the N-terminal domain, encompassing amino acids 1–53, is the site of DNA interaction, and the C-terminal half is required for multimerization. The N-terminal, DNA binding domain, AbrBN53, is highly conserved among Gram-positive, spore-forming *Bacteria* (*Bacillus stearothermophilus*, *B. anthracis*, *B. halodurans*, *Clostridium acetobutylicum*), and each species appears to contain more than one AbrBN53 homolog. An alignment of three AbrB-like proteins of *B. subtilis* (AbrB, Abh and SpoVT) reveals over 70% sequence identity in the DNA-binding domain, which includes arginines 23 and 24 that are crucial for DNA interaction (Strauch, M.A.,1995).

Coles *et al*. (2006) have determined the structure of AbrB, showing that it indeed forms a dimeric barrel structure. However, each monomer consists of four βstrands, with the core βαβ element elaborated by both N- and C-terminal strands (Coles *et al*., 2005). The four strands formed two β hairpins, which dimerized in a singly interdigitated way to form an eight-stranded barrel, which we named the *swapped-hairpin* barrel similar to, but topologically distinct from, double-psi barrels**.** (**Fig. 9**).

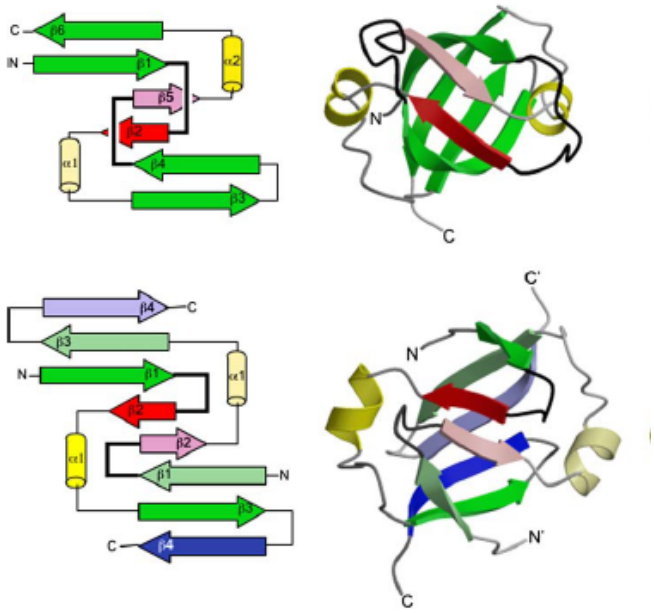

Fig. 9 Comparison of Double-Psi barrels and Swapped-Hairpin barrels
The top row shows the double-psi β barrel of VatN-N, the substrate recognition domain of the AAA ATPase VAT from the archaeon *Thermoplasma acidophilum* and the bottom row the swapped-hairpin barrel of AbrB. Common β strands are shown in green. The β2 and β2' strands, which have swapped positions in the two folds, are shown in red and pink, respectively. The two additional strands in the swapped-hairpin barrel are shown in blue.

Biotechnological features of this work

The main reason for which *S. solfataricus* has become a model *archaeon* is its ability to grow aerobically on a wide variety of heterotrophic substrates like peptides and sugars, in both rich and chemically defined media, in liquid media, as well as on plates. Moreover, *Sulfolobus* isolates from all over the world appeared to possess a wide array of extrachromosomal genetic elements (Zillig *et al*. 1998; Rice *et al*. 2001). These findings have encouraged researchers to develop tools for genetic manipulation.
Many attempts have been made to generate shuttle vectors that replicate stably in both *E.coli* and *Sulfolobus,* while maintaining a relatively high copy number in both organisms. Hybrid vectors were constructed containing the replicase gene from euryarchaeal rolling-circle plasmid pGT5, which could be transformed to *S. acidocaldarius* by electroporation, and maintained by applying butanol or benzyl alcohol tolerance selection (Aravalli and Garrett 1997). In the absence of selectable markers, the highly efficient self-spreading capabilities of conjugative plasmid pNOB8 (Schleper *et al*. 1995) and *S. shibatae* fusellovirus SSV1 (Reiter *et al*. 1987) were exploited to generate a population in which the majority of the cells contained the recombinant plasmid (Elferink *et al*. 1996; Stedman *et al*. 1999). A thermostable hygromycin phosphotransferase mutant gene enabled selection of *S.solfataricus* Gθ transformants containing plasmid pEXSs, a small shuttle vector replicating from a 1.7 kb SSV1 origin (Cannio *et al.* 1998, 2001). Most success, however, has been obtained by using β-galactosidase (*lacS*) deficient strains derived from spontaneous insertion element disruption (Worthington *et al.* 2003) or large chromosomal deletions at this locus (Bartolucci *et al*. 2003). Phenotypically, β-galactosidase complementation allows blue/white activity screening using chromogenic compound 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-β-Gal), but has in some cases allowed selection for lactose utilization as a sole carbon and energy source

(Bartolucci *et al*. 2003; Worthington *et al*. 2003). The latter property has led to a successful homologous recombination system for *S. solfataricus* strain 98-2, in which a specific gene can be disrupted by a suicide vector (Worthington *et al*. 2003). For the first time in hyperthermophilic *Crenarchaeota*, this system has enabled researchers to construct knock-out mutants and to study gene function *in vivo*. Significant progress has also been made with the construction of viral shuttle vector pMJ03. This SSV1-derived vector, which includes *lacS* for blue/white screening, also contains genes coding for orotidine-5'-monophosphate pyrophosphorylase (*pyrE*) and orotidine-5'-monophosphate decarboxylase (*pyrF*), which serve as a selectable marker complementing uracil auxotrophic *Sulfolobus* strains. This selection marker appears to be crucial for the stability of the shuttle vector; without selective pressure, rearrangement of the shuttle vector was frequently observed during prolonged cultivation (Jonuscheit *et al*. 2003).

Apart from fundamental studies like the identification of physiological functions such as DNA replication and transcription, *Sulfolobus* vectors can also be employed for biotechnological purposes such as the over-expression of endogenous genes and the consequent overproduction of thermostable enzymes. In fact, industrially relevant enzymes include extracellular proteases and polysaccharide hydrolases, such as glucoamylases, cellulases, mannanases and xylanases; however, it appeared to be difficult to produce this class of enzymes from *Sulfolobus* in common production hosts like *E. coli*, *Bacillus subtilis*, *Lactococcus lactis*, *Pichia pastoris*, and *Aspergillus niger*. In part this might be due to the fact that these enzymes are only capable to fold correctly under the specific conditions characterizing *Sulfolobus'* natural environments, like low pH and high temperatures. The availability of efficient expression system for *Sulfolobus* may represent the unique possibility to overcoming this kind of problems.

In industrial field, enzymes are increasingly being sought as alternatives to chemical catalysts. Biocatalysis, the use of microbial cells or isolated enzymes in the production of fine chemicals, is steadily moving towards becoming accepted as an indispensable tool in the inventory of modern synthetic chemistry [1 da art. Sulfo cell factory]. It is estimated that in 10 % of the cases biocatalysis will provide an overall superior synthetic strategy over traditional organic chemistry. This remarkable development in a field coined "white biotechnology" is due to the growing recognition in the industry of the capabilities and performance of enzymes as exemplified in a growing number of implemented processes, examples running at a scale of >1000 tons product/year. However when natural enzymes are recruited for industrial and biotechnological applications they are often not well suited to these tasks. Due to poor substrate solubility, breakdown of unstable products, or competing chemical reactions, the conditions for an enzyme reaction may be unsuitable for large-scale applications. Reflecting their participation in complex biochemical networks inside living cells, enzymes are often inhibited by their own substrates or products, either of which may severely limit the productivity of a biocatalytic process. Evolution is usually the culprit: enzymes are optimized and often highly specialized for specific biological functions within the context of a living organism. Biotechnology, in contrast, needs enzymes which are stable at high temperature and active over long periods of time (a feature that might clash with the need for rapid protein turnover inside a cell), enzymes which are active in non-aqueous solvents (a feature probably not required in most biological milieu), and enzymes which can accept different substrates (substrates not present in nature). Over the past 20 years, a great deal of interest has been focused on *Archaea* because of their ability to flourish in the most extreme

environments and hence provide a new generation of novel bioprocesses and microbial products of interest to industry. *Archaea* bring promising secrets from their hostile homes, and "produce" enzymes more stable than their mesophilic counterparts which can be recruited for industrial and biotechnological applications. The optimal growth temperature of thermophilic *Bacteria* generally has an upper limit around 70 °C; thus, a shift forward the more thermophilic archaeal expression system (growth up to 113 °C) would offer the greatest advantage for expanding the temperature limit for an efficient production of thermoproteins when their expression in mesophilic cells is seriously hampered. Even more interestingly, a more efficacious stabilization of mesophilic or moderately thermophilic proteins or enzymes can likely be achieved by "directed evolution" in this class of hosts. In fact, the directed evolution techniques allow the generation of enzymes with greatly enhanced characteristics and, in some examples, enzymes with new and completely novel substrate specificities. A mutated version of the hygromycin B phosphotransferase (*hph* $_{mut}$) gene from *E. coli*, isolated by directed evolution at 75 °C in transformants of a thermophilic strain of *S. solfataricus,* has been characterized with respect to its stability in both the original mesophilic and the new thermophilic hosts by Cannio R. *et al* (2001). This gene has been demonstrated to be able to express the hygromicin B resistance phenotype and to be steadily maintained and propagated also in other, more thermophilic strains of *S.solfataricus*, up to 82°C. Comparative characterization has revealed that the mutant enzyme had acquired significant thermoresistance and displayed higher thermal activity with augmented catalytic efficiency.

The aim of the work

This work was developed through different steps:

- Physiological characterization of the *S.islandicus* REY 15/4 host-SSV2 virus interaction
- Construction of the pSSVx-derived *E. coli-S. solfataricus* shuttle vector
- Structural-functional analysis of the gene product ORF c68 of the pSSVx

Host/virus interaction modes have provided windows to study microbial diversity as well as genetic processes at the molecular level, in particular for prokaryotes, and hence have helped in clarifying the physiological mechanisms, the dependence on the specific biochemical environment and evolution of their host cells. A striking finding among *Sulfolobus* virus studies has been the identification of the first archaeal helper and satellite virus system: SSV2 and pSSVx viruses in *S. islandicus* REY 15/4 which were deeply described above. The characterization of the SSV2 virus physiology was essential to clarify the helper functions of the fusellovirus respect to pSSVx and to explain the relationships between the virus replication activity and the growth rate of its natural host. On the basis of the characterization of the host-SSV2 virus interaction I could use the pSSVx, which requires the presence of the helper virus SSV2 for spreading, as the start-point to construct a shuttle vector which might be an efficient cloning vehicle for the expression of passenger genes in *S. solfataricus.*

In this study I report the findings of a comparative study of SSV2 physyology in the natural host *S.islandicus* REY 15/4, versus the foreign host *S. solfataricus,* and provide the evidence of differently regulated SSV2 life cycles in the two hosts. In fact, whereas a significant induction of SSV2 replication takes place during the growth of the natural host REY15/4, the cellular content of SSV2 DNA remains fairly low

throughout the incubation of the foreign host. The accumulation of episomal DNA in the former case cannot be traced to decreased packaging activity because of a simultaneous increase in the virus titre in the medium. In addition, the interaction between SSV2 and its natural host is characterised by the concurrence of host growth inhibition and the induction of viral DNA replication.
The different recombinant *E.coli–S.solfataricus* shuttle vectors constructed retained the wild-type capability to replicate at high copy number and to spread in cell cultures in the presence of its helper virus SSV2. *Sulfolobus* transformants were demonstrated to be stable and propagate the pSSVx derived plasmids in a reproducible and constant fashion without any rearrangement, recombination or integration into the chromosome. Moreover, stable complementation of a β-galactosidase mutant of *S.solfataricus* previously isolated and characterized in our laboratory (Bartolucci S. *et al*., 2003) and reproducible gene expression levels were also obtained by introducing the β-galactosidase gene (*lacS*) as a reporter under the control of a strong and heat-inducible promoter into the shuttle vector.
The *orf* c68 is the only reverse oriented pSSVx putative gene and has no homologs in the crenarchaeal family of pRN plasmids (Arnold H.P. *et al*., 1999a). It is highly similar (42/53% identity/similarity) to *orf* c56 of pSSVi (Whang Y. *et al* 2007) and also finds its homologues in the integrated elements of the *S.tokodaii* and *S.acidocaldarius* genome. The *orf* c68 encodes a protein containing a SpoVT/Abr-B like domain, typical of a bacterial family of transcription factors demonstrated to be repressors or activators or ambiactive, depending on the specific case.
There is still a lack of information on the specific characteristics of DNA-binding proteins from hyperthermophiles. In the present study I report the heterologous expression of ORF c68, its purification and characterization of its DNA-binding activity. The ORF c68 may be involved in regulating the gene expression of other *orfs* of the pSSVx . Further investigation on the physiological function of this protein will provide useful tools to optimize the characteristics of the shuttle vector pSSVx-derived. Moreover, the understanding of the ORF c68 functional role may contribute to clarify some informations the mechanism of the transcriptional regulation in *Archaea*.

Materials and Methods

Strains, enzymes and reagents used in this study

The reagents used for preparation of buffers and growth media of *Sulfolobus solfataricus* were supplied by Sigma Aldrich; the yeast extract and casaminoacids were from DIFCO.
The reagents for polyacrylamide gels electrophoresis were supplied by Biorad. The restriction enzymes, the "modification enzymes" (alkaline phosphatase, T4 DNA ligase, T4 DNA polymerase) and the molecular weight markers for nucleic acids were supplied by Roche. The *Taq* DNA polymerase was supplied by Promega. The *fmol DNA Cycle Sequencing System* Kit was distributed by Promega. The synthesis of the oligonucleotides was committed at MWG-Biotech; the radioactive material was supplied by Perkin Elmer. The *E.coli* pUC 19 plasmid *Sma* I digested and dephosphorilated was supplied by MBI Fermentas; the pET 30 a+ vector was supplied by Novagen.
Escherichia coli TOP F'10 strain (F'{ laclq Tn*10* (Tet^R)} *mcr*A Δ (*mrrhsd*RMS*mcr*BC) φ80Δ *lacZ* ΔM15 Δ *lacX*74 *rec*A1 *deo*R *ara*D139 Δ (*ara-leu*) 7697*gal*U *gal*K *rps*L(Str^R) *end*A1 nupG), used for cloning, were supplied by Invitrogen; *E.coli* BL21-*Codon Plus* (DE3) RIL cells (F– *ompT hsdS*(rB–mB–) *dcm+ Tet*^{r} *gal λ* (*DE3) endA The* [*argU ileY leuW Cam*^{r}*),* used for overexpression, were supplied by Stratagene.
S.islandicus REY15/4 and related *Sulfolobus* enrichment cultures were kindly provided by Wolfram Zillig (Zillig *et al.* 1994).
S.islandicus REY31A is a pSSVx-cured strain isolated from REY15/4. The plasmid curation was done by growing REY15/4 continuously for approx. 100 generation times at an optical density below 0.4 (OD_{600}). Thereafter, six single colonies were isolated via Gelrite plating and purified by re-streaking on plate for three times. Since pSSVx was hardly detected at low optical density values in REY15/4 cells, the REY31 isolates, were carefully checked for the absence of pSSVx by polymerase chain reactions and Southern analysis over all the growth phases.
S. islandicus REY15A is virus-free isolate obtained by Gelrite plating of the same enrichment culture from which the REY15/4 strain had been isolated. Six single clones (REY15A-F) were isolated and purified as single colonies as described for REY31A. They were then tested for the absence of extrachromosomal DNA and for the incapability in conferring growth inhibition to *S.solfataricus* P2 strain.
S.solfataricus P2 (DSM1617) was purchased from the German Collections of Microorganisms and Cell Cultures.
The strain *S.solfataricus* GθW, a spontaneous derivative mutant of Gθ strain lacking the β-galactodidase activity, was previously isolated in our laboratory (Bartolucci S. *et al*., 2003).

E.coli cells transformation techniques

Transformation of *E.coli* TOP F'10 cells by electroporation

2,5 ml of an over-night culture of *Escherichia coli* TOP F'10 cells were inoculated into 250 ml of LB medium. The cells were grown up to midlogarithmic phase (0,6 OD_{600}) at 37°C, stored on ice for 20-30 min and then harvested by centrifugation. The pellet was washed in 250 ml of sterile H_2O. After the second centrifugation the cells were washed in 125 ml of sterile H_2O. The third washing was performed in 5 ml 10%

glycerol; the cells were then harvested by centrifugation and the pellet was resuspended in 750 μl of 10% glycerol.
Aliquots of 10^{10} cells/ml (40 μl) were mixed with 1 μl of the DNA ligase reaction, incubated for 1 min on ice and transferred into chilled plastic cuvettes with an electrode gap of 0.2 cm (BioRad). High voltage electroporation (25 μF) was performed with a BioRad Gene Pulser Xcell™ at a field strength of 2.5 kV/cm and 200 Ohm a shock pulse was applied to competent cells producing pulse length of ~5.0-5.5 ms. Immediately after electroporation cell mixtures were diluted to 1 ml with LB medium and incubated for ~1 h at 37°C under shacking. The cells were then plated onto selective solid medium supplemented with 100 μg/ml ampicillin to isolate the recombinant clones.

Transformation of *E.coli* BL 21-*Codon Plus* (DE3) cells
2,5 ml of an over-night culture of BL21-*Codon Plus* (DE3) RIL cells were inoculated into 250 ml of LB medium. The culture was grown up to midlogarithmic phase (0,6 OD_{600}) at 37°C, stored on ice for 20-30 min and then harvested by centrifugation. The pellet was washed in 125 ml of cold 50 mM $CaCl_2$ and stored on ice for 30 min. The cells were then harvested by centrifugation and the pellet was resuspended in 16 ml of cold 50 mM $CaCl_2$.
Aliquots of 200 μl of competent cells were mixed with 50 ng of plasmidic DNA and stored on ice for ~20 min. The cells mixtures were transferred at 42 °C for 90 sec, on ice for two min and then diluted to 1 ml with LB medium. An incubation for ~1 h at 37°C under shacking was performed before plating the cells onto selective solid medium supplemented with 50 μg/ml kanamicin and 33 μg/ml chloramphenicol.

Southern blot analyses

DNA-labelling
The probes used in hybrizitation experiments were labelled by Random Primed DNA-labelling kit (Roche) using α-[^{32}P] d-ATP (Perkin Elmer). The labelled products were then purified by using a Nick ™ column (Pharmacia) and the radioactivity was measured with a Beckman scintillator.

Hybridization reactions
The pre-hybridization of the nylon membrane was performed at 65°C for 4 hours in a solution composed by 6x SSC (sodium saline citrate), 5x Denhart, 0.5% SDS and salmon sperm 100μg/ml. After the addition of the labelled probes ($1x10^6$ cpm/ml of hybridization solution), previously denaturized at 100°C for 5min, the hybridization reaction was performed over-night. The membrane was then washed as follows:

- 5 min at room temperature in 2x SSC, 0.5% SDS
- 15 min at room temperature in 2xSSC, 0.1% SDS
- 30 min-4 hours at 65°C in 0.1x SSC, 0.1% SDS

The hybridization signals were revealed by exposition of an autoradiographic sheet for 12 hours at -80°C.

Proteins analyses

Determination of the protein concentration

The concentration of the proteins in solution was determined according to the Bradford' s method (Bradford, 1976). The Coomassie Brilliant (Bio-Rad) reagent was added to the samples and the absorbance at 595 nm was monitored. A solution of 1 μg/μl of bovine serum albumin (BSA) was used as standard.

Electrophoretic analysis of proteins (SDS-PAGE)

The electrophoresis on polyacrylamide gel in the presence of SDS (SDS-PAGE) was performed according Laemmli's protocol (Laemmli, 1970). The samples were denaturized at 100°C for 20 min in 2% SDS, 5% β-mercaptoethanol, 0.001% bromophenol blue and 10 % glycerol. The samples were then loaded on a 15% polyacrylamide gel and electrophoresed in 0.025 M Tris-HCl, 0.2M glycine pH 8.3 and 0.1% SDS. A better resolution of the ORF c68 protein band was obtained by substituting the glycine with trycine in the electrophoresis buffer. The electrophoresis was performed at 25 mA for ~1.5 hours; the proteins were then revealed by Coomassie Briliant-Blue; the gel was submerged in the staining solution (0.1% Coomassie Brilliant-Blue R250, 25% isopropilic alcohol and 10% acetic acid) for ~30 min with gentle agitation. The gel was washed in a solution containing 30% ethanol and 10% acetic acid to remove the excess of Coomassie and then stored in 10% acetic acid.

Alternatively the proteins were transferred by elettro-blotting from gel to a polivinildienefluoride (PVDF) membrane as described in the paragraph "Western blot analysis".

Growth of *Sulfolobus* strains and isolation of SSV2-infected *S.solfataricus* GθW and *S. islandicus* REY15A

S.solfataricus strains P2 (DSM 1617), Gθ (Cannio R. *et al*., 2001) and the derivative mutant GθW [Δ*lacS,*] as well as *S.islandicus* REY 15/4 (Zillig W. *et al*., 1994) were grown at 75 or 80°C in TYS medium:

Ammonium sulfate	3g/l
Potassium sulfate	0,5g/l
Potassium chloride	0,1g/l
Glycine	0,7g/l
Yeast extract	0,5g/l

Brock's basal salts:

$MnCl_2 \times 4H_2O$	0,8 mg/l
$Na_2B_4O_7 \times 2H_2O$	2,1 mg/l
$ZnSO_4 \times 7H_2O$	0,11mg/l
$CuSO_4 \times 2H_2O$	25 μg/l
$Na_2MoO_4 \times 2H_2O$	15 μg/l
$VoSO_4 \times 5H_2O$	15 μg/l
$CoSO_4 \times 7H_2O$	5 μg/l
$NiSO_4 \times 6H_2O$	5 μg/l
$MgCl_2$	1 mM
$Ca(NO_3)_2$	0,3 mM
$FeSO_4 \times 7H_2O$	2 mg/l
HCl	0,1 mM

with 0.1% tryptone and 0.2% sucrose, at pH 3.2 adjusted using concentrated H_2SO_4. The shacking rate of the cultures was 150 rpm. For electroporation and plaque assays, cells were grown with phosphate buffered medium N.182 (M182) suggested by the DSMZ Catalogue of strains containing 0.1% glucose:

Ammonium sulfate	2,5 g/l
Magnesium sulfate	0,2 g/l
Potassium phosphate	3,1 g/l
Calcium chloride	0,25 g/l
Casaminoacids	1,0 g/l
Yeast extract	1,0 g/l

Basal salts:

$MnCl_2$ x $4H_2O$	1,8 mg/l
$Na_2B_4O_7$ x$10H_2O$	4,5 mg/l
$ZnSO_4$ x $7H_2O$	0,22 mg/l
$CuCl_2$ x $2H_2O$	0,05 mg/l
Na_2MoO_4 x $2H_2O$	0,03 mg/l
$VoSO_4$ x $2H_2O$	0,03 mg/l
$CoSO_4$ x $7H_2O$	0,01 mg/l

The optical density of liquid cultures was monitored at 600 nm. For solid media, gellan gum (Gelrite, Sigma) was added to a final concentration of 0.8% (0.35% for overlays), and $MgCl_2$ and $CaCl_2$ were added to 10.0 and 3.0 mM, respectively.
The *S.solfataricus* strains and *S.islandicus* REY15A (SSV2-REY15A) transfected with the SSV2 and pSSVx were obtained as following: 1-2 μl of the REY15/4 surnatant containing SSV2 and pSSVx virons were spotted onto the soft layer of a Gelrite plate seeded with *S.solfataricus* P2 or *S.islandicus* REY 15A; the plates were incubated at 80°C for 2-3 days and turbid halos (plaques) resulting from the inhibition of host growth by SSV2 appeared. Clones of GθW and REY 15A SSV2-infected cured selectively for pSSVx were obtained by isolation of single colonies formed on plates by the cells extracted from plaques and revitalized in liquid cultures.

pSSVx-SSV2 analyses

SSV plaque assay

The virus titre of a culture was determined by means of plaque assay using *S.solfataricus* P2 as an indicator strain according to Schleper *et al*. (1992). The plates were incubated for 2–3 days at 80°C and examined for the appearance of the turbid halos which the inhibition of host growth by SSV2 virus generates.

DNA isolation

The Plasmid DNAs of SSV2 and pSSVx were prepared from *Sulfolobus* cells using the QIAGen Spin Miniprep kit. Total DNAs were isolated from 10-15 ml of *Sulfolobus* cultures. Summing up, cell pellets were re-suspended in 400 μl of the lysis buffer (50 mM Tris-HCl, pH7.5, 50 mM EDTA and 0.2% N-Lauroyl Sarcosine) and incubated at

55°C in the presence of proteinase K (380 mg/ml) for two hours. To extract the samples phenol twice, phenol-chloroform once and, finally, chloroform were used. Thereafter, the DNAs were precipitated using ethanol.

Estimation of cellular SSV2 and pSSVx content during host growth

To visualize SSV2 and pSSVx DNA accumulation during the growth of host cells, total DNAs were extracted from the samples collected at different incubation stages (as indicated in the individual experiments) from the cultures of *S.islandicus* REY15/4, REY31A, and SSV2-REY15A. Total DNAs were isolated from the same amount of the cells (normalized by OD_{600}) using the procedure described above. Similar amount of DNAs (estimated by OD_{260} ad ethidium bromide fluorescence) was used for restriction digestion and Southern blot analysis. Normalization of the loaded DNAs was performed using the alba gene (*Sso*0962), or the host-SSV2 hybrid fragment derived from SSV2 integration into the host genome (one copy per genome). To determine the SSV2 and pSSVX copies per chromosome, three different sequences, i.e. the alba gene, the above mentioned 4.69-kb SSV2 (or the coding region of vp3 gene) and a *Hind*III 1.7-kb pSSVx fragments, were ^{32}P-radiolabelled by random priming, yielding to a specific activity of ca.10^9 cpm/μg DNA. The same amounts of the radiolabelled probes (1x10^6 cpm/ml for each) were mixed and used for Southern hybridization, as described in "Hybridization reactions" paragraph. Since only one copy of the alba gene (or the host-SSV2 hybrid fragment) is present in *Sulfolobus* genomes, SSV2 and pSSVx copy numbers were determined by the ratios of extrachromosomal-derived hybridization signals to chromosomal-derived ones which were quantified using a phosphoroimager and the Quantity One 4.2 software (Bio-Rad).

pSSVx-derived shuttle vector construction

pSSVx-derived plasmid constructions

Plasmid pSSVrt was constructed by cloning the pSSVx DNA linearized at the position 2812 with *Afl*III and modified with Klenow DNA polymerase into the *Sma*I site of the pUC19 *E.coli* vector. Linearization was obtained by *Afl*III partial digestion in the presence of ethidium bromide using a protocol already described for the SSV1 virus DNA linearized with *Sau*3AI (Stedman K.M. *et al*., 1999). Clones with insertion at the specific position were selected by restriction analysis of the resulting plasmid collection. A size reduced derivative of pSSVrt was also constructed; the polycloning sequence (between the *Aat*II and *Eco*RI sites) of the pUC28 vector was inserted into a 1812 bp *Aat*II/*Afl*III DNA fragment from pUC18 after suitable modification of incompatible ends. The pSSVx sequence was excised from the pSSVrt plasmid with *SacI* and *PstI* and inserted into the same sites of the minimal plasmid obtained to produce the pMSSV vector. An expression cassette of the *lacS* gene (Cubellis M.V. *et al*.,1990) was PCR amplified by the vector pMJ03 (Jonuscheit M., *et al*. 2003) and inserted between the *Xho*I and *Pst*I sites of pMSSV generating the expression vector pMSSV*lacS*. Excision of the *E.coli* minimal plasmid was obtained by digestion of pMSSV*lacS* (1.0 μg) with *Sac*I and purification of the pSSVx/*lacS* moiety from agarose gel. The DNA was re-circularized by ligation in a final volume of 5.0 μl, diluted in water to 20 ng/μl final concentration, and 2 μl were used for electroporation.

Transformation procedure and analyses of *Sulfolobus* transformants

S.solfataricus cells of SSV2 lysogenes were grown up to midlogarithmic phase (0.3–0.45 OD_{600}), harvested by centrifugation and repeatedly washed in 20 mM sucrose as described previously by Schleper *et al*. (1992). Aliquots of 10^{10} cells/ml (50 μl) were mixed with 1 μl DNA (10–100 ng/μl), incubated for 1 min on ice and transferred into chilled plastic cuvettes with an electrode gap of 0.1 cm (BioRad). High voltage electroporation (25 μF) was performed with a BioRad Gene Pulser Xcell™ at a field strength of 1.5 kV/cm and 400 W resistance; two successive shock pulses were applied to competent cells producing pulse length of ~10.0 and 9.8 ms, respectively. Immediately after electroporation cell mixtures were diluted with 1 ml of medium M182 containing 0.1% glucose, transferred to glass vials and incubated for 3 h at 75 or 80°C. After suitable scale-up, 5–15 ml aliquots of the cultures (the volumes varying in order to withdraw the same number of cells per aliquot) were harvested at increasing cell density for DNA extractions. For monitoring propagation of the recombinant satellite virus, extrachromosomal DNA mini preparations and plaque assays were performed. Transformants were stored at -80°C in 15% glycerol stocks.

For Southern blot analysis, ~2 μg of total cellular DNAs, extracted according to Arnold *et al*. (1999a), and 5 ng of pMSSV*lacS* plasmid purified from *E.coli*, were cut with *Hind*III and *Bgl*II, and electrophoresed in a 0.8% agarose gel; DNA digests were blotted and hybridized (see "Hybridization reactions" paragraph). The probe was prepared by cutting out and purifying a *Hind*III restriction fragment from the pMSSV*lacS* vector encompassing the *lacS* gene and a portion (up to -302) of the tf55α 5' flanking region. This restriction DNA fragment was randomly labelled as previously described.

β-galactosidase complementation

pMSSV*lacS* transformed cells were tested for β-galactosidase activity; 1 ml aliquots of cultures were centrifuged and cell pellets overlaid with a X-Gal solution (2 mg/ml in phosphate buffered medium) and incubated for 15 min at 75°C for blue color development. *In situ* assays were performed on the same cultures seeded on plates and grown as circular colonized areas, as previously described (Bartolucci S. *et al*., 2003). Supernatants of cultures grown up to 0.5, 1.0 and 1.3 OD_{600} were checked for plaque formation as already described and analyzed by X-Gal staining of plaques formed on continuous lawns of the *S.solfataricus* GθW strain. For infection in liquid culture, 400 μl supernatants from pMSSV*lacS* transformants were added to a 20 ml culture of cells transfected only with SSV2 and grown up to 0.5 OD_{600}. After incubation under shaking for 48 h, cells were diluted 1:50, grown up to 0.36 and 1.0 OD_{600} and tested for β-galactosidase activity as already described.

ORF c68 analyses

Bioinformatics analysis

Data banks (non-reduntant) were screened for homologues to ORF c68 of pSSVx with the program BLAST, PSI-BLAST and CDD. Sequence alignments were performed with ClustalW and the predictions of the protein secondary structures were carried out with different algorithms and programs (GOR(IV), AGADIR, JPred, PredictProtein, NNPredict) that are provided by Expasy web site.

Cloning of *orf c68* and overexpression of ORF c68

Escherichia coli TOP F'10 strain was used for cloning. The *orf c68* gene was amplified from pSSVrt plasmid by PCR reactions. The first amplification was performed by using the following primers:

Forward: 5'- ACGAACAAATTGTTTCATATGAGACCAGGCATACG -3'
Reverse: 5'-TTATCAAAAAAACACGTCTCGAGAATTTTTAATTCCTTCACCC-3'

The *forward* primer indicated above contained the *Nde*I restriction site, while the *reverse* primer was designed to substitute the STOP-translation codon of the *orf c68* gene with part of the *Xho*I restriction site. This cloning strategy was used to obtain the ORF c68 gene product fused with the 6Histidine-tag of the pET30 a+ (Novagen) expression vector (see below).
The PCR reaction necessary to obtain the untagged ORF c68 protein, was performed with these primers:

Forward: 5'- ACGAACAAATTGTTTCATATGAGACCAGGCATACG -3'
Reverse: 5'-TTATCAAAAAAACACTCGAGTTAAATTTTTAATTCCTTCACC-3'

In this case, the *forward* primer was the same used in the first reaction while the *reverse* primer was designed so that the cut of the *Xho*I restriction enzyme could be downstream the STOP-translation codon, which was so left unchanged. The *Xho* I restriction sites on the *reverse* primers are underlined and the complementary sequence of the STOP- translation codon of the *orf c68* gene is coloured in red.
The two PCR products were cut with *Nde*I and *Xho*I (Roche) and ligated into the expression vector pET30a+ cut with the same enzymes. The identity of the inserts in the resulting plasmids (pET30/6His-*orf c68* and pET30/*orf c68*) was confirmed by DNA sequencing. These constructs were transferred into *E.coli* BL21-*Codon Plus* (DE3) RIL cells. Aliquots of 10 ml of an over-night culture were inoculated into 1L of LB medium supplemented with 50 μg/ml kanamycin and 33 μg/ml chloramphenicol and grown at 37°C under shaking. When the culture reached 0,7 OD_{600} 0,5 mM IPTG was added; after 3 h the cells were collected by centrifugation and kept frozen at -20 °C until use.

Purification of ORF c68 gene products

The frozen cells from 1 L of culture were resuspended in 20 ml ice-cold lysis buffer (50 mM sodium phosphate, 100 mM NaCl, 1 mM EDTA, pH 7,0) with protease inhibitor cocktail (Roche) added. Lysis was performed by addition of 0,17 mg/ml lysozyme and incubation for 20 min at room temperature under gentle shaking. Sodium deoxycholate 1,6 mg/ml was added and the suspension incubated at room temperature for 30 min in the presence of 30 μg DNase I and centrifuged (20817 x g,

20 min, 4°C). Addition of protamine sulphate to a final concentration of 1mg/ml at 4°C, and a 20 min-centrifugation step (20817 x g at 4°C) allowed the removal of the nucleic acid component as described by Cannio *et al*., 1994.

1. In the case of the 6His-tagged ORF c68 protein, the cleared lysate was dialysed over night against 20 mM Tris-HCl pH 8,0 and then incubated at 60°C for 20 min to remove the most of the host mesophilic proteins. The sample was then centrifuged for 20 min at 20817 x g (at 4°C) and 10 mM imidazole was added to the supernatant. The lysate was loaded onto a His-trap column connected to AKTA™ chromatography system. The column was equilibrated with 50 mM Tris-HCl pH 8,0 and developed with three steps: 1) 50 mM Tris-HCl pH 8,0, 300 mM NaCl and 20 mM imidazole; 2) 50 mM Tris-HCl pH 8,0, 300 mM NaCl and 250 mM imidazole; 3) 50 mM Tris-HCl pH 8,0, 300 mM NaCl and 500 mM imidazole. The protein was purified to homogeneity during the third step of the His-trap affinity chromatography.

2. To purify the ORF c68 untagged protein the cleared lysate was dialysed over night against 20 mM sodium phosphate, pH 7,5 and then loaded onto a cationic exchange RESOURCE S column connected to AKTA™ chromatography system. The column was equilibrated with 20 mM sodium phosphate, pH 7,5 and developed with a linear gradient of 0-300 mM NaCl followed by a final step of 500 mM NaCl. ORF c68 was eluted at ~300 mM NaCl and then purified to homogeneity after a gel filtration step on a Superdex 75 PC 3.2/30 column. The gel filtration chromatography was performed in 50 mM sodium phosphate and 200 mM KCl, pH 7,5.

Determination of the molecular weight of the ORF c68 untagged protein

Gel filtration was carried out onto a Superdex 75 PC 3.2/30 column connected to AKTA™ chromatography system. An Aliquot of 100 μl of ORF c68 (5mg/ml) was loaded at 0,04 ml/min onto a column equilibrated with 50 mM sodium phosphate, 200 mM KCl, pH 7,5. ORF c68 was eluted at 1,64 ml. The molecular weight (M) was determined by interpolation of a calibration curve. This curve was prepared by measuring the elution volumes of four standards (LMW Gel Filtration Calibration Kit, Amersham Biosciences) and plotting their values versus the logarithm of their molecular weight.

Electrophoretic mobility shift assay (EMSA)

The 80 bp and 130 bp fragments (promoter and *CopG*/*c68* intergenic region respectively, **Fig. 10**) were amplified by PCR using the pSSVrt as template. The amplification of the 80 bp promoter region was performed whith these primers:

- *Forward:* 5'-GTCTCATAATAAACAATTTGTTCGT- 3'
- *Reverse*: 5'- CCACTTTTGGGTATTACCCCGTGGC- 3'

The 130 bp intergenic fragment was amplified by using the following primers:

- *Forward:* 5'- GTCTCATAATAAACAATTTGTTCGT- 3'
- *Reverse:* 5'- ACCATTTTGTCACCAGGTACAGTA- 3'

The PCR products were ligated into pGEM t-easy vector (Promega). The final constructs were digested with *Eco*RI enzyme to generate DNA fragments with 3'

recessed termini. The probes labelling was then performed by *fill-in* reaction with Klenow polymerase (Roche) and [α-^{32}]dATP (Perkin Elmer, Italy).
A thermal pre-incubation of the purified ORF c68 protein was conduced for 15 min at 50°C in assay buffer (25 mM Tris-HCl pH 8,0, 50 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 5% glycerol) in the presence of 500 ng of poly dI/dC as competitor. The labelled probes were added at a concentration of 0,030 μM (for the 50bp promoter region) and 0,015 μM (for the *CopG/c68* intergene). The binding reactions were performed with increasing amounts of ORF c68 (4-8 μM) for 15 min at 50°C in a final volume of 15 μl. The samples were loaded with 1-2 μl of loading buffer (0,1% xylene cyanol, 0,1% bromophenol blue in 50 % glycerol) onto a 6 % polyacrylamide gel in TBE 0,5 x, after a pre-run of the gel at 150 Volts for 30 min. After the electrophoresis the gel was dried and the radioactive signals were revealed by autoradiography.
In the displacement experiments the binding reactions were performed under the same conditions described above, except for the addition of increasing amounts of unlabelled probe (50 bp promoter region) with 1:1, 1:10, 1:100 and 1:1000 ratio of labelled /unlabelled DNA.

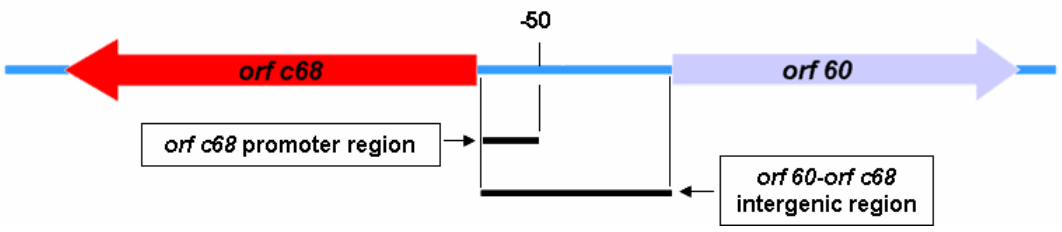

Fig. 10 Schematic representation of the DNA fragments used as probes for the EMSA experiments

Hydroxyl radical foot-printing
The *CopG/c68* intergenic region used as probe was amplified by PCR reaction with one of the two oligos labelled at the 5' end with T4 Polynucleotide Kinase (PNK, Roche). The pGEM-*CopG/c68* construct was used as template. The PCR product was purified by PCR Clean-up kit (Qiagen).
The binding-reactions were performed as described below:

- 1μl of probe DNA (~ 500 cps on a mini-monitor)
- 100 ng poly dI/dC
- 5μl 10 x TB buffer (500mM Tris-HCl pH 8,0, 800mM KCl, 250mM $MgCl_2$, 10mM DTT)
- ORF c68 (1-6 μg)
- To 50μl with water
- Incubate at 50°C for 20 min

Then the "hydroxyl radical cleavage" reactions were conduced as schematized below:

- To each binding-reaction tube 2μl 400 μMFe/800 μMEDTA, 2μl 20mM Sodium Ascorbate and 2μl 0,3% H_2O_2 were added to the side of the eppendorfs; to initiate reactions a pipette was used to push drops into binding reactions
- The cleavage was allowed to proceed for 10 min at room temperature
- To stop the cleavage reactions 5μl of stop solution (2μl 100 mM thiourea, 1μl 250 mM EDTA and 2μl 3M Sodium Acetate) were added

- The DNAs were then precipitated with 60 μl of 95% ethanol over night at -20°C

The sequencing reactions were performed according the instructions of the *fmol* DNA Cycle Sequencing system (Promega):

- 100 ng pGEM-*CopG/c68*
- 1,5 pmol oligo 5'end labelled
- 5μl DNA Sequencing 5X Buffer
- 1μl *Taq* enzyme
- Sterile water to final volume of 16 μl

4μl of the mix described above were added to each tube containing 2μl of the appropriate d/ddNTP Mix (G, A, T, C); the reactions were then conduced in a thermal cycler:
95°C for two min, then

95°C for 30 sec (denaturation)
42°C for 30 sec (annealing)
70°C for 1 min (extension)..... } 30 cycles total

The samples were then loaded on a 6% polyacrylamide (19:1), denaturing gel (7,6 M urea). The gel was dried and the radioactive signals were revealed exposition of an autoradiographic sheet for 12 hours at -80°C, as previously described.

Western blot analysis

The western blot analysis of ORF c68 expression was performed on protein extracts from cells of *S.islandicus* REY 15/4. 250 ml of cell cultures were harvested by centrifugation at two different growth phases and resuspended in 5ml 25 mM HEPES buffer pH7,0. The samples were then sonicated at 20KHz for 5 min (30'' ON, 30'' OFF) and centrifuged for 20 min at 20817 x g. 100μg of protein extracts were then separated by electrophoresis on a 15% polyacrylamide SDS-PAGE. The gel was then equilibrated in Transferring Buffer (25 mM Tris pH 8.0, 190mM Glycine and 10% methilic alcohol) and the samples were transferred from gel to a PVDF membrane by elettro-blotting. The PVDF membrane was submerged in methilic alcohol for 10 sec before using, to eliminate the hydrophobic properties, washed in H_2O for 2 min and equilibrated in Transferring Buffer for 5 min. The electro-blotting was performed for 2 hours (room temperature) at 100 Volts in Transferring Buffer with 20% ethanol added. .
The membrane was then treated as described below:

- staining with Ponceau Red to verify proteins transferring
- treatment in Blocking solution [5% dry milk in TBS (20 mM Tris, 150 mM NaCl, pH 7,5, with 0,1% tween added)] for two hours at room temperature to saturate the aspecific antibody-binding sites
- Incubation with anti-ORF c68 antibody for two hours at room temperature with gentle agitation
- washing in TBS with 0,1% tween added and then incubation with the secondary antibody (anti–FC rabbit) for 1 hour at room temperature
- extensive washing in TBS
- revelation by ECL.

Results

1. Characterization of the *Sulfolobus* host-SSV2 virus interaction

Accumulation of SSV2 and pSSVx DNA during the growth of *S. islandicus* REY15/4

At an early stage of our experiment, we prepared episomal SSV2 DNA using approximately equal amounts of *S. islandicus* REY15/4 cells from different cultivation batches. In some experiments, we obtained a large yield of episomal SSV2 and pSSVx DNAs; in others, a fairly small amount of plasmid DNAs. To account for this finding, we prepared episomal and total DNA using REY15/4 cells at different optical densities. We used Southern hybridization and the 4.69-kb SSV2 as well as the *Hind* III 1.7-kb pSSVx fragments as probes, (see material and methods) to identify the episomal SSV2 and pSSVx. Fig. 11 shows that whereas in early growth phases (0.5 to 0.9, OD_{600}) episomal SSV2 and pSSVx DNA was produced at a low rate, large amounts of SSV2 and pSSVx DNA accumulated at maximum cell density (1.25-1.3 OD_{600}). This suggests that there was an induction of SSV2 and pSSVx replication during the growth of the natural host. Furthermore, multiple topological forms of the SSV2 and pSSVX episomal genomes were observed after the induction of viral replication (**Figure 11**, 1.25-1.3 OD_{600}), consistent to the results obtained from the SSV1 studies (Nadal *et al.*, 1986).

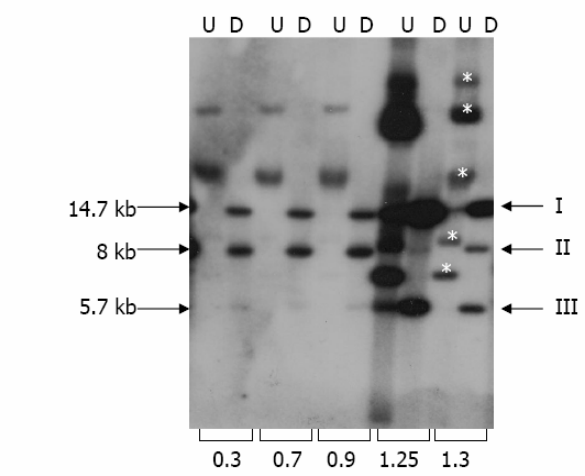

Fig. 11 Content of the virus DNAs in REY 15/4 cells of different optical densities The OD_{600} values of the *S. islandicus* REY 15/4 culture used for preparing the genomic DNA are reported at the bottom. Lanes labelled with "D" indicate the DNA samples digested with *ClaI* restriction enzyme (one cut for both SSV2 and pSSVx genomes). Lanes marked with "U" indicate undigested total DNAs. The linear SSV2 and pSSVx are indicated." I" indicates the linearized episomal SSV2 (14,7 Kb), "II" indicates the integrated SSV2 (8 Kb) and "III" indicates the linearized episomal pSSVx (5,7 Kb). The amount of total DNAs at 1,25 and 1,3 OD_{600} values was respectively ten and fivefold lower, than that used for samples at 0,3-0,9 OD_{600}. The *asterisks* indicate the multiple topological forms of the SSV2 and pSSVx episomal genomes observed after the induction of viral replication

Characterizing the SSV2 and pSSVx replication induction process

To shed further light on the increase in SSV2 and pSSVx observed during the growth of the natural REY host we collected REY15/4 cell samples at 2-hour intervals during incubation (see *Materials and Methods*). Over the first 24 incubation hours, no episomal SSV2 and pSSVx bands were visible in the total DNAs prepared from rapidly growing REY15/4 cells; thereafter, episomal SSV2 and pSSVx DNA accumulated rapidly reaching a maximum within 4 hours, i.e. between the 26th and 30th incubation hours (**Fig. 12**). This indicates that the increase of SSV2 and pSSVx must be traced to the induction of SSV2 and pSSVx replication which sets in 4 hours after the growth of host cells came to a stop.

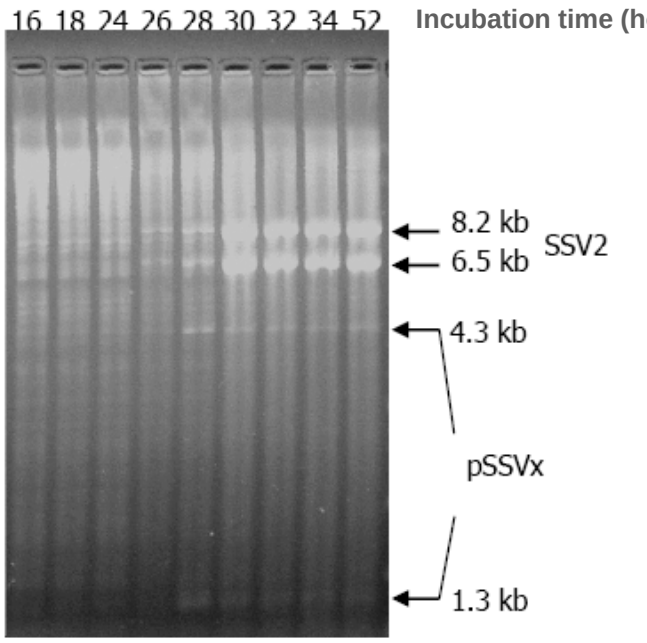

Fig. 12 Relative content of the episomal SSV2 and pSSVx genomes in the REY15/4 cells during incubation. The genomic DNA was prepared from REY15/4 samples indicated by solid cycles, digested with *Bam*HI, and analyzed by agarose gel electrophoresis. The agarose gel was stained with ethidium bromide and a picture was taken using a Kodak system. Two *Bam*HI fragments of the episomal SSV2 and pSSVx are indicated with their sizes.

It was previously shown (Arnold *et al*., 1999; Stedman *et al*., 2003) that SSV2 replication is not dependent on pSSVx in *S.solfataricus* P1 cells. We were interested in establishing if pSSVx contributes to the SSV2 replication induction observed in the natural host REY15/4. To obtain a pSSVx-free strain for investigation, a REY15/4 culture was left to grow continuously for approx. 100 generation times at an optical density below OD_{600} 0.4. Thereafter, we obtained single colonies via Gelrite plating and one of these, REY31A, was used to investigate the physiology of SSV2 replication, after checking carefully the absence of pSSVx (see material and methods). The analysis of SSV2 DNA content performed by quantification of the radioactive signals after Southern blot, revealed that the induction of SSV2 takes place in REY31A cells with the same modality and at the same extent described for REY15/4 cells (**Fig. 13**). These finding indicate that the presence of pSSVx is dispensable to the SSV2 replication induction process.

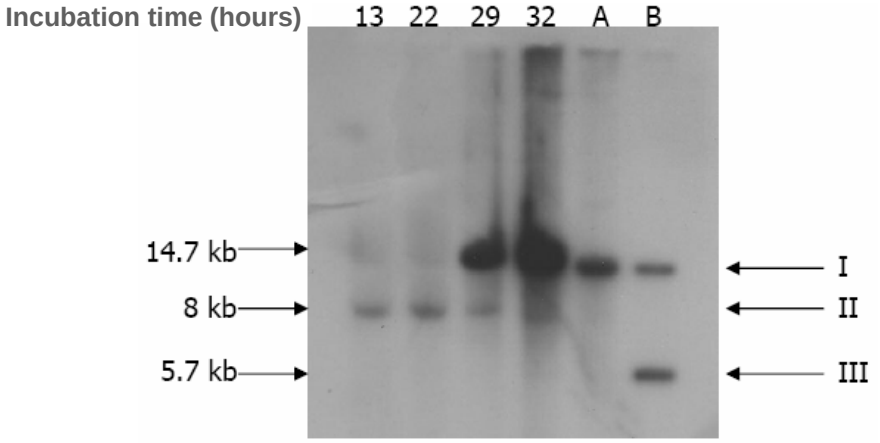

Fig. 13 Analysis of the growth characteristics of REY31A and evaluation of SSV2 induction level
Total DNAs from equal amounts of REY31A cells (normalised by OD_{600}) were prepared at the incubation times reported at the top. DNAs were digested with *Cla I* (one single cut for both SSV2 and pSSVx genomes) and the probes used are the SSV2 4,69 kb and a pSSVx-*Hind*III 1,7 kb restriction fragments, this latter included to demonstrate the absence of pSSVx. The size of SSV2 (I), pSSVx (III) and the host-virus hybrid fragment (II) are indicated on the left. *Cla*I digests of purified SSV2 (lane A) and SSV2 together with pSSVx (lane B) were loaded onto the same membrane as positive control for hybridisation. The radioactive signal of the host-virus hybrid fragment was used to normalise the amount of the DNA loaded and to determine the SSV2 amount at every growth stage (see material and methods).

To establish if SSV2-determined growth inhibition is a reversible process, we carried out a dilution experiment, i.e. we released the REY15/4 host cells into fresh medium at a 10-20:1 ratio after the virus induction process had set in. The growth rates and quantities of SSV2 DNA in the diluted cultures were estimated as described in Materials and Methods. Our results showed that host cells grow exponentially after a lag phase of varying length and that thereafter the average SSV2 cellular content reverts to ca. one copy per chromosome, as shown below in the Southern blot of **Fig. 14**:

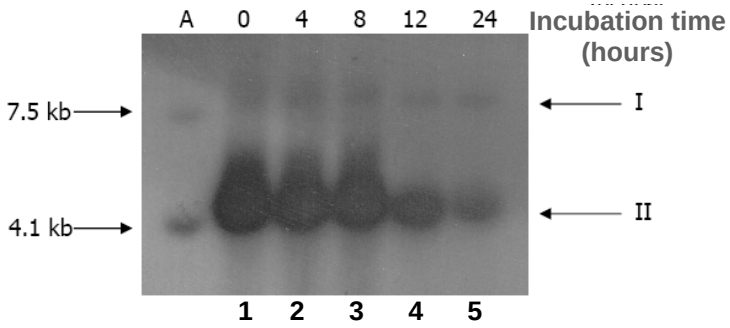

Fig. 14 The SSV2 induction process is reversible. Total DNAs from equal amounts of REY15/4 cells (normalized by OD_{600}) were digested with Hind III and the membrane hybridized with the coding region of the vp3 gene (13652-13930 bp on the SSV2 genome) and the REY 15/4 alba gene (*Sso*0962), as probes. The sizes of the radioactive signals of the chromosomal restriction fragment containing the alba gene (I) and the episomal SSV2 (II) are reported on the left. Lane A: Total DNA from REY15/4 cells, before the SSV2 replication had set in. Lanes 1-5: Total DNAs from REY15/4 cells once SSV2 replication induction had set in and harvested at different incubation hours after dilution into fresh medium. A progressive decrease of episomal SSV2 content occurs within 24 hours after dilution of SSV2-induced REY15/4 cells into fresh medium.

Virus packaging capacity during host growth

To decide if SSV2 DNA accumulation was caused by the lower virus packaging capacity of the host, we also measured the SSV2 virus titre of the REY15/4 culture. The virus particle titre was found to be fairly low (2-4 x 10^3 cfu/ml) in the exponential growth phase but rose 10-fold after the host had entered the late exponential growth-phase. Thereafter, it remained approx. 1 x 10^5 cfu/ml until the latest sample-taking points. These results indicate that when SSV2 replication is induced a concurrent, consistent packaging and extrusion of the virus particles occurs, and that the fact that free episomal SSV2 DNA accumulates in the cytoplasm must be traced back, not to any reduced viral DNA packaging and/or exportation capacity, but rather to a higher replication rate of SSV2.

The Mechanisms governing the induction of SSV2 replication

Compared to virus-free *Sulfolobus* cultures, for REY15/4 cells we always observed a sudden halt in growth and a substantial difference in final optical density values (typically 1.3 versus 2.5, OD_{600}), i.e. a sign pointing to growth inhibition. Since growth inhibition coincided with the onset of the viral DNA replication induction, we resolved to investigate the mechanisms governing these two processes and the relations between them. For this purpose we isolated a *S. islandicus* strain (which is closely related to the REY15/4 strain) and obtained virus-harbouring strain (SSV2-REY15A) as described in Materials and Methods. **Figure 15** shows that the three SSV2-REY15A cultures grew in the same manner as the REY15/4 strain. However, in the infected cultures, growth came to a stop at 1.3 (OD_{600}), whereas the uninfected REY15A culture continued to grow for 20 hours reaching cell density of 2.0-2.5 (OD_{600}). Analysing the total DNA in the SSV2- infected REY15A cells during growth, we found that SSV2 replication sets in within 4 hours after the end of cell growth, as in the natural host of SSV2. To establish whether the uninfected REY15A was actively growing beyond 1.3 OD_{600}, when we observed the sudden halt of SSV2-REY15A cells, we took 9 cell samples between 1.1-2.2 OD_{600} , determined the colony formation units (CFU) by Gelrite plating (see material and methods) and found a linear relationship between the CFUs and OD_{600} values of REY15A cells (not shown). Moreover, the CFU/OD_{600} fell into the range of 7.29 ± 0.66 x 10^8 indicating that the *Sulfolobus* cells were still able to divide when the OD_{600} of a REY15A culture increased up to 2.2 OD_{600}. Thus, the sudden halt of the optical density increase of SSV2-REY15A cultures reaching 1.3 OD_{600}, reflected growth inhibition occurring in actively dividing culture at a growth stage corresponding to an exponential growth phase.
Since REY15A and REY15/4 are very closely related to each other, we inferred that the same would be applicable to the natural host REY15/4, i.e. growth inhibition occurred at a late exponential growth phase and REY15/4 cells would continue to grow and reach a higher cell density level (as for REY15A) if it was cured of SSV2.

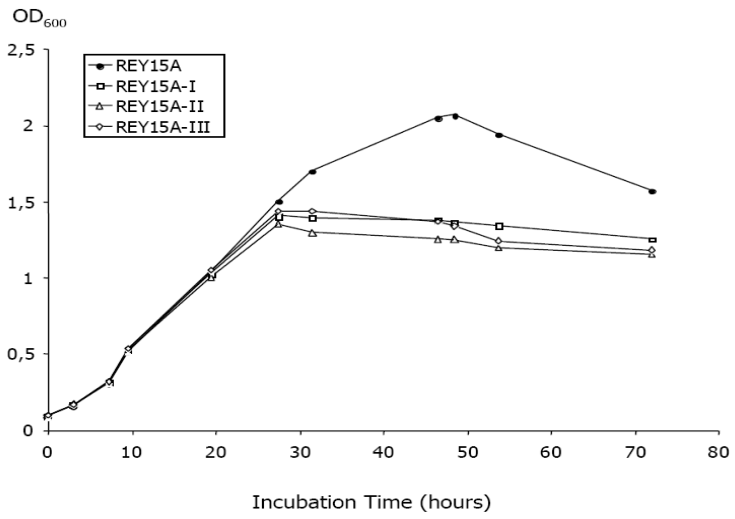

Fig. 15 Coupling of host growth inhibition and SSV2 replication induction in REY15A
S. islandicus REY15A and its three SSV2-containing strains, REY15A-I, II, III were grown in a TYS medium. The results were the same in three experiments carried out in succession

2. Construction of the pSSVx-derived *E. coli-S. solfataricus* shuttle vector

Transfection and isolation of *S.solfataricus* SSV2 transfectants

The strains P2 and GθW, a β-glycosidase defective mutant isolated earlier (Bartolucci *et al*., 2003), were tested for susceptibility to infection by SSV2 and pSSVx viruses produced by *S.islandicus* REY15/4. Supernatants of REY15/4 cultures generated growth inhibition halos on continuous lawns of both *S.solfataricus* strains (**Fig. 16**), similar to those observed for the strain P1 (Arnold *et al*., 1999a); like for SSV1 virus (Stedman *et al*., 1999), plaques did not form on *S.acidocaldarius* cells, namely no infection occurred. Cells extracted from areas of these plaques and revitalized in liquid cultures contained extrachromosomal DNA indistinguishable from that isolated from REY15/4, indicating the presence and active replication of both virus and satellite elements (**Fig. 17**). From this culture, clones infected only by the SSV2 virus, namely cured for the pSSVx, were isolated as single colonies on plates after suitable dilutions, as described for the P1 strain by Arnold *et al.* (1999a). The absence of the satellite virus was confirmed by restriction analysis of extrachromosomal DNA (**Fig. 17**) and Southern hybridization (data not shown).

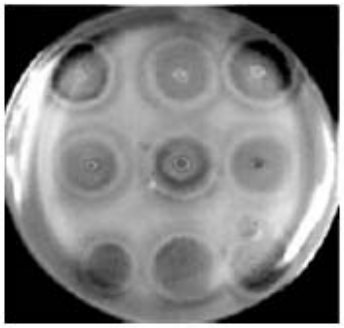

Fig.16 Transfection of *S.solfataricus* GθW with conditioned medium of *S.islandicus* REY15/4.
The spontaneous mutant GθW lacking the β-glycosidase activity was first tested for infection with both SSV2 and pSSVx by formation of growth inhibition areas in cell lawns on plate. Virus plaques of SSV2 and pSSVx were obtained by spotting aliquots of REY 15/4 supernatants (2 μl for the three plaques in the first row, 4 μl for those in the second row and 6 μl for those in the third row) onto a continuous indicator lawn of the GθW cells.

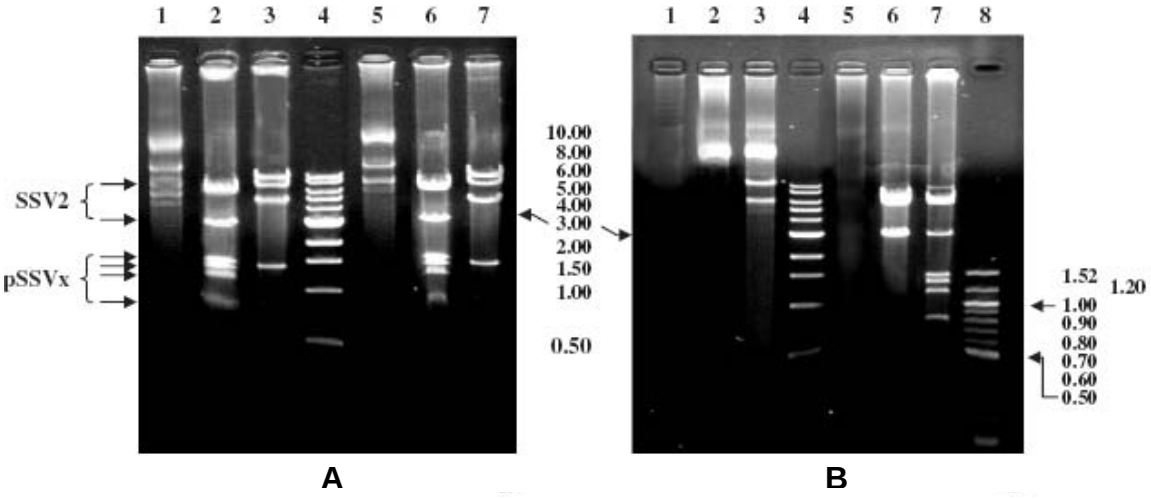

Fig. 17
Panel A Extrachromosomal DNAs (uncut in lanes 1 and 5), extracted from both the natural source *S.islandicus* (lanes 1–3) and the infected *S.solfataricus* (lanes 5–7), propagated after extraction from the plaques and revitalization, was cut with *EcoRI* (lanes 2 and 6) and *BamHI* (lanes 3 and 7). Episomal DNAs were confirmed to be identical and to contain both SSV2 and pSSVx DNAs, as demonstrated by the two distinct *EcoRI* restriction patterns. Lane 4: DNA molecular weight standards (kb).

Panel B The cells entrapped inside the plaques have been extracted from the plates, revitalized in liquid medium and then streaked onto solid medium to obtain single colonies. The comparative analysis with extrachromosomal DNAs from non-transformed recipient cells (lanes 1 and 5) and from the cells transfected with both viral and satellite DNAs (lanes 3 and 7) revealed the absence of the pSSVx plasmid in one colony isolated on plate and analyzed (lanes 2 and 6). Lanes 1–3: uncut DNAs; lanes 5–7: *EcoRI* digests; lanes 4 and 8: DNA molecular weight standards (Kb)

Shuttle vector construction

A sequence analysis on the pSSVx DNA was performed in order to locate regions that could be manipulated without affecting DNA replication and particle proliferation/ spreading, and thus representing candidate targets for site-specific insertion of foreign DNA sequences. A 200 bp segment was identified which contains the tail-to-tail intergenic region between the ORFs *C68* and *288*. The segment shows archaeal transcription termination signals (Reiter *et al*., 1988) and tendency to form hairpin loops as for rho-independent termination mechanisms. Moreover, it has an *Afl*III site (cut at the position 2812 on the pSSVx genome), useful for cloning, since it is situated in the 3' direction beyond the ORF288 transcription termination region, and far upstream of the ORF *c68* stop codon. Since the *Afl*III site is present in five copies on the pSSVx sequence (positions 677, 814, 982, 2812 and 4994), singly cut pSSVx genomes were generated by digestion with the restriction enzyme under conditions that allowed single cleavage of the DNA molecules presumably at every specific site (Stedman *et al*., 1999). After suitable modification of termini, these DNAs were inserted into *E.coli* pUC19 plasmid vector and specific insertion at the position 2812 produced the fusion plasmid pSSVrt (**Fig. 18**)

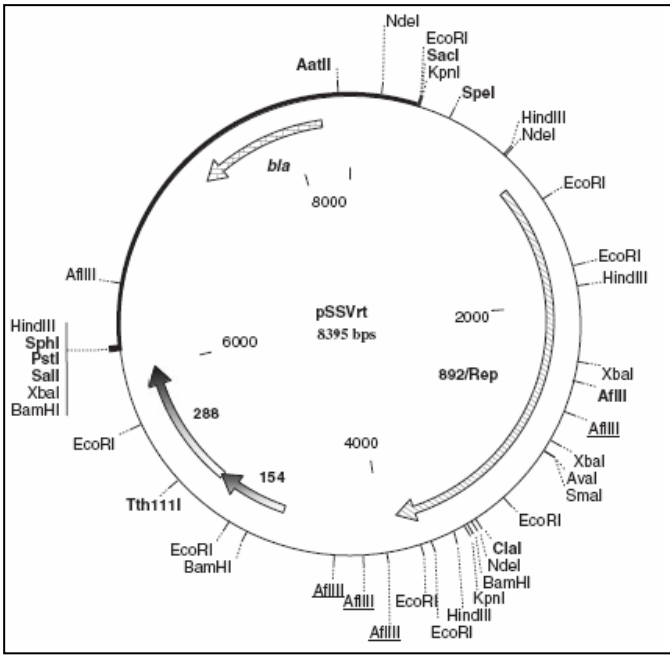

Fig.18 Plasmid map of the pSSVx/pUC19 shuttle vector.
The pSSVrt shuttle vector has been obtained by cloning the pUC19 bacterial sequence (heavy bar) into an *Afl*III restriction site (all *Afl*III sites are underlined) downstream of the ORF288 of the pSSVx plasmid and by propagation in *E.coli*. Main ORFs allow easy location of archaeal (replicase ORF892/Rep as well as ORF154 and ORF288 putative DNA packaging components of pSSVx) and bacterial (*bla*, ampicillin resistance gene) sequences; bold face letters indicate unique restriction sites

pSSVx-derived shuttle vectors for the *lacS* gene transfer and expression in the β-galactosidase deficient mutant GθW

β-D-galactosidase activity of *S.solfataricus* is displayed by the *lacS* gene product (Pisani *et al*., 1990; Grogan *et al*., 1991) and is responsible for the typical blue stain of the cells when exposed to the chromogenic substrate, X-gal; the lack of gene function in defective mutant strains is hence responsible for colorless colonies (Schleper *et al*., 1994). The *lacS* coding sequence with a 648 bp 3'-untranslated region (3'-UTR) was fused to a 448 bp fragment containing the promoter region of tf55a and the first five codons of its ORF; this 2578 bp expression cassette was inserted into the polycloning site of the pSSVrt vector, increasing its size to ~11 kb. A similar gene fusion was already demonstrated to efficiently complement a β-galactosidase defect in *Sulfolobus* when carried by a viral SSV1-derived shuttle vector (Jonuscheit *et al*., 2003). The plasmid was used to transform the strain GθW (ΔlacS) and after electroporation, cell culture was regenerated to allow spreading and subsequently seeded as spots onto plates as previously described (Bartolucci *et al*., 2003). After 2–3 days of incubation, the colonized areas were overlaid with X-gal and reincubated at 75°C. No color was developed even after prolonged incubation, namely no expression of β-galactosidase activity could be detected, suggesting the failure of the recombinant plasmid to transport and/or to express the *lacS* gene. Extrachromosomal DNA preparation did not contain the recombinant pSSVx confirming that interference could have occurred either at the level of DNA propagation or of the recombinant particle proliferation/spreading. Southern analysis of extrachromosomal DNAs from this primary transformants demonstrated that the DNA transfer and maintenance were unaffected whereas spreading was impeded, as indicated by specific bands that could be detected at constant but very weak intensity for cells withdrawn at different generation stages (data not shown). To confirm this preliminary results a size reduction of the pSSVrt vector was obtained by eliminating a redundant sequence in the *E.coli* plasmid moiety not necessary for replication and

ampicillin selection and producing the vector pMSSV. Moreover a smaller (2025 bp) *lacS* expression cassette provided by Dr C. Schleper (Jonuscheit *et al*., 2003) and containing a shorter 3'-UTR replaced the one used in the first attempt. This *lacS* expression cassette was inserted into the polycloning site of pMSSV, generating the expression vector pMSSV*lacS* (**Fig. 19**).

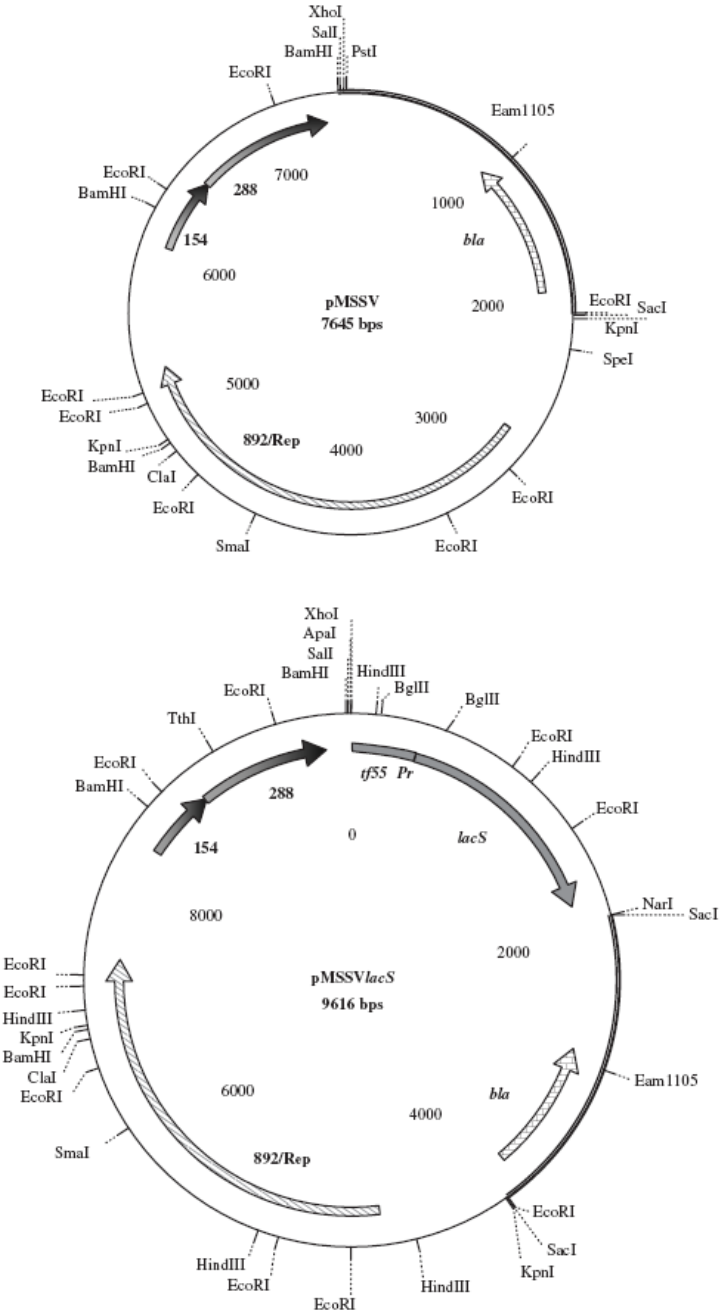

Fig. 19 Plasmid maps of the minimal plasmid pMSSV and of the β-glycosidase expression vector pMSSV*lacS*.
The bacterial moiety (indicated as a solid bar) of the plasmid pSSVrt was reduced in size, by eliminating every redundant sequence from the pUC19 vector and maintaining its ColE1 replicon and the ampicillin resistance marker (*bla*), to generate a new shuttle vector named pMSSV. An expression cassette containing the *lacS* gene fused to the thermosome tf55α subunit promoter was inserted into the newly constructed vector producing the expression vector pMSSV*lacS*.

Both pMSSV and pMSSV*lacS* were transferred into *S.solfataricus* GθW/SSV2 cells by electroporation. After culture propagation, extrachromosomal DNAs were prepared and analyzed by agarose electrophoresis (**Fig. 20**). The presence and growth-dependent accumulation of both vectors in GθW/ SSV2 revealed successful transformation and DNA replication, confirming that the plasmid size was critical for particle formation and spreading; moreover the *lacS* gene is harmless and does not induce recombination of the vector. Identical results were obtained when SSV2 lysogens of *Sulfolobus* GθW were transformed with the plasmid pSSVx*lacS* lacking the *E.coli* sequences and obtained by cleavage with *Sac*I and re-ligation (**Fig. 20**). All pSSVx derived plasmid failed to transform *Sulfolobus* if linearized prior to transfer. In fact, X-Gal staining test was positive on *Sulfolobus* GθW/SSV2 transformed with pMSSV*lacS* (or pSSVx*lacS*, data not shown) on liquid-cultured (**Fig. 21A**), plated (**Fig. 21B**), and primary infected cells (**Fig. 21C**). Plaques depicted in Figure 21C stained after incubation with the chromogenic substrate as a result of propagation of the engineered satellite virus. Since cell growth retardation in the plaques is directly proportional to the SSV2 virus titer which is maximal in the stationary phase cultures, turbidity and color faded progressively with increasing cell density. Therefore, the β-galactosidase was expressed from the recombinant plasmid and the engineered satellite virus had spread efficiently throughout the culture.

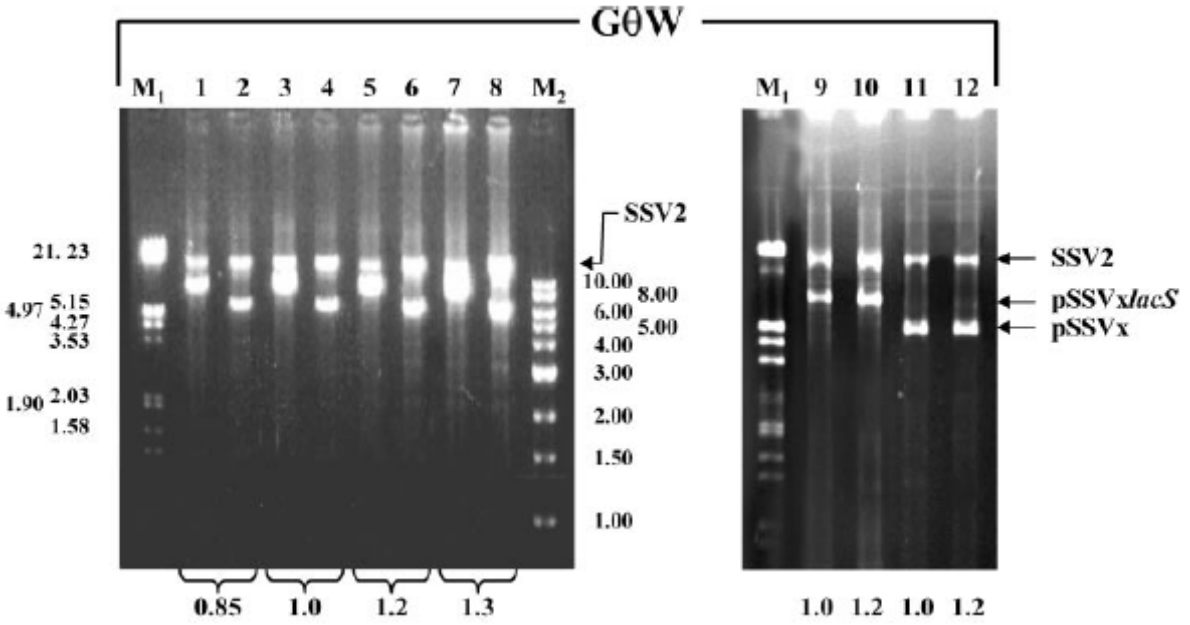

Fig. 19 Transformation of *S.solfataricus* with the vectors pMSSV and pMSSVlacS.
Competent cells of *S.solfataricus* GθW infected with the SSV2 virus were electroporated with the vector carrying the β-glycosidase gene pMSSV*lacS* and also with pSSVx*lacS*, a pMSSV*lacS*-derived plasmid obtained by excision of the *E.coli*-specific sequences and re-ligation. pMSSV and wild-type pSSVx were also transferred into the GθW/SSV2 lysogen for comparison of vector transfer and propagation efficiency in the presence and absence of the *lacS* gene. Extrachromosomal DNAs extracted from transformed cells after propagation of the cultures were checked by agarose electrophoresis for the presence of pMSSV*lacS* (lanes 1, 3, 5 and 7) and pMSSV (lanes 2, 4, 6 and 8) at the different growth stages indicated as cell densities (OD_{600}). Similarly, recovery of the pSSVx*lacS* plasmid (*SacI* cut, lanes 9 and 10) was monitored in comparison to the parental pSSVx (*Sma*I cut, lanes 11 and 12). Mobility of the virus SSV2 is highlighted. M1 and M2: molecular weight markers (Kb)

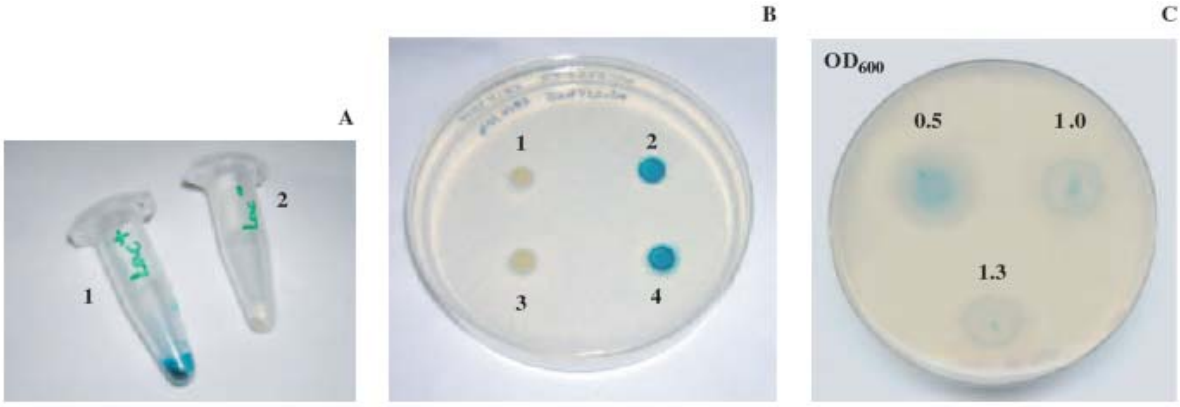

Fig. 21 Complementation of the β-glycosidase mutation in *S.solfataricus* GθW and spreading of the pMSSV*lacs* vector.
The expression vector pMSSV*lacS* was transferred into *S.solfataricus* GθW cells lysogens for SSV2. Successful transformation was checked after propagation for several generation by direct exposure of cell pellets from liquid cultures to X-Gal and development of the blue color [**A** (1) Lac+, pMSSVlacS transformed cells; (2) Lac–, cells transfected with pSSVx]. Maintenance of the plasmid was also confirmed for colonized areas on plates and X-Gal test [**B** (1) Strain GθW; (2) wild-type strain Gθ; (3) SSV2/ pMSSV-infected GθW; (4) SSV2/ pMSSV*lacS* -infected GθW].
C: 4-ml aliquots of culture supernatant of SSV2/pMSSV*lacS*-infected GθW, withdrawn at the different cell densities indicated, were spotted onto a continuous lawn of uninfected GθW cells. The spreading of the recombinant satellite particles was revealed by the blue color developed on the plaques (primary infected cells) upon X-Gal exposure. Stain intensity depended on the extent of cell growth retardation in the plaques and was progressively less intense for increasingly higher virus titers.

pMSSV*lacS* DNA prepared from *S.solfataricus* could be re-transferred into *E.coli* without suffering recombinational rearrangements. Plasmid preparations and total DNA from *Sulfolobus* transformants were analysed by restriction analyses and Southern hybridization that confirmed the maintenance of the vector at the same levels of the parental pSSVrt. No vector integration into the host chromosome occurred since the recombinant vector exhibited the same restriction pattern as the DNA prepared from *E.coli* and no signal relative to the plasmid could be detected on genomic DNAs in the Southern blots (**Fig. 22**). Bands corresponding to chromosomal tf55α promoter (in all strains tested) and *lacS* gene (absent in GθW) could also be visualized as internal controls for detection of single copy chromosomal sequences. Signal assignment of the restriction fragments of tf55α gene was performed on the basis of the sequence and localization on the P2 strain genome, whereas the mapping performed by Bartolucci *et al*. (2003) allowed the identification of *lacS* gene specific signals. A similar Southern analysis confirmed that the pMSSV*lacS* copy number varied from 10–15 (mid-log phase) to ~130 molecules per cell (stationary phase).

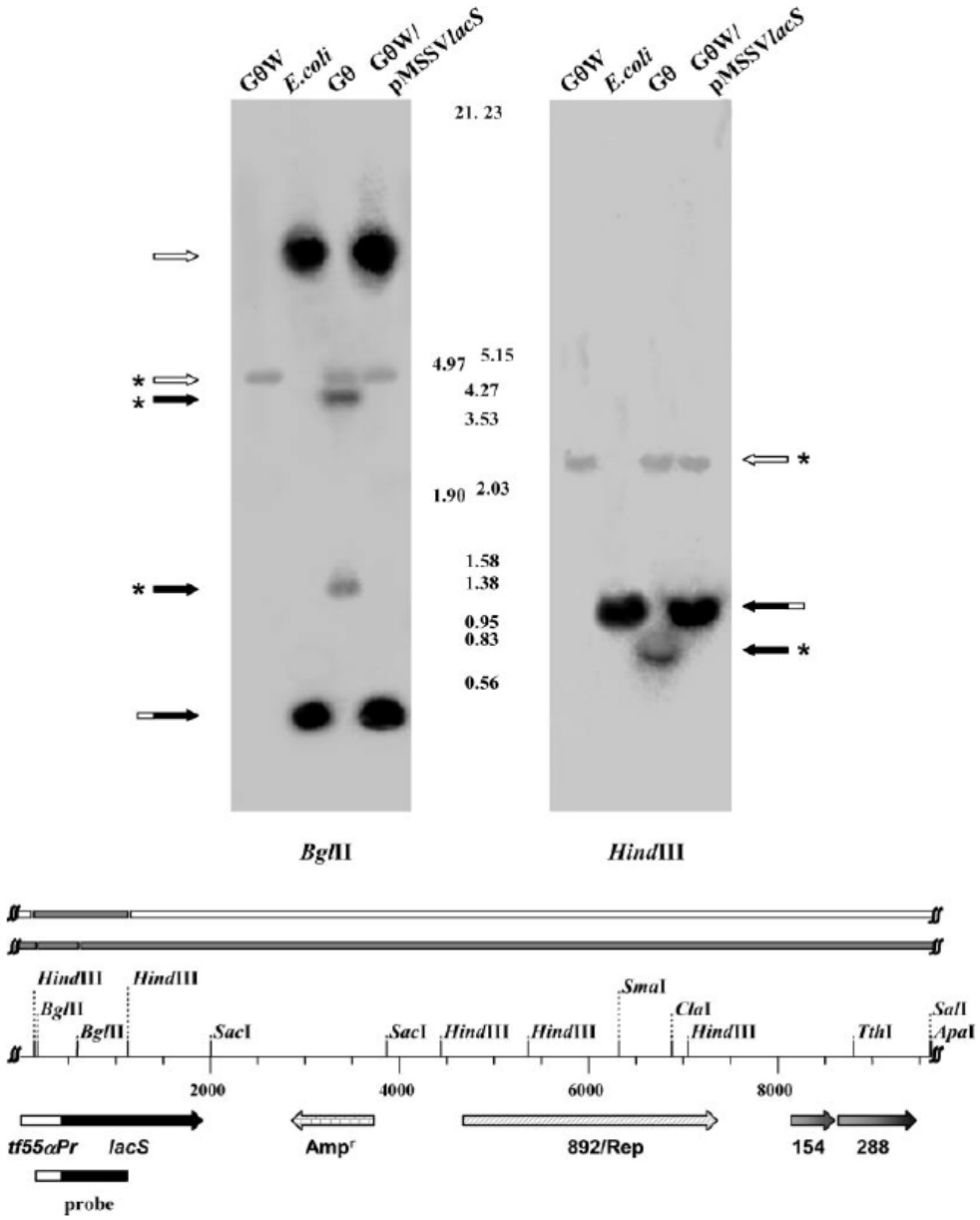

Fig. 22 Southern analysis of pMSSV*lacS* transformants.
For hybridization, total DNAs from transformant cells (GθW/pMSSV*lacS*) and from the recipient mutant strain GθW as well as from wild-type cells Gθ were cut with *Bgl*II and *Hind*III as indicated. The recombinant pMSSV*lacS* vector prepared from *E.coli* before transformation was used as a reference for correct restriction patterns (molecular weight standards are indicated). White and black arrows indicate hybridization to tf55α and *lacS* gene sequences, respectively, the asterisks distinguishing signals of the chromosomal copies. Signals of the tf55α–*lacS* gene fusion on the vector are highlighted by white/black arrows. The scheme on the bottom represents the restriction patterns for the two enzymes on the map of the linearized pMSSV*lacS* and the fragments producing positive signals for hybridization (grey bars).

3. Structural-functional analysis of the gene product ORF c68 of the pSSVx

Expression and purification of the 6Histidine-tagged ORF c68 protein

The gene *orf c68* was cloned into the *E.coli* expression vector pET30 a+ (Novagen) under control of the T7 promoter; the cloning was performed so that a 6Histidine-tag could be added to the C-terminal portion of the recombinant protein. BL21 (DE3) RIL *E.coli* cells were used as host for the heterologous expression. Upon induction crude extracts were prepared and the ORF c68 protein was purified to homogeneity by two steps of purification: 1) a thermal denaturation at 60 °C for 20 min to remove most of the host mesophilic proteins, 2) a Ni-affinity chromatography (**Fig. 23**).

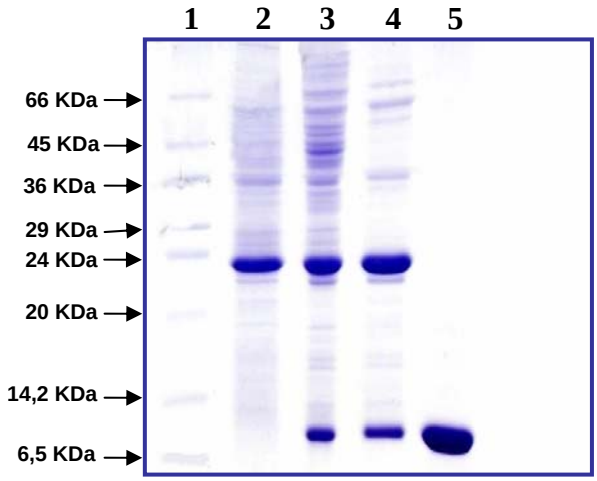

Fig. 23 Purification of recombinant 6His-tagged ORF c68.

SDS-PAGE of protein extracts at various stages of the purification of 6His-tagged ORF c68. Lane 1: Molecular weight markers; lane 2: crude extracts from *E.coli* BL21(DE3)/pET30-c68 not induced with IPTG; lane 3: crude extracts from *E.coli* BL21(DE3)/pET30-c68 induced with 0,5 mM IPTG; lane 4: sample after incubation at 60°C for 20 min ; lane 5: sample after His-trap affinity chromatography

The 6His-tagged ORF c68 showed a great instability in solution and considerable decreases of its concentration were observed. Because of the small dimension of the ORF c68 recombinant protein (~8KDa), the presence of the 6Histidine-tag probably constituted a destabilizing factor for a correct folding.
However, the 6His-tagged protein was used to produce anti-ORF c68 antibodies in rabbit, necessary for the *in vivo* expression analysis of the recombinant protein.

Expression and purification of the untagged ORF c68 protein

The gene *orf c68* was cloned into the *E.coli* expression vector pET30 a+ but the STOP-translation codon of the *orf c68* gene was unchanged, so that no tag could be fused to the recombinant protein. BL21 (DE3) RIL *E.coli* cells were used as host for the heterologous expression. Upon induction crude extracts were prepared; the recombinant ORF c68 represented about 10% of the total protein. The ORF c68 protein was purified to homogeneity by a cation exchange and gel filtration chromathographies (**Fig. 24A**).
To assess the quaternary structure of the ORF c68, analytical gel filtration (on Superdex 75 PC 3.2/3.0 column) of purified ORF c68 was performed. The protein was eluted at a volume accounting for a dimeric form (**Fig. 24B**).

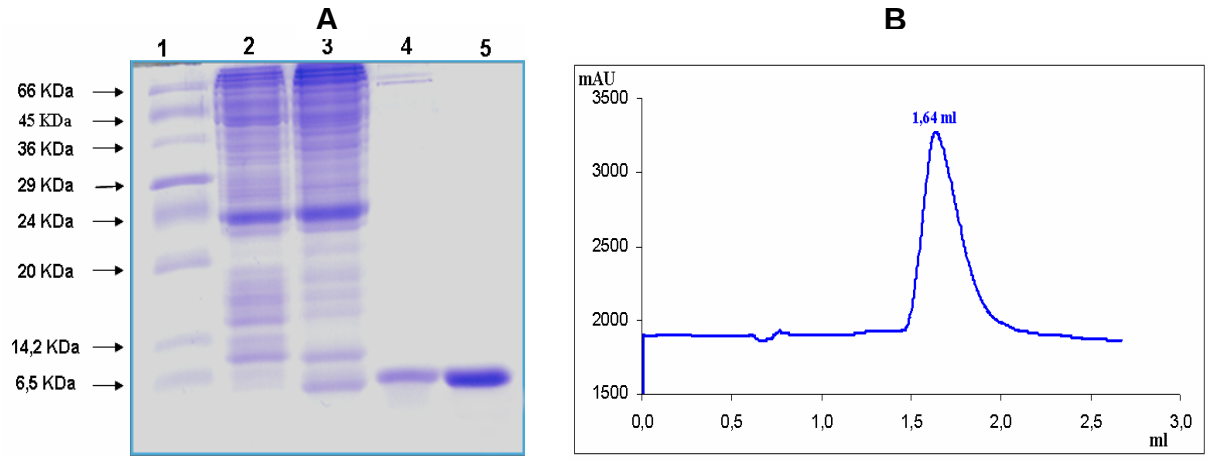

Fig. 24 Purification of recombinant untagged ORF c68 and determination of its quaternary structure.
SDS-PAGE of protein extracts at various stages of the purification of ORF c68. Lane 1: Molecular weight markers; lane 2: crude extracts from *E.coli* BL21(DE3)/pET30-c68 not induced with IPTG; lane 3: crude extracts from *E.coli* BL21(DE3)/pET30-c68 induced with 0,5 mM IPTG; lane 4: sample after cationic-exchange chromatography; lane 5: sample after gel-filtration chromatography (**Panel A**)
The purified recombinant ORF c68 protein was loaded onto a Superdex 75 PC 3.2/30 column connected to AKTA^{TM} chromatography system and eluted at 1,64 ml. The corresponding molecular weight (M) was determined by interpolation of a calibration curve as described in materials and methods (**Panel B**).

In vivo expression analysis of ORF c68

As previously described, an induction of SSV2 and pSSVx replication occurs during the growth of the natural host. The replication of the two genetic elements was induced at maximum cell density (1.25-1.3 OD_{600}) and large amounts of SSV2 and pSSVx DNA were detectable whereas in early growth phases (0.5 to 0.9, OD_{600}) episomal SSV2 and pSSVx DNA was produced at a low rate.
Moreover, it has been reported by Northern blot analysis that the *orf c68* gene gives rise to two differently sized transcripts of about 200 and 500 nt originated from the same strand (Contursi *et al*., 2007). The content of the 200-nt transcript, which became the predominant one, increased gradually even after the pSSVx copy number had reached its maximum. This expression profile has suggested an up-regulation for the *orf c68* and a complete transcription repression of the 500 nt species as soon as the pSSVx has reached the maximum copy number.
To analyse the expression levels of ORF c68 *in vivo,* cells of the natural host REY 15/4 were harvested at middle (0.6 OD/ml) and late exponential (1.3 OD/ml) phases, before and after the induction of the SSV2 and pSSVx replication, respectively. Western blot analysis was performed in relation to the induction-dependent increase of the pSSVx copy number. The native ORF c68 protein was detected only in the late logarithmic phase of cell growth, namely after the steep increase of pSSVx copy number (**Fig. 25**)

Fig. 25 *In vivo* expression analysis of ORF c68 by Western blot
Western blot was performed on total proteins extracted from cells harvested at two different growth phases. Lane 1: purified ORF c68 protein used as a positive control (100 ng); lane 2a and 3a: protein extracts from cells grown up to middle (0.6 OD/ml) and late exponential (1.3 OD/ml) phases, respectively. Extrachromosomal pSSVx DNA extracted from cells at the same two phases of growth is shown in the agarose gel, the induction of the pSSVx replication is detectable in the late exponential growth phase (lane 3b).

Bioinformatic analysis

The *orf c68* is the only reverse oriented pSSVx putative gene and has no homologs in the crenarchaeal family of pRN plasmids (Arnold *et al*. 1999). It is highly similar (42/53% identity/similarity) to ORFc56 of pSSVi (Wang *et al*. 2007) and also finds its homologues in the integrated elements of the *Sulfolobus tokodaii* (Kawarabayasi *et al.* 2001) and *S.acidocaldarius* (Chen *et al.* 2005) genomes. In order to get insight on the function of ORF c68, we resolved to perform secondary structure prediction. The presence of an α-helix flanked by four β- strands in the ORF c68 protein was revealed. A bioinformatic analysis performed with the PSI-BLAST and CDD search revealed that this putative fold is present in a family of prokaryotic factors represented by Abr-B (**Fig. 26**). These factors were demonstrated to be repressors or activators or ambiactive, depending on the specific case. As previously described, Coles *et al*. (2006) have shown that in the case of AbrB transcription factor, the four strands formed two β hairpins, which dimerized in a singly interdigitated way to form an eight-stranded barrel, which has been named the *swapped-hairpin* barrel similar to, but topologically distinct from, double-psi barrels**.**

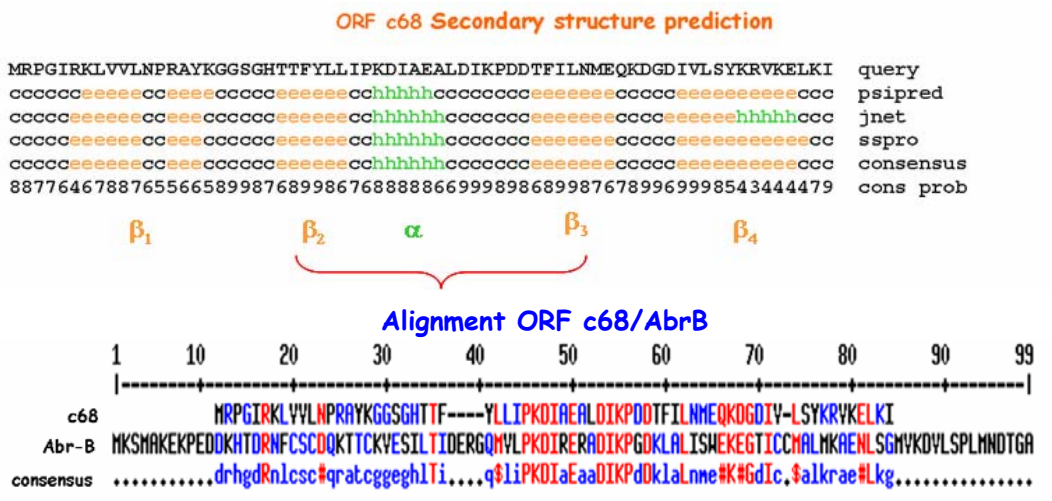

Fig. 26 A Secondary structure prediction of the ORF c68 The four predicted β-strands are indicated in orange and the α-helix in green; the β-α-β core, typical of the β-barrel fold, is present. The **panel B** shows the alignment of the ORF c68 protein and the transcriptional factor Abr-B from *Bacillus subtilis* in which the mostly conserved residues are indicated in red.

c68 is a DNA-binding protein and binds to its own promoter

The DNA-binding activity of the purified recombinant ORF c68 was investigated by gel retardation assays with DNA fragments upstream of the *orf c68* gene and using amount of ORF c68 ranging from 4 to 8 μM. The first probe used for EMSA experiments was a [α-^{32}]dATP labelled 50 bp DNA fragment containing the promoter region of the *orf c68* gene (ups c68). The DNA fragment was shifted upon addition of increasing amounts of the ORF c68 protein even in the presence of 500 ng of competitor DNA (polidI/dC) (**Fig. 27**). The strongest shif signal was revealed when 8μM of ORF c68 was used.

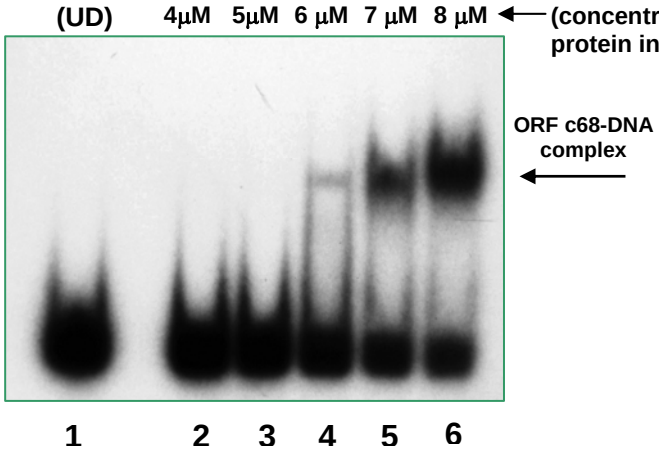

Fig. 27 Binding of ORF c68 to DNA analysed by EMSA
The ORF c68 was able to bind a 50 bp DNA fragment containing its own promoter region, in the presence of a large amount (500ng) of non-specific competitor DNA (poli dI/dC).

Lane 1: unbound DNA probe (UD); lane 2-6: 0.03 ng (3000 cpm) of labelled probe incubated with increasing amounts of purified ORF c68 protein. The ORF c68 concentrations in the final reaction volume (15 μl) are indicated on the top.

Displacement experiments to determine the specificity of binding

In order to verify the specificity of binding of ORF c68 to its own promoter, displacement experiments were performed. As aspected for a specific binding, when EMSA assay was performed in the presence of increasing amounts of specific competitor DNA (non labelled 50 bp DNA fragment containing the *orf c68* promoter), the binding of ORF c68 to its own promoter region was gradually displaced (shown in **Fig. 28 A**)
Conversely, the binding of the ORF c68 to its own promoter was not affected by the presence of increasing amounts of non-specific competitor DNA (the polilinker region of the pUC 28 vector) (**Fig. 28 B**).

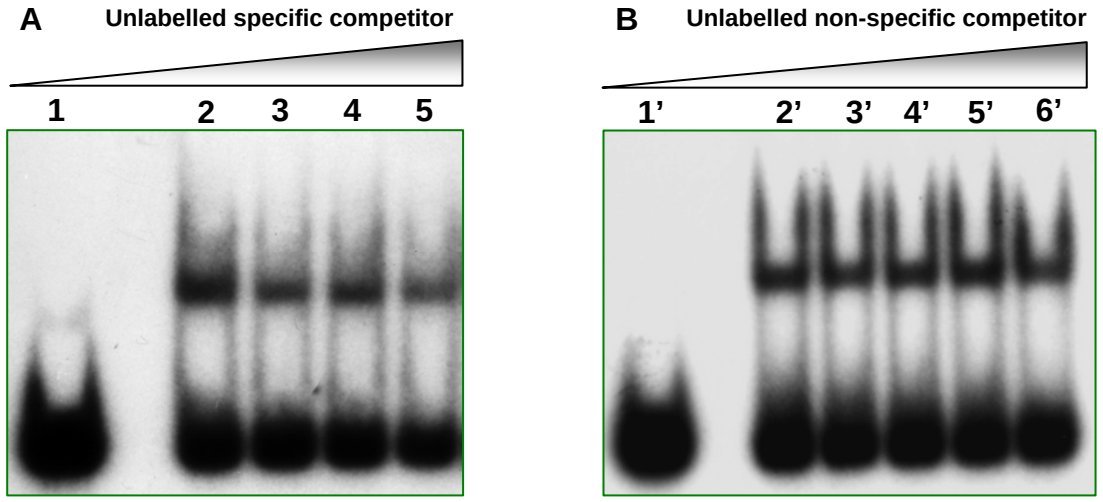

Fig. 28 Displacement assays of the binding of ORF c68 to its own promoter
ORF c68 protein binds specifically to its own promoter.
A Lane 1: 0.03 ng (3000 cpm) of labelled 50 bp DNA fragment containing the *orf* c68 promoter region; lane 2-5 labelled probe incubated with purified ORF c68 (8μM) in presence of increasing amounts (0.03 ng-0.3 ng-3 ng and 30 ng respectively) of specific competitor DNA.
B Lane 1': unbound DNA probe (as described for the experiment shown in the panel A); lane 2': labelled probe incubated with ORF c68 (8μM) without competitor DNA; lanes 3'-6': labelled probe incubated with ORF c68 protein (8μM) in presence of increasing amounts (0.03 ng-0.3 ng-3 ng and 30 ng respectively) of non-specific competitor DNA (polilinker region of the pUC 28 vector).

Binding of ORF c68 to *orf 60-orf c68* intergenic region

Interaction of the ORF c68 protein with a larger region encompassing the promoter of the *orf c68* as well as of the adjacent *orf 60* gene, was also investigated. A schematic representation of this DNA region is reported in **Fig. 10** (see materials and methods). As shown in **Fig. 29** ORF c68 protein binds the *orf 60-orf c68* 130 bp intergenic region. At relatively higher concentrations, the ORF c68-DNA interaction gives rise to an additional retardation band (complex 2).

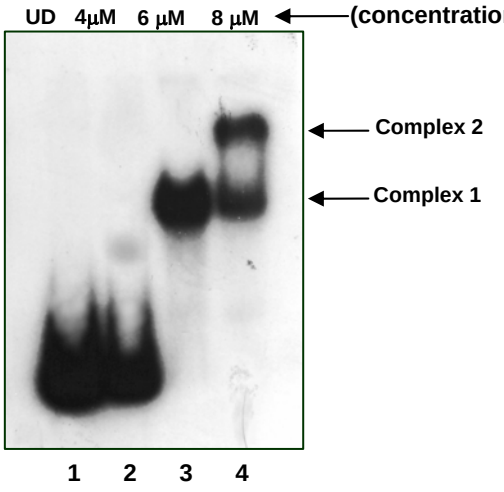

Fig. 29 Binding of ORF c68 to *orf 60/orf c68* intergenic region ORF c68 protein was able to bind also the 130 bp intergenic region in presence of 500ng of competitor DNA (polidI/dC).
Lane 1: unbound DNA probe (UD); lane 2-4: 0.03 ng (3000 cpm) of intergenic *orf 60/orf c68* DNA fragment incubated with increasing amounts of purified ORF c68 protein.
At concentration of the ORF c68 protein higher than 6μM was used, an additional retardation band was clearly visible (complex 2).

Foot printing analysis

The results described above indicated that ORF c68 was able to bind specifically its own promoter region as well as the *orf 60/orf c68* intergenic DNA fragment which encompasses the *orf 60* promoter. To map the binding sites of the ORF c68 protein, the whole *orf c68*/*orf 60* intergenic region was used for a hydroxyl radical foot-printing analysis. Increasing amounts of purified ORF c68 protein (from 1 μg up to 6 μg) were assayed as described in *materials and methods*. As shown in **Fig. 30** the ORF c68 protects three main contiguous nucleotide stretches; one is included in the 50 bp c68 5'-flanking region (1, indicated in red) and the other two stretches, clearly identified by the flanking hypersensitive nucleotides, are located further upstream (2 and 3, indicated in blue) upstream the *orf 60* promoter region.

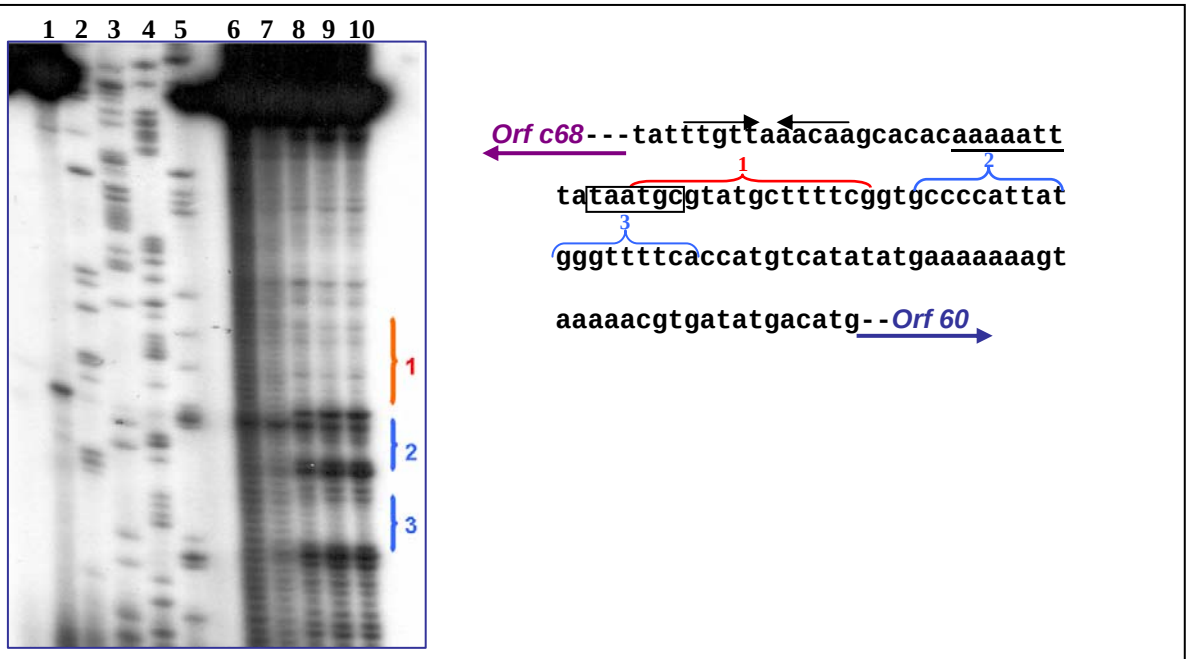

Fig. 29 Interaction of the ORF c68 with the *orf 60/orf c68* intergenic region Lane 1: 5'-labelled *orf 60/orf c68* intergenic region used as control; lane 2-6: sequencing reactions performed on the DNA probe (G, A, T, C respectively); lane 6: 5'-labelled *orf 60/orf c68* intergenic region after hydroxyl radical cleavage but not pre-incubated with the ORF c68; lane7-10: the intergenic region after pre-incubation with increasing amount of the ORF c68 protein (1-2-4-6 μg, respectively) and then hydroxyl radical cleavage. In the lanes 8,9 and 10 the three main areas of the intergenic region protected by the interaction with ORF c68 protein are clearly detectable.
On the right the sequence of the *orf 60/orf c68* intergenic region is shown. The putative TATA box and BRE are under lined and boxed respectively; palindromic repeats sequences are highlighted by opposite arrows.

Discussion

This work is part of a project aimed at developing innovative genetic systems for hyperthermophiles *Archaea*. The *in vivo* and *in vitro* studies of gene function as well as genetic manipulation of microorganisms need appropriate DNA-transformation systems, genetic markers and suitable vectors and hosts. Moreover, the homologous and heterologous expression of genes is a prerequisite for most biochemical studies of thermostable proteins and thermozymes that, because of their peculiar stability, are good candidates for biothecnological industry.

A vast variety of systems have been developed for protein production in members of the *Bacteria* and *Eukarya*, using numerous combinations of vector and promoter systems. Conversely, members of the *Archaea* are much less amenable to genetic manipulation because of their harsh growth conditions and low transformation efficiency. Systems for the production of recombinant proteins have been set up only for few archaeal *genera*, including halophiles and methanogens. Mesophilic hosts, in particular *Escherichia coli*, have been used to produce thermostable proteins for biochemical characterization and crystallographic studies. However, a considerable number of proteins of hyperthermophiles fold into their native state only under natural conditions of high temperature or in the presence of their native cofactors. Furthermore, the production of recombinant and tagged proteins in native thermophilic hosts allows the identification of associated factors or even larger protein complexes. At present, the state of art about manipulative genetics in hyperthermophilic *Archaea*, although still at a stage of infancy, strongly proposes *S.solfataricus* as a good candidate among organisms that can be used for biothecnological applications as "*cell factories*".

S.solfataricus represents one of the more extensively studied species among the hyperthermophiles *Archaea* and, thanks to its versatility, it is considered a model for biochemical and molecular characterizations. Since many studies on the transcription, translation, and replication of the extremophile *S.solfataricus* have been performed *in vitro*, it is highly desirable to develop genetic tools for *in vivo* studies and for high-level production of proteins in this organism. The construction of an innovative vector for *S.solfataricus* was the purpose of the present work. Sulfolobales harbour a vast variety of virus/plasmids that represent a promising base to construct genetic systems. In particular, the pSSVx/SSV2 system appeared to be interesting for the construction of an innovative cloning/expression vector. The SSV2-pSSVx viruses have been isolated from *S.islandicus* REY 15/4 and represent the first archaeal example of helper and "satellite" virus system. Interestingly pSSVx and its helper SSV2 can be successfully transferred from their natural host into *S.solfataricus* cells (Arnold *et al*., 1999). In the foreign host the two genetic elements show a spreading-efficiency comparable to that observed in the natural one.

A physiological characterization of the SSV2-pSSVx genetic system in the natural host as well as in the well characterized *S.solfataricus* host, was performed (Contursi *et al*., 2006). We have demonstrated that SSV2 DNA replication is characterized by a physiological induction in the natural host REY 15/4. The episomal SSV2 was produced at a rate of just one copy per chromosome during the exponential growth phase and a 50-fold increase occurred within 4h upon the growth stop. Interestingly, the pSSVx genome was induced in a growth-dependence fashion, similarly to its helper SSV2. We have shown that growth inhibition of the host concurs with the virus replication induction and the inhibition takes place at a late exponential growth

phase. Moreover, the inhibition effect is a reversible. Upon being released into a fresh medium, the SSV2-induced *Sulfolobus* culture resumes its exponential growth and host regains control over SSV2 replication.
Whereas it appears that the replication induction must have been turned on by unknown internal signal(s), either originated from virus or host, it is not clear if the growth inhibition constitutes part of host self-defence mechanism or it is conferred by a signal related to the SSV2 replication induction and/or another SSV2 signal. The physiological analysis pointed up a role of the SSV2 virus in host survival and a mutually beneficial interaction (like those of some bacteriophage-host systems), which increase the host's metabolic activity and/or environmental fitness (Weinbauer 2004). Clearly, there is interesting molecular biology behind these phenomena and further investigation should reveal mechanisms governing host-virus interactions as well as co-evolution of viruses and their hosts in *Archaea*.
The physiology of the pSSVx/SSV2 system was investigated also in *S.solfataricus* since, as mentioned above, it represents a suitable host for developing genetic systems. We demonstrated that in *S.solfataricus* cells infected with SSV2 and pSSVx the replication of the two elements is correlated with the growth rate of the host but the replication induction observed in the natural host does not occur in the foreign one (Contursi *et al*., 2006). The virus particles production in the *S.solfataricus* host starts during the exponential growth, reaching its maximum peak before cells approach to the stationary phase and eventually decreases.
As described above, in the natural host REY 15/4 growth inhibition of the host concurs with the SSV2 and pSSVx replication induction. Once settle in the foreign *S.solfataricus* host, SSV2 no longer inhibits host growth and this is in strict contrast to what is seen in early infected cells. Consequently, there is a major difference in growth rates between the early SSV2-infecting cells and SSV2-harbouring strain of *S.solfataricus*: the former showed a greatly retarded growth and the latter exhibited a similar rate of growth as an uninfected control. This suggests that a series of interactions between SSV2 and its foreign host will eventually lead to a co-existence harmony between them as seen for the SSV2-carrying *S.solfataricus* strains. It is thus very interesting to investigate whether this conversion of host response is a general scheme of the host defence of *S.solfataricus* to viral infection and to study the molecular mechanisms governing it.
pSSVx genetic element was chosen to develop a genetic system since it could be manipulated more easily than SSV2 thanks to the small dimension of its genome and to its genetically plasmidic nature (Aucelli *et al*., 2006). pSSVx is also able to spread in cell culture like a virus, packaged in lemon-shaped viral particles, only in the presence of SSV2. The partial viral nature of the pSSVx make it a powerful tool to overcome the low transformation efficiency of *S.solfataricus*, a serious limit for the development of efficient genetic systems for *Archaea*.
Hence, to realize a shuttle vector *E.coli/S.solfataricus* the selection of a *S.solfataricus* strain stably infected with SSV2 virus, able to guarantee the propagation of the pSSVx-derived vector, was strongly required. As already described (see "Results" session), the SSV2 viral infection generated "plaques of inhibition growth" on a continuous layer of host cells; the plaques were constituted by cells which, after revitalization in liquid medium, were stably infected with SSV2. This mechanism allowed us to obtain the GθW/SSV2 strain as lysogen host for our shuttle vector pSSVx-derived. This host strain presented two important features: (1) the presence of the SSV2 virus, necessary for the spreading of the pSSVx-based vector, (2) the absence of the β-glycosidase gene, responsible for the unique cellular β-

galactosidase activity (Cubellis M.V. *et al*. 1990; Schleper C. *et al.* 1994), and therefore the opportunity to use the complementation for the selective isolation of the transformants. The efficient propagation in both bacterial and archaeal hosts represented an important requirement of an *E.coli/S.solfataricus* shuttle vector. The pUC 19 *E.coli* vector was chosen for propagation in the bacterial host. On the other hand, an accurate analysis of the pSSVx sequence was necessary for the choice of the site-specific insertion pf pUC19 into the pSSVx genome. This preliminary study was crucial to preserve physiological the characteristics as well as propagation and replication of the pSSVx also in the deriving pSSVx-based shuttle vector. Hence, an extended intergenic region between the *orfs* c68 and 288 on the pSSVx genome appeared potentially useful for the insertion of the *E.coli* pUC19 vector. This region showed archaeal transcription termination signals and tendency to form hairpin loops as for the rho-independent termination mechanisms. The vector resulting by this molecular fusion pUC19-pSSVx was named pSSVrt.

An interesting aspect emerged in this study was the limiting size of the foreign genes which could be cloned into the shuttle vector. The insertion of foreign sequences into the pSSVrt vector allowed also to determine the limiting upper size (~11 kb) of the DNA to be accommodated in the virus particles. The additional sequences severely affected packaging and spreading rather than the transfer and/or the replication/maintenance of the viral DNA. These results also indicated that transport in viable virions and no other mechanism, such as conjugation, was responsible for cell-to-cell transfer of this genetic element. Hence, a “minimal” vector pMSSV was devised by reducing in size the pSSVrt DNA by:

- elimination of a non-essential sequence from the bacterial moiety of the original pSSVrt vector, without affecting the ColEI region and the *Amp* ^{r} gene necessaries for the replication in the bacterial host and the selection of the transformants, respectively.
- the use of a *lacS* reporter gene with a shortened 3’-UTR

pMSSV allowed the insertion of genes of a wider size range at least up to the 2.0 kb of a smaller tf55α/*lacS* expression cassette (Jonuscheit *et al*., 2003). The high propagation efficiency of the pMSSV vector, in SSV2 lysogens of the strain GθW, was found to be comparable to the wild-type pSSVx. Hence this genetic system is suitable to overcome the low efficiency typical of the common transformation techniques of *Sulfolobus* (10^3-10^4 cells/μg of plasmidic DNA used).The efficient propagation of the pMSSV shuttle vector could be traced back to the partial viral nature of the pSSVx, which was inherited by the vector.

Another important feature of the pMSSV shuttle vector lies on its capacity to steadily remain in episomal form inside host cells. The vector did not suffer either integration into the host chromosome or rearrangements and resided with undetectable to 130 copies per chromosome, the number depending on the cell growth phase. Thanks to this feature, the expression level of the passenger genes would not be affected by the decrease of the copy number which may result from the integration of the vector into the host chromosome.

The availability of this new transformation system based on SSV2 and the engineered pSSVx will contribute to clarify the mechanisms responsible for the pSSVx/SSV2 dependence as well as for replication, gene regulation and packaging of the episomal DNAs.

In order to improve the features of the pSSVx-based genetic system, we resolved to characterize the ORFs of the pSSVx, since most of them have unknown function. While the characterization of ORF91, ORF76, ORF154 and ORF288 is still

underway, interesting results were obtained by the study of the ORF c68 protein. The *orf c68* gene is the only one with opposite orientation on the pSSVx genome; it has no homologs in the crenarchaeal family of pRN plasmids (Arnold H.P. *et al*., 19992). ORF c68 shares high similarity to ORFc56 of pSSVi (Wang *et al*. 2007) and also finds its homologues in the integrated elements of the *S.tokodaii* and *S.acidocaldarius* genomes. The biochemical characterization of a putative protein without any useful informations about its function *in vivo* appeared very tricking. Hence, we started with bioinformatic analyses of ORF c68 performed by a secondary structure prediction and CDD and PSI-BLAST searches. This analyses highlighted that the *orf* c68 encodes a protein containing a SpoVT/Abr-B like domain, typical of a bacterial family of transcription factors.
At an early stage of our experiments, the recombinant ORF c68 protein was obtained in *E.coli* mesophilic host with a 6His-tag fused to its C-terminal portion and it was purified to homogeneity. Since the presence of the six additional aminoacidic residues made the ORF c68 protein very unstable in solution, the tag was eliminated by using a new cloning strategy. The ORF c68 untagged protein was also purified to homogeneity and used for all the following experiments. The expression level of the ORF c68 *in vivo* was detectable only at late exponential growth phase of the natural host simultaneously to the induction of the pSSVx replication. The findings of the secondary structure prediction as well as the intrinsic properties of the protein (highly basic isoelectric point, very low molecular weight) encouraged us to investigate the DNA-binding capacity of the ORF c68 purified protein. We found that ORF c68 binds specifically its own promoter region, this suggesting a role as auto-regulator of its own expression. Foot printing and EMSA experiments have revealed that ORF c68 binds also a region encompassing the promoter region of ORF 60. This finding suggests an hypothetical role of ORF c68 also in modulating the expression of the contiguous *orf 60* gene. The existence of a relation in the transcription of the *orf c68* and *orf 60* genes is supported by the transcriptional analysis performed by Contursi *et al.* (2007): the *orf* c68 and *orf 60* belong to the same complex transcriptional network, involving also *orf* 91 and *orf 892 RepA* (encoding a putative replication protein) adjacent genes, probably involved in regulating the *rep* gene expression. Crystallographic studies on the ORF c68 are currently underway and it is expected that structural informations will corroborate our functional data. Further investigation on the *in vivo* function of this protein may contribute to shed light on the mechanisms of the transcriptional regulation and of plasmid-virus maintenance and propagation in *Archaea.*
From an evolutionary point of view, since the SpoVT/Abr-B like transcription factor gene is missing on other prokaryotic plasmids but is present on the genomes of pSSVi and of the viruses STSV1 and SIFV from *S. tengchongensis* and *S. islandicus*, respectively (Prangishvili D. *et al*., 2006), it is tempting to speculate that the acquisition by the pSSVx genome of *orf* c68 (as well as of the *orfs* 154-288) has been crucial for the development of a viral nature and for the ability to respond to diverse viral stimuli.
When a deeper characterization/optimization of the pMSSV will be completed, another interesting future perspective of this work will be the construction of a SSV2-derived vector usable in combination with pMSSV. Two-vectors system will support functional genomic and proteomic approaches to clarify metabolic pathways and, more in general, peculiar physiological processes in *S.solfataricus*.

REFERENCES

Aravalli, R. N. & Garrett, R. A. (1997). "**Shuttle vectors for hyperthermophilic archaea"** *Extremophiles* **1**, 183–191.

Arnold HP, She Q, Phan H, Stedman KM, Prangishvili D, Holz I, Kristjansson JK, Garrett RA, Zillig W (1999a) "**The genetic element pSSVx of the extremely thermophilic renarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus"** *Mol Microbiol* **34**, 217-226

Arnold HP, Stedman KM, Zillig W (1999b) **"Archaeal phages"**. P76-89. In R. G. Webster A. Granoff (ed.), Encyclopedia of Virology. Academic Press, London

Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T., Weidmann, J.F., Kristjanson, J.K., Klenk, H.P., Nelson, K.E., Fraser, C.M. (2000). "**A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus"*** *Virology* 267, 252-266.

Aucelli T., Contursi P., Girfoglio M., Rossi M. and Cannio R. (2006) **"A spreadable, non-integrative and high copy number shuttle vector for *Sulfolobus solfataricus* based on the genetic element pSSVx from *Sulfolobus islandicus*"** *Nucleic Acid Research* **34**, 17

Bartolucci,S., Rossi,M. and Cannio,R. (2003) "**Characterization and Functional complementation of a nonlethal deletion in the chromosome of a beta-glycosidase mutant of *Sulfolobus solfataricus"*** *J. Bacteriol.*, **185**, 3948-3957.

Bell, S.D., and Jackson, S.P. (2000) "**Mechanism of autoregulation by an archaeal transcriptional repressor"** *J Biol Chem* **275,** 31624-31629.

Bettstetter, M., Peng, X., Garrett, R.A., Prangishvili, D. (2003). "**AFV1, a novel virus infecting hyperthermophilic archaea of the genus acidianus"** *Virology* **315**, 68-79.

Blum H, Zillig W, Mallock S, Domday H, Prangishvili D (2001) "**The genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses"**. *Virology* **281**, 6- 9

Brinkman, A.B., Dahlke, I., Tuininga, J.E., Lammers, T., Dumay, V., de Heus, E., *et al.* (2000) "**An Lrp-like transcriptional regulator from the archaeon *Pyrococcus furiosus* is negatively autoregulated"** *J Biol Chem* **275,** 38160-38169.

Brinkman, A.B., Ettema, T.J.G., de Vos, W.M., and van der Oost, J. (2003) **"The Lrp family of transcriptional regulators"** *Mol Microbiol* **48,** 287-294.

Brown, J.R., Doolittle, W.F. (1997). "**Archaea and the prokaryote-to-eukaryote transition"** *Microbiol. Mol. Biol. Rev*. **61**, 456-502.

Brugger, K., Redder, P., She, Q., Confalonieri, F., Zivanovic, Y. & Garrett, R. A. (2002). "**Mobile elements in archaeal genomes".** *FEMS Microbiol Lett* **206**, 131-141.

Byeon, W.H., and Weisblum, B. (1990) "**Replication genes of plasmid pE194 - cop and repF transcripts and encoded proteins"** *J Bacteriol* **172**, 5892-5900.

Cannio, R., Contursi, P., Rossi, M. & Bartolucci, S. (1998) **"An autonomously replicating transforming vector for *Sulfolobus solfataricus*"** *J Bacteriol* **180**, 3237-3240

Cannio,R., Contursi,P., Rossi,M. and Bartolucci,S. (2001) "**Thermoadaptation of a mesophilic hygromicin B phosphotransferase by directed evolution in hyperthermophilic Archaea: selection of a stable genetic marker for DNA transfer into *Sulfolobus solfataricus"*** *Extremophiles* **5,** 153-159.

Casjens, S.R., (2005). "**Comparative genomics and evolution of the tailed bacteriophages"** *Curr. Opin. Microbiol.* **8**, 451-458.

Coles, M., Djuranovic, S., Soding, J., Frickey, T., Koretke, K., Truffault, V., Martin, J., Lupas, A.N. (2005). **"AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi beta barrels"** *Structure (Camb)* **13**, 919-928.

Coles, M.H., Hulko, M, Djuranovic, S. Truffault, V. Koretke, K. Martin, J.(2006) "**Common Evolutionary Origin of Swapped-Hairpin and Double-Psi beta Barrels**" *Structure* **14**, 1489-1498

Contursi P., Jensen S., Aucelli T., Rossi M., Bartolucci S. and She Q. (2007) **"Characterization of the *Sulfolobus* host-SSV2 virus interaction"** *Extremophiles*

Contursi, P., Cannio, R., Prato, S., She, Q., Rossi M. and Bartolucci S. (2007) **"Transcriptional analysis of the genetic element pSSVx: differential and temporal regulation of gene expression reveals correlation between transcription and replication"** *J Bacteriol.* **17**,6339-50.

Cordes, M.H., Walsh, N.P., McKnight, C.J., Sauer, R.T. (1999). "**Evolution of a protein fold *in vitro***" *Science* **284**, 325-328.

Cubellis,M.V., Rozzo,C., Monteucchi,P. and Rossi,M. (1990) "**Isolation and sequencing of a new beta-glycosidase-encoding archaebacterial gene**" *Gene*, **94**, 89-94.

Dong, T. C., S. M. Cutting, and R. J. Lewis. (2004). "**DNA-binding studies on the *Bacillus subtilis* transcriptional regulator and AbrB homologue, SpoVT"** *FEMS Microbiol. Lett.* **233,** 247-256.

Elferink, M. G., Schleper, C. & Zillig, W. (1996). "**Transformation of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus* via a self-spreading vector"** *FEMS Microbiol Lett* **137**, 31-35.

Enoru-Eta, J., Gigot, D., Glansdorff, N., and Charlier, D. (2002) "**High resolution contact probing of the Lrp-like DNAbinding protein Ss-Lrp from the hyperthermoacidophilic crenarchaeote *Sulfolobus solfataricus* P2"***Mol Microbiol* **45**,1541–1555.

Gaël Erauso, Kenneth M. Stedman, Harmen J. G. van de Werken, Wolfram Zillig, and John van der Oost (2004) **"Two novel conjugative plasmids from a single strain of *Sulfolobus*"** *Microbiology* **152**, 1951-1968

Esposito D, Scocca JJ (1997) "**The integrase family of tyrosine recombinases: evolution of a conserved active site domain"** *Nucleic Acids Res* **25**,3605-3614

Ettema, T.J.G., Brinkman, A.B., Tani, T.H., Rafferty, J.B., and van der Oost, J. (2002) "**A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes"***J Biol Chem* **277,** 37464-37468.

Garrett, R.A., P. Redder, B. Greve, K. Brügger, L. Chen and Q. She. 2004. **"Archaeal plasmids"**. *In* Plasmid Biology. Eds. B.E. Funnell and G.J. Phillips. A.S.M. Press, Washington, DC, pp 377–392.

Gomis-Ruth, F.X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., Gonzalez, A., Espinosa, M., del Solar, G., Coll, M. (1998). **"The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator"** *EMBO J*. **17**, 7404-7415.

Greve, B., Jensen, G. B., Brügger, K., Zillig, W. & Garrett, R. (2004). "**Genomic comparison of archaeal plasmids from *Sulfolobus*"** *Archaea* **1**, 231-239.

Grogan D, Palm P, Zillig W (1990) **"Isolate B12, which harbours a virus like element, represents a new species of the archaebacterial genus *Sulfolobus*, *Sulfolobus shibatae*, sp. nov."** *Arch Microbiol* **154**, 594-599

Grogan,D.W. (1991) **"Evidence that beta-galactosidase of Sulfolobus solfataricus is only one of several activities of a thermostable beta-D-glycosidase"** *Appl. Environ. Microbiol.* **57**, 1644-1649.

Haring, M., Peng, X., Brugger, K., Rachel, R., Stetter, K.O., Garrett, R.A., Prangishvili, D. (2004). "**Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus*: a novel virus family, the Globuloviridae"** *Virology* **323**, 233-242.

Haring, M., Rachel, R., Peng, X., Garrett, R.A., Prangishvili, D. (2005a). "**Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, *Acidianus* bottle-shaped virus, from a new family, the Ampullaviridae"** *J. Virol.* **79**, 9904-9911.

Haring, M., Vestergaard, G., Brugger, K., Rachel, R., Garrett, R.A., Prangishvili, D. (2005b). **"Structure and genome organization of AFV2, a novel archaeal lipothrixvirus with unusual terminal and core structures"** *J. Bacteriol.* **187**, 3855-3858.

Haring, M., Vestergaard, G., Rachel, R., Chen, L., Garrett, R.A., Prangishvili, D. (2005c).“**Virology: independent virus development outside a host”** *Nature* **436**,1101-1102.

Hendrix, R.W., Smith, M.C., Burns, R.N., Ford, M.E., Hatfull, G.F. (1999). **“Evolutionary relationships among diverse bacteriophages and prophages: all the world’s a phage”** *Proc. Natl. Acad. Sci.* U.S.A. **96**, 2192-2197.

Janekovic, D., Wunderl, S., Holz, I., Zillig, W., Gierl, A., Neumann, H. (1983). “**TTV1, TTV2 and TTV3, a family of viruses of the extremely thermophilic, anaerobic sulfur reducing archaebacterium *Thermoproteus tenax*”** *Mol. Gen. Genet.* **192**, 39-45.

Jonuscheit, M., Martusewitsch, E., Stedman, K. and Schleper, C. (2003) “**A reporter gene system for the hypertermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector**” *Mol. Microbiol.,* **48**, 1241-1252.

Kessler A., Sezonov G., Guijarro J. I., Desnoues N., Rose T., Delepierre M., Bell S. and Prangishvili D. (2006) **A novel archaeal regulatory protein, Sta1, activates transcription from viral promoters** *Nucleic Acids Research* **0,** 1-9

Kletzin A., Lieke A., Urich T., Charlebois R.L. and Sensen C.W. (1999) **“Molecular analysis of pDL10 from *Acidianus ambivalens* from extremely thermophilic and acidophilic Archaea”** *Genetics* **152**, 1307-1314.

Kyrpides N.C. and C.A. Ouzounis (1999) **“Transcription in Archaea”** *Proc. Natl. Acad. Sci.* USA **96**, 8545-8550.

Leonard, P.M., Smiths, S.H.J., Sedelnikova, S.E., Brinkman, A.B., de Vos, W.M., van der Oost, J., *et al*. (2001) “**Crystal structure of the Lrp-like transcriptional regulator from the archaeon *Pyrococcus furiosus*”** *EMBO J* **20,** 990-997.

Lindquist, B.H., Deho, G., and Calendar, R. (1993) “**Mechanisms of genome propagation and helper exploitation by satellite phage P4”**.*Microbiol Rev* **57**, 683-702.

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.

Mann, B.A., and Slauch, J.M. (1997) **"Transduction of low-copy number plasmids by bacteriophage P22”** *Genetics* **146**, 447-456.

Martin A, Yeats S, Janekovic D, Reiter WD, Aicher W, Zillig W (1984) **“SAV1, a temperate U. V. inducible DNA virus like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12”** *EMBO J* **3**, 2165-2168

Martusewitsch, E., Sensen, C. W. & Schleper, C. (2000) **"High spontaneous mutation rate in the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by transposable elements"** *J Bacteriol* **182**, 2574-2581

Nadal M, Mirambeau G, Forterre P, Reiter WD, Duguet M (1986) **"Positively supercoiled DNA in a virus-like particle of an archaebacterium"**.*Nature* **321**,256-258

Nunes-Duby SE, Kwon HJ, Tirumalai RS, Ellenberger T, Landy A (1998) **"Similarities and differences among 105 members of the Int family of site-specific recombinases"** *Nucleic Acids Res* **26**, 391-406

Ouhammouch, M., Dewhurst, R.E., Hausner, W., Thomm, M., and Geiduschek, E.P. (2003) **"Activation of archaeal transcription by recruitment of the TATA binding protein"** *Proc Natl Acad Sci USA* **100,** 5097-5102.

Palm P, Schleper C, Grampp B, Yeats S, McWilliam P, Reiter WD, Zillig W (1991) **"Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae"*** *Virology* **185**, 242-250

Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., & 16 other authors (2003). **"Origins of highly mosaic mycobacteriophage genomes"** *Cell* **113**, 171-182.

Peng,X., Holz,I., Zillig,W., Garrett,R.A. and She,Q. (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.

Peng X, Blum H, She Q, Mallok S, Brügger K, Garrett RA, Zillig W, Prangishvili D (2001) **"Sequences and replication of genomes of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses"** *Virology* **291**, 226-234

Perez-Rueda, E., Collado-Vides, J., Segovia, L., (2004) **"Phylogenetic distribution of DNA-binding transcription factors in Bacteria and Archaea"** *Comput. Biol. Chem.* **28**, 341-350.

Pisani,F.M., Rella,R., Raia,C.A., Rozzo,C., Nucci,R., Gambacorta,A., De Rosa,M. and Rossi,M. (1990) **Thermostable beta-galactosidase from the archaebacterium Sulfolobus solfataricus. Purification and properties.** *Eur. J. Biochem*. **187**, 321-328.

Prangishvili, D., S. V. Albers, I. Holz, H. P. Arnold, K. Stedman, T. Klein, H. Singh, J. Hiort, A. Schweier, J. K. Kristjansson, and W. Zillig. (1998) **"Conjugation in Archaea—frequent occurrence of conjugative plasmids in *Sulfolobus*"** *Plasmid* **40**,190-202.

Prangishvili, D., Stedman, K. & Zillig, W. (2001) **"Viruses of the extremely thermophilic archaeon *Sulfolobus*"** *Trends Microbiol* **9**, 39-43.

Prangishvili D, Garrett RA, Koonin EV (2006a) **"Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life"** *Virus Res* **117**, 52-67.

Prangishvili D, Vestergaard G, Haring M, Aramayo R, Basta T, Rachel R, Garrett RA (2006b) **"Structural and Genomic Properties of the Hyperthermophilic Archaeal Virus ATV with an Extracellular Stage of the Reproductive Cycle"** *J Mol Biol* 2006 Apr 27 pub ahead.

Reilly, M. S. & Grogan, D. W. (2001) **"Characterization of intragenic recombination in a hyperthermophilic archaeon via conjugational DNA exchange"***J Bacteriol* **183**, 2943-2946.

Reiter WD, Palm P, Yeats S, Zillig W (1987) **"Gene expression in archaebacteria: physical mapping of constitutive and UV-inducible transcripts from the *Sulfolobus* virus-like particle SSV1"** *Mol Gen Genet* **209**, 270-275

Reiter,W.D., Zillig,W. and Palm,P.(1988) **"Archaebacterial viruses"** *Adv. Virus Res.* **34**, 143-188.

Reiter WD, Palm P (1990) **"Identification and characterization of a defective SSV1 genome integrated into a tRNA gene in the archaebacterium *Sulfolobus sp.* B12"** *Mol Gen Genet* **221**, 65-71

Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S., McDermott, T. & Young, M. J. (2001) **"Viruses from extreme thermal environments"** *Proc Natl Acad Sci* USA **98**, 13341-13345.

Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J.E., Douglas, T., Young, M. (2004) **"The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life"** *Proc. Natl. Acad. Sci.* USA **101**, 7716-7720.

Schleper, C., Kubo, K. & Zillig, W. (1992) **"The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA"** *Proc Natl Acad Sci* U S A **89**, 7645-7649

Schleper,C., Roder,R., Singer,T. and Zillig,W. (1994) **"An insertion element of the extremely thermophilic archaeon Sulfolobus solfataricus transposes into the endogenous β-galactosidase gene"** *Mol. Gen. Genet*., **243**, 91-96.

Schleper, C., Holz, I., Janekovic, D., Murphy, J. & Zillig, W. (1995) **"A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating"** *J Bacteriol* **177**, 4417-4426.

She, Q., Phan, H., Garrett, R. A., Albers, S. V., Stedman, K. M. & Zillig, W. (1998) **"Genetic profile of pNOB8 from *Sulfolobus*: the first conjugative plasmid from an archaeon"** *Extremophiles* **8**, 417-425.

She,Q., Brugger,K. and Chen,L. (2002) **"Archaeal integrative genetic elements and their impact on genome evolution"** *Res Microbiol.* **153**, 325-332.

Snyder JC, Stedman K, Rice G, Wiedenheft B, Spuhler J, Young MJ (2003) **"Viruses ofhyperthermophilic Archaea"** *Res Microbiol* **154**, 474-482

Somers, W.S., Phillips, S.E. (1992). **"Crystal structure of the met repressor operator complex at 2.8 A resolution reveals DNA recognition by beta-strands"** *Nature* **359**, 387-393.

Stedman, K. M., Schleper, C., Rumpf, E. & Zillig, W. (1999) **"Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors"** *Genetics* **152**, 1397-1405.

Stedman, K. M., She, Q., Phan, H., Holz, I., Singh, H., Prangishvili, D., Garrett, R. & Zillig, W. (2000). **"pING family of conjugative plasmids from the extremely thermophilic archaeon *Sulfolobus islandicus*: insights into recombination and conjugation in *Crenarchaeota*"** *J. Bacteriol* **182**, 7014-7020.

Stedman KM, She Q, Phan H, Arnold HP, Holz I, Garrett RA, Zillig W (2003) **"Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2"** *Res Microbiol* **154**, 295-302

Strauch, M.A. (1995) **"Delineation of AbrB-binding sites on the *Bacillus subtilis* spo0H, kinB, ftsAZ, and pbpE promoters and use of a derived homology to identify a previously unsuspected binding site in the bsuB1 methylase promoter"***J. Bacteriol.* **177**, 6999–7002.

Vaughn, J.L., Feher, V., Naylor, S., Strauch, M.A., and Cavanagh, J. (2000). **"Novel DNA-binding domain and genetic regulation model of *Bacillus subtilis* transition state regulator AbrB"** *Nat. Struct. Biol. 7*

Wang, Y., Z. Duan, H. Zhu, X. Guo, Z. Wang, J. Zhou, Q. She, and L. Huang. (2007) **"A Novel *Sulfolobus* non-conjugative extrachromosomal genetic element capable of integration into the host genome and spreading in the presence of a fusellovirus"** *Virology* doi:10.1016/j.virol.2007.01.035.

Werner F. and Weinzierl. R.O. (2002) **"A recombinant RNA polymerase II-like enzyme capable of promoter-specific transcription"** *Mol.Cell* **10**, 635-646.

Wiedenheft B, Stedman KM, Roberto F, Willits D, Gleske AK, Zoeller L, Snyder J, Douglas T, Young M (2004) **"Comparative genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses"** *J Virol* **78**, 1954-1961

Worthington,P., Hoang,V., Perez-Pomares,F. and Blum,P. (2003) **"Targeted disruption of the alpha-amylase gene in the hyperthermophilic archaeon *Sulfolobus solfataricus J. Bacteriol"* 185**, 482-488.

Xiang, X., Chen, L., Huang, X., Luo, Y., She, Q., Huang, L. (2005) **"*Sulfolobus tengchongensis* spindle-shaped virus STSV1: virus-host interactions and genomic features"**J. Virol. **79**, 8677-8686.

Yeats S, McWilliam P, Zillig W (1982) **“A plasmid in the archaebacterium *Sulfolobus acidocaldarius”*** *EMBO J* **1**, 1035-1038

Zuber P. (2000) **“Specificity through flexibility”** *Nat. Struct. Biol. 7*

Zillig, W., Stetter, K. O., Wunderl, S., Schulz, W., Priess, H. & Scholz, I. (1980) **“The *Sulfolobus* -"*Caldariella*" group: taxonomy on the basis of the structure of DNA-dependent RNA polymerases”** *Arch Microbiol* **125**, 259-269.

Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janecovik, D., Hain, J., Lanzendorfer, M. & Kristjansson, J. (1994) **“Screening for Sulfolobales, their plasmids and viruses in Icelandic solfataras”** *Syst Appl Microbiol* **16**, 609-628.

Zillig W., Prangishvilli D., Schlepr C., Elferink M., Albers S., Janekovic D., Gotz D. (1996) **“Viruses, plasmids and other genetic elements of thermofilic and hypertermofilic *Archaea*”** *FEMS Microbiol. Rev***. 18**, 225-236

Zillig, W., Arnold, H. P., Holz, I. & 7 other authors (1998) **“Genetic elements in the extremely thermophilic archaeon *Sulfolobus*”** *Extremophiles* **2**, 131-140.

PUBLICATIONS

Tiziana Aucelli, Patrizia Contursi, Michele Girfoglio, Mosè Rossi and Raffaele Cannio **"A spreadable, non-integrative and high copy number shuttle vector for *Sulfolobus solfataricus* based on the genetic element pSSVx from *Sulfolobus islandicus"*** *Nucleic Acids Research* (2006)

Patrizia Contursi, Susanne Jensen, Tiziana Aucelli, Mosè Rossi, Simonetta Bartolucci and Qunxin She **"Characterization of the Sulfolobus host-SSV2 virus interaction"** *Extremophiles* (2006)

COMUNICATIONS

Tiziana Aucelli, Patrizia Contursi, Mosè Rossi e Raffaele Cannio **"Spreadable and high copy number shuttle vectors for Sulfolobus solfataricus based on the genetic elements pSSVx and SSV2 from Sulfolobus islandicus"** *International Conference on Extremophiles, 17-21 Sept. 2006 Brest, France.*

Tiziana Aucelli, Patrizia Contursi, Emilia Pedone, Raffaele Cannio, Mosè Rossi and Simonetta Bartolucci **"Structural-functional analysis of the gene product ORF c68 of the plasmid-virus hybrid pSSVx from *Sulfolobus islandicus* REY 15/4"** *International Conference on Thermophiles, 24-27 Sept. 2007 Bergen, Norway.*

Ringraziamenti

Nessuno di noi riesce a conseguire un successo se non è attorniato dalle persone giuste. E' pertanto doveroso da parte mia rivolgere un sentito ringraziamento a tutti coloro che hanno scelto di starmi accanto, soprattutto quando le circostanze non si sono rivelate delle più favorevoli.

Ringrazio la Prof.ssa Simonetta Bartolucci per avermi accolta nel suo gruppo di ricerca e supportata durante questi tre anni di lavoro.

Una indiscussa e netta predominanza di merito per la comprensione e lo smisurato amore dimostratimi durante questo mio percorso di vita spetta innanzitutto ai miei Genitori ed a mio fratello Marco, nonché al resto della mia preziosa famiglia.

Sento, inoltre, di rivolgere al mio Andrea un simbolico "grazie" per avermi incessantemente dimostrato quanto sostanziale sia la presenza di un giusto compagno…..sin dai tempi del Liceo!

Desidero che tutta la mia gratitudine vada alla Dott.ssa Rossella De Mizio per avermi insegnato quanto bello possa essere che qualcuno apprenda da te e grazie a questo "cresca", nonostante le avversità, nell'amore per la Scienza…….e per essere divenuta mia amica.

Un ringraziamento speciale alla Dott.ssa Ilaria Galdi, al Dott. Raffaele Ronca e al fantastico "binomio" di quasi-dottori Anna Vitale e Ilario Cuozzo per le preziose qualità umane e la grande ricchezza interiore che li caratterizza.

Un affettuoso grazie anche alle nuove "reclute" del gruppo: Gaetano, Valentina, Raffaella ed Imma……semplicemente per quello che sono !

Non posso esimermi, infine, dal tributare al Dott. Raffaele Cannio una profonda riconoscenza per essere stato il mio prezioso mentore sin dai tempi della Tesi di Laurea con le sue ineguagliabili doti scientifiche ed umane.

Published online 13 September 2006 *Nucleic Acids Research, 2006, Vol. 34, No. 17* ***e114***
doi:10.1093/nar/gkl615

A spreadable, non-integrative and high copy number shuttle vector for *Sulfolobus solfataricus* based on the genetic element pSSVx from *Sulfolobus islandicus*

Tiziana Aucelli[1], Patrizia Contursi[1], Michele Girfoglio, Mosè Rossi and Raffaele Cannio*

Istituto di Biochimica delle Proteine, Via P. Castellino 111, 80131 CNR, Napoli, Italy and [1]Dipartimento di Biologia Strutturale e Funzionale, Università degli Studi di Napoli Federico II, Via Cinthia, 80126 Napoli, Italy

Received July 5, 2006; Revised and Accepted August 5, 2006

ABSTRACT

The pSSVx genetic element from *Sulfolobus islandicus* REY15/4 is a hybrid between a plasmid and a fusellovirus, able to be maintained in non-integrative form and to spread when the helper SSV2 virus is present in the cells. In this work, the satellite virus was engineered to obtain an *Escherichia coli–Sulfolobus solfataricus* shuttle vector for gene transfer and expression in *S.solfataricus* by fusing site-specifically the pSSVx chromosome with an *E.coli* plasmid replicon and the ampicillin resistance gene. The pSSVx-based vector was proven functional like the parental virus, namely it was able to spread efficiently through infected *S.solfataricus* cells. Moreover, the hybrid plasmid stably transformed *S.solfataricus* and propagated with no rearrangement, recombination or integration into the host chromosome. The high copy number of the artificial genetic element was found comparable with that calculated for the wild-type pSSVx in the new host cells, with no need of genetic markers for vector maintenance in the cells and for transfomant enrichment.

The newly constructed vector was also shown to be an efficient cloning vehicle for the expression of passenger genes in *S.solfataricus*. In fact, a derivative plasmid carrying an expression cassette of the lacS gene encoding the β-glycosidase from *S.solfataricus* under the control of the *Sulfolobus* chaperonine (thermosome tf55) heat shock promoter was also able to drive the expression of a functional enzyme. Complementation of the β-galactosidase deficiency in a deletion mutant strain of *S.solfataricus* demonstrated that *lacS* gene was an efficient marker for selection of single transformants on solid minimal lactose medium.

INTRODUCTION

Host/virus interaction modes have provided windows to study microbial diversity (1) as well as genetic processes at the molecular level, in particular for prokaryotes, and hence have helped in clarifying the physiological mechanisms, the dependence on the specific biochemical environment and evolution of their host cells (2,3).

Very few viruses have been identified from Archaea (4) as compared with Bacteria and Eukarya and detailed description has been reported for those from hyperthermophilic archaea (5,6) with representatives that replicate in the genus *Sulfolobus* being the majority within the kingdom Crenarchaeota (6–8). To date, the *Fuselloviridae* are the most widespread on earth in the *Sulfolobus* genus with viruses sharing similar morphology as well as DNA genome size and organization (9–12).

Sulfolobus spindle-shaped virus 1 (SSV1) is the best studied member of this family and demonstrated to be temperate both in *Sulfolobus shibatae* and in non-natural but related *Sulfolobus* hosts, such as *Sulfolobus solfataricus* (13,14); infection, integration of DNA into the host chromosome and production of virions cause apparently no phenotype change but a significant growth retardation of the host cells which can be visualized as turbid plaques around propagation foci on plated lawns of indicator host cells (14–18).

More recently, another fusellovirus, SSV2 from *Sulfolobus islandicus* strain 15/4 was isolated, characterized and its complete genomic sequence determined. SSV2 shares with SSV1 similar morphology, replication and DNA size (19). The overall genome architecture is conserved but the low similarity in the sequences should be responsible for the higher copy number and the lack of a strong ultraviolet induction of episomal SSV2 DNA and particle production, as well as for the different integration of the SSV2 genome which occurs into the host chromosome at the site of a glycyl tRNA instead of arginyl-tRNA (12).

S.islandicus REY15/4 harbors also a small plasmid, pSSVx, assigned to the pRN family (20,21) of *Sulfolobales* plasmids; pSSVx is also capable of spreading in the cell cultures of

*To whom correspondence should be addressed at Istituto di Biochimica delle Proteine, Consiglio Nazionale delle Ricerche, Via Pietro Castellino 111, 80131, Naples, Italy. Tel: +39 081 613 2285; Fax: +39 081 613 2248; Email: r.cannio@ibp.cnr.it

S.solfataricus but only in the presence of either SSV2 or SSV1, necessary as helpers (9). In fact, pSSVx contains two open reading frames showing high-sequence similarity to a tandem of ORFs in both SSV1 and SSV2 genomes; the proteins encoded by these ORFs are probably necessary for specific recognition of the pSSVx DNA but need viral helper components for capsid formation and packaging (9,19).

In general, the choice of *S.solfataricus* as a model for fundamental understanding of the genetics of extremely thermophilic archaea is due to growth conditions operatively non-prohibitive (22) and capability of maintaining and propagating either natural or genetically modified extrachromosomal DNAs (23,24) from other sources. The complete genome of *S.solfataricus* has also been determined (25), the biochemical characterization of many gene products obtained (26) and the development of post-genomics tools such as proteomics and metabolic pathway reconstruction recently attempted (27,28). Some progress has been made to develop stable transformation (29–32), specific gene disruption methods (33) as well as overexpression of foreign and homologous genes (34); nevertheless none of the systems described so far has been proven efficient for reproducibility and stability of gene cloning and protein expression levels in *Sulfolobus*, probably due, in most cases, to low transformation efficiencies, inefficient selection and/or instability of the vectors in the host as well as changes in the ratio episomal versus integrative forms occurring in the cell during replication of virus-derived constructs (35).

In this study, a genetic system for *Sulfolobus* was developed that is based on the satellite virus pSSVx from *S. islandicus* 15/4. The different recombinant *Escherichia coli–Sulfolobus solfataricus* shuttle vectors constructed retained the wild-type capability to replicate at high copy-number and to spread in cell cultures in the presence of its helper virus SSV2. *Sulfolobus* transformants were demonstrated to be stable and propagate the pSSVx derived plasmids in a reproducible and constant fashion without any rearrangement, recombination or integration into the chromosome.

Moreover, stable complementation of a β-galactosidase mutant of *S.solfataricus* previously isolated and characterized in our laboratory (32) and reproducible gene expression levels were also obtained by introducing the β-galactosidase gene (*lacS*) as a reporter under the control of a strong and heat-inducible promoter into the shuttle vector.

MATERIALS AND METHODS

Growth of *Sulfolobus* strains and isolation of SSV2-infected *S.solfataricus* GθW and P2

S.solfataricus strains P2 (DSM 1617), Gθ (23) and the derivative mutant GθW [Δ*lacS*, (32)] as well as *S.islandicus* REY 15/4 (22) were grown at 75 or 80°C in glycine buffered Brock's medium (36) with 0.1% tryptone, 0.05% yeast extract and 0.2% sucrose at pH 3.2. For electroporation and plaque assays, cells were grown with phosphate buffered medium N.182 (M182) suggested by the DSMZ Catalogue of strains containing 0.1% glucose. For isolation of independent clones, medium M182 contained only 0.25% lactose and no other nutrient. The optical density of liquid cultures was monitored at 600 nm. For solid media, gellan gum (Gelrite, Sigma) was added to a final concentration of 0.8% (0.35% for overlays), and $MgCl_2$ and $CaCl_2$ were added to 10.0 and 3.0 mM, respectively.

The *S.solfataricus* strains GθW and P2 transfected with the SSV2 and pSSVx were extracted from zones of growth inhibition (plaques) formed on indicator lawns around spots of the *S.islandicus* REY 15/4 culture supernatants (2–6 μl for each spot), as described by Arnold *et al*. (9). Clones cured selectively for pSSVx were obtained by isolation of single colonies formed on plates by the cells extracted from plaques and revitalized in liquid cultures.

Viral DNA isolation and plasmid constructions

Extrachromosomal SSV2 and pSSVx DNAs from both *S.islandicus* REY 15/4 and transfectants of *S.solfataricus* strains were performed with Qiaprep Spin Miniprep kit (Qiagen) following the standard procedure suggested by the manufacturer for *E.coli* cells.

Plasmid pSSVrt was constructed by cloning the pSSVx DNA linearized at the position 2812 with AflIII and modified with Klenow DNA polymerase into the SmaI site of the pUC19 *E.coli* vector. Linearization was obtained by AflIII partial digestion in the presence of ethidium bromide using a protocol already described for the SSV1 virus DNA linearized with Sau3AI (37). Clones with insertion at the specific position were selected by restriction analysis of the resulting plasmid collection. A size reduced derivative of pSSVrt was also constructed; the polycloning sequence (between the AatII and EcoRI sites) of the pUC28 vector was inserted into a 1812 bp AatII/AflIII DNA fragment from pUC18 after suitable modification of incompatible ends. The pSSVx sequence was excised from the pSSVrt plasmid with SacI and PstI and inserted into the same sites of the minimal plasmid obtained to produce the pMSSV vector. An expression cassette of the *lacS* gene (38) was PCR amplified by the vector pMJ03 (35) and inserted between the XhoI and PstI sites of pMSSV generating the expression vector pMSSV*lacS*. Excision of the *E.coli* minimal plasmid was obtained by digestion of pMSSV*lacS* (1.0 μg) with SacI and purification of the pSSVx/*lacS* moiety from agarose gel. The DNA was re-circularized by ligation in a final volume of 5.0 μl, diluted in water to 20 ng/μl final concentration, and 2 μl were used for electroporation.

Transformation procedure and analyses of *Sulfolobus* transformants

S.solfataricus cells of SSV2 lysogenes were grown up to mid-logarithmic phase (0.3–0.45 OD_{600}), harvested by centrifugation and repeatedly washed in 20 mM sucrose as described previously by Schleper *et al*. (14). Aliquots of 10^{10} cells/ml (50 μl) were mixed with 1 μl DNA (10–100 ng/μl), incubated for 1 min on ice and transferred to chilled plastic cuvettes with an electrode gap of 0.1 cm (BioRad). High voltage electroporation (25 μF) was performed with a BioRad Gene Pulser Xcell™ at a field strength of 1.5 kV/cm and 400 Ω resistance; two successive shock pulses were applied to competent cells producing pulse length of ∼10.0 and 9.8 ms, respectively. Immediately after electroporation cell mixtures were diluted with 1 ml of medium M182 containing 0.1%

glucose, transferred to glass vials and incubated for 3 h at 75 or 80°C. After suitable scale-up, 5–15 ml aliquots of the cultures (the volumes varying in order to withdraw the same number of cells per aliquot) were harvested at increasing cell density for DNA extractions. For monitoring propagation of the recombinant satellite virus, extrachromosomal DNA mini preparations and plaque assays were performed. Transformants were stored at −80°C in 15% glycerol stocks.

For Southern blot analysis, ∼2 μg of total cellular DNAs, extracted according to Arnold *et al*. (9), and 5 ng of pMSSV*lacS* plasmid purified from *E.coli*, were cut with HindIII and BglII, and electrophoresed in a 0.8% agarose gel; DNA digests were blotted and hybridized according to standard procedures (39). The probe was prepared by cutting out and purifying a HindIII restriction fragment from the pMSSV*lacS* vector encompassing the *lacS* gene and a portion (up to −302) of the tf55α 5′ flanking region. This restriction DNA fragment was randomly labeled using the random prime DNA labeling kit (Boehringer Mannheim).

β-galactosidase complementation and isolation of mixed and single transformants

pMSSV*lacS* transformed cells were tested for β-galactosidase activity; 1 ml aliquots of cultures were centrifuged and cell pellets overlaid with a X-Gal solution (2 mg/ml in phosphate buffered medium) and incubated for 15 min at 75°C for blue color development. *In situ* assays were performed on the same cultures seeded on plates and grown as circular colonized areas, as previously described (32). Supernatants of cultures grown up to 0.5, 1.0 and 1.3 OD_{600} were checked for plaque formation as already described and analyzed by X-Gal staining of plaques formed on continuous lawns of the *S.solfataricus* GθW strain. For infection in liquid culture, 400 μl supernatants from pMSSV*lacS* transformants were added to a 20 ml culture of cells transfected only with SSV2 and grown up to 0.5 OD_{600}. After incubation under shaking for 48 h, cells were diluted 1:50, grown up to 0.36 and 1.0 OD_{600} and tested for β-galactosidase activity as already described.

Single transformants were selected either on rich (M182, glucose 0.1%) or on minimal (lactose 0.25% as the only nutrient) solid media, by plating or by streaking of electropored cells propagated until they showed positive staining with X-Gal and diluted after they had reached different cell densities (0.3–1.0 OD_{600}). Single clones formed after 2 weeks of incubation at 75°C were stained with X-Gal (incubation at 75°C for 3–4 h for color development on lactose plates), picked and resuspended in 100 μl lactose minimal medium and seeded as spots onto fresh rich (M182) medium plates. After *in situ* X-Gal test on plates, cell spots were transferred to M182 medium containing glucose, propagated up to early stationary phase (1.0 OD_{600}) and analyzed for extrachromosomal DNA content as described above.

β-galactosidase enzyme activity

β-galactosidase activity of transformants was visualized and quantified as follows. Crude extracts were prepared by a freeze thaw technique, suspension of the cells in 10 mM Tris–HCl buffer, pH 8.0, placing at 80°C for 10 min, then at 50°C for 5 min. This procedure was repeated four times, and the extract was spun for 30 min at ∼10 000*g*. The supernatant was either assayed immediately or stored at −80°C before testing. Protein concentrations of the crude cell extracts were determined by the Bradford assay (Bio-Rad). Detection of the β-galacosidase in protein extracts was performed on 10% acrylamide SDS–PAGE gels after electrophoretic separation and extensive washing in 10 mM Tris–HCl buffer, pH 8.0. The specific enzyme band was visualized incubating the gel 30 min at 75°C in the same buffer containing 2 mg/ml X-gal.

The β-galactosidase assay procedure followed essentially the protocol of Pisani *et al*. (40). A sample of 10 μl extract was transferred to a preheated (75°C) quartz cuvette containing 990 μl assay buffer (2.8 mM ONPG in 50 mM sodium phosphate buffer, pH 6.5). The ONPG hydrolysis reaction was followed spectrophotometrically at 75°C by measuring the increase in absorbance at 405 nm in a Beckman spectrophotometer with heatable cuvettes. One unit was defined as the amount of enzyme catalyzing the hydrolysis of 1 μmol of ONPG min^{-1} at 75°C with a molar absorption coefficient of 3100 M^{-1} cm^{-1} at 405 nm for ONP.

RESULTS

Transfection and isolation of *Sulfolobus solfataricus* SSV2 transfectants

The strains P2 and GθW, a β-glycosidase defective mutant isolated earlier (32), were tested for susceptibility to infection by SSV2 and pSSVx viruses produced by *S.islandicus* REY15/4. Supernatants of REY15/4 cultures generated growth inhibition halos on continuous lawns of both *S.solfataricus* strains (Figure 1A), similar to those observed for the strain P1 (9); like for SSV1 virus (37), plaques did not form on *Sulfolobus acidocaldarius* cells, namely no infection occurred. Cells extracted from areas of these plaques and revitalized in liquid cultures contained extrachromosomal DNA indistinguishable from that isolated from REY15/4, indicating the presence and active replication of both virus and satellite elements (Figure 1B). From this culture, clones infected only by the SSV2 virus, namely cured for the pSSVx, were isolated as single colonies on plates after suitable dilutions, as described for the P1 strain by Arnold *et al*. (9). The absence of the satellite virus was confirmed by restriction analysis of extrachromosomal DNA (Figure 1C) and Southern hybridization (data not shown).

Shuttle vector construction

A sequence analysis on the pSSVx DNA was performed in order to locate regions that could be manipulated without affecting DNA replication and particle proliferation/spreading, and thus representing candidate targets for site-specific insertion of foreign DNA sequences. A 200 bp segment was identified which contains the tail-to-tail intergenic region between the ORFs C68 and 288. The segment shows archaeal transcription termination signals (41) and tendency to form hairpin loops as for rho-independent termination mechanisms (42,43). Moreover, it has an AflIII site (cut at the position 2812 on the pSSVx genome), useful for cloning, since it is situated in the 3′ direction beyond the ORF288

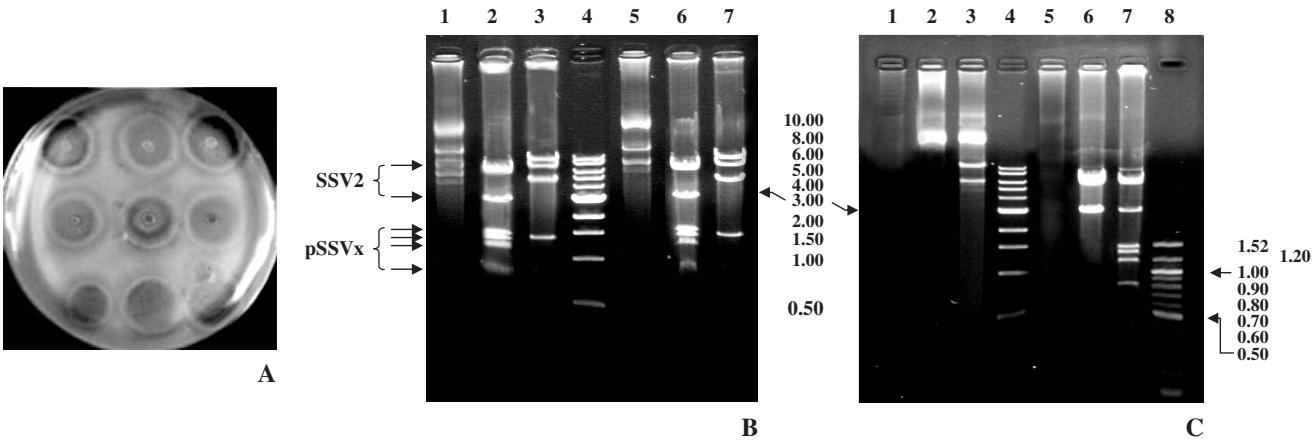

Figure 1. Transfection of *S.solfataricus* GθW with conditioned medium of *S.islandicus* REY15/4. The spontaneous mutant GθW lacking the β-glycosidase activity was first tested for infection with both SSV2 and pSSVx by formation of growth inhibition areas in cell lawns on plate. Virus plaques of SSV2 and pSSVx were obtained by spotting aliquots of REY 15/4 supernatants (2 µl for the three plaques in the first row, 4 µl for those in the second row and 6 µl for those in the third row) onto a continuous indicator lawn of the GθW cells (**A**). (**B**) Extrachromosomal DNAs (uncut in lanes 1 and 5), extracted from both the natural source *S.islandicus* (lanes 1–3) and the infected *S.solfataricus* (lanes 5–7), propagated after extraction from the plaques and revitalization, was cut with EcoRI (lanes 2 and 6) and BamHI (lanes 3 and 7). Episomal DNAs were confirmed to be identical and to contain both SSV2 and pSSVx DNAs, as demonstrated by the two distinct EcoRI restriction patterns. Lane 4: DNA molecular weight standards (kb). (**C**) The cells entrapped inside the plaques have been extracted from the plates, revitalized in liquid medium and then streaked onto solid medium to obtain single colonies. The comparative analysis with extrachromosomal DNAs from non-transformed recipient cells (lanes 1 and 5) and from the cells transfected with both viral and satellite DNAs (lanes 3 and 7) revealed the absence of the pSSVx plasmid in one colony isolated on plate and analyzed (lanes 2 and 6). Lanes 1–3: uncut DNAs; lanes 5–7: EcoRI digests; lanes 4 and 8: DNA molecular weight standards (kb).

transcription termination region, and far upstream of the ORF c68 stop codon. Since the AflIII site is present in five copies on the pSSVx sequence (positions 677, 814, 982, 2812 and 4994), singly cut pSSVx genomes were generated by digestion with the restriction enzyme under conditions that allowed single cleavage of the DNA molecules presumably at every specific site (37). After suitable modification of termini, these DNAs were inserted into *E.coli* pUC19 plasmid vector and specific insertion at the position 2812 produced the fusion plasmid pSSVrt (Figure 2).

pSSVrt shuttle vector transfection and spreading

The plasmid DNA pSSVrt selected from *E.coli* was transferred by electroporation into *Sulfolobus*, using different strains, namely SSV2 lysogens of the strains P2, MT4 and GθW, which is a stable β-galactosidase mutant with an extended deletion in the *lacS* genetic locus (32). To test transformation efficiency, *S.solfataricus* was transformed with varying amounts of the plasmid pSSVrt (10–100 ng) and then checked for the presence and amount of the vector at different cell densities and after several generations. Three hours after electroporation no extrachromosomal DNA could be detected, confirming the low transformation efficiency already determined both by Schleper *et al*. (14) and Cannio *et al*. (23). Nevertheless the plasmid pSSVrt spread efficiently throughout a culture after transformation; in fact after the first scale-up from 3 to 50 ml of the culture, the amount of the plasmid and of the helper virus SSV2 increased and could be detected by ethidium bromide fluorescence on agarose gels (Figure 3). Identical results were obtained with all the different DNA concentrations used for electroporation but a lower amount (~5-fold) of the hybrid plasmid could be detected in the P2 and MT4 (data not shown) strains when compared with the GθW strain (Figure 3).

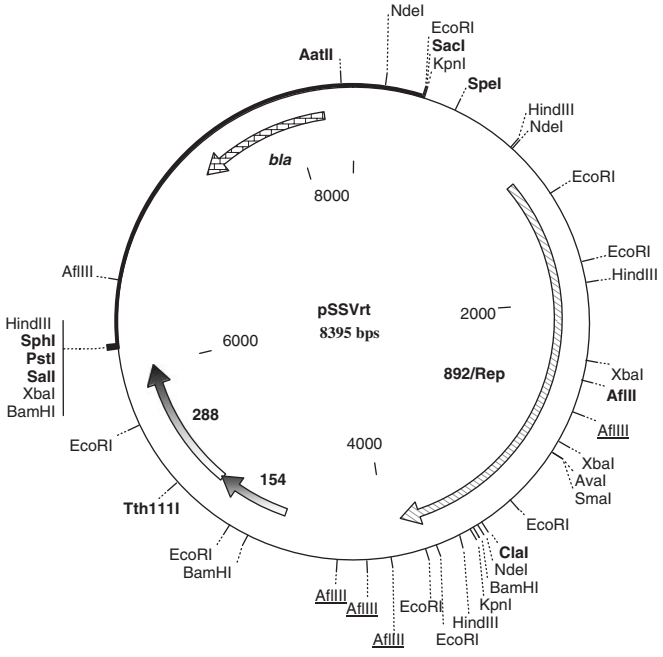

Figure 2. Plasmid map of the pSSVx/pUC19 shuttle vector. The pSSVrt shuttle vector has been obtained by cloning the pUC19 bacterial sequence (heavy bar) into an AflIII restriction site (all AflIII sites are underlined) downstream of the ORF288 of the pSSVx plasmid and by propagation in *E.coli*. Main ORFs allow easy location of archaeal (replicase ORF892/Rep as well as ORF154 and ORF288 putative DNA packaging components of pSSVx) and bacterial (*bla*, ampicillin resistance gene) sequences; bold face letters indicate unique restriction sites.

As expected, the same experiments performed on wild-type strains not infected with SSV2 produced no transformation by the shuttle vector, confirming the need of the helper virus for the propagation of the engineered plasmid.

A long-term growth experiment was also carried out with transformed *S.solfataricus* GθW. A culture containing pSSVrt was grown under standard conditions until it reached an optical density of 0.8, then diluted 1:10 in the same medium and grown and diluted again twice in the same manner. At this point samples were withdrawn at different cell densities and the DNA from the cells was analyzed; no change in the DNA replication and accumulation was observed (Figure 4) even after storing the propagated culture as a frozen −80°C glycerol stock and repeating the dilution/growth cycles. The relative fluorescence intensities of the pSSVrt and pSSVx were almost identical and the comparison with DNA fragments loaded at known concentrations on the gels allowed the estimation of ∼130 and 150 copies per cell (density of 1.0–1.2 OD_{600}) for the engineered and the unmodified viral DNAs.

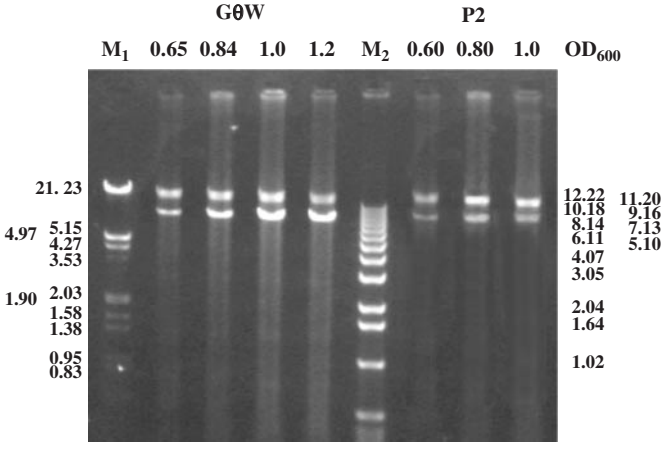

Figure 3. Transformation of *S.solfataricus* with the shuttle vector pSSVrt. Transfer into *S.solfataricus* GθW and P2 strains was performed by electroporation and after propagation of the cultures, the cells were harvested at different growth phases (indicated as optical density values at 600 nm on the top) for extrachromosomal DNA preparation. The upper and lower bands in each preparation correspond to SSV2 virus and recombinant satellite pSSVrt, respectively, both linearized with SalI. M_1 and M_2: DNA molecular weight markers (kb).

pSSVx-derived shuttle vectors for the *lac*S gene transfer and expression in the β-galactosidase deficient mutant GθW

β-D-galactosidase activity of *S.solfataricus* is displayed by the *lacS* gene product (40,44) and is responsible for the typical blue stain of the cells when exposed to the chromogenic substrate, X-gal; the lack of gene function in defective mutant strains is hence responsible for colorless colonies (45). The *lacS* coding sequence with a 648 bp 3′-untranslated region (3′-UTR) was fused to a 448 bp fragment containing the promoter region of tf55α and the first five codons of its ORF; this 2578 bp expression cassette was inserted into the polycloning site of the pSSVrt vector, increasing its size to ∼11 kb. A similar gene fusion was already demonstrated to efficiently complement a β-galactosidase defect in *Sulfolobus* when carried by a viral SSV1-derived shuttle vector (35). The plasmid was used to transform the strain GθW (Δ*lacS*) and after electroporation, cell culture was regenerated to allow spreading and subsequently seeded as spots onto plates as previously described (32). After 2–3 days of incubation, the colonized areas were overlaid with X-gal and reincubated at 75°C. No color was developed even after prolonged incubation, namely no expression of β-galactosidase activity could be detected, suggesting the failure of the recombinant plasmid to transport and/or to express the *lacS* gene. Extrachromosomal DNA preparation did not contain the recombinant pSSVx confirming that interference could have occurred either at the level of DNA propagation or of the recombinant particle proliferation/spreading. Southern analysis of extrachromosomal DNAs from this primary transformants demonstrated that the DNA transfer and maintenance were unaffected whereas spreading was impeded, as indicated by specific bands that could be detected at constant but very

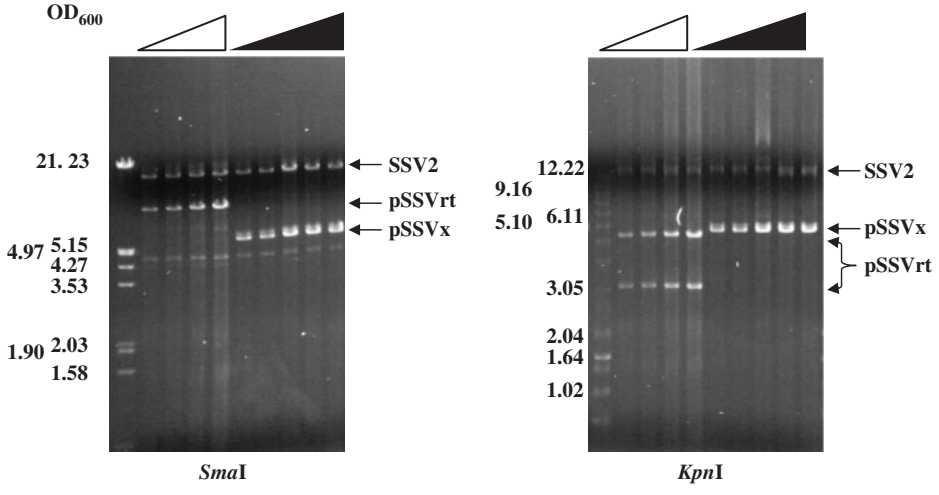

Figure 4. Maintenance and propagation of pSSVrt in *S.solfataricus*. After storage, SSV2 lysogens of *S.solfataricus* GθW transformed with pSSVrt and wild-type pSSVx were grown for several generations to check transformation stability. The propagated cultures were tested by restriction analysis of extrachromosomal DNAs (extracted at increasing cell densities from ∼0.5 to 1.2 OD_{600}); the SmaI enzyme produced single cuts in both satellite and SSV2 elements, whereas KpnI had 1 and 2 recognition sites in pSSVx and pSSVrt, respectively, and no site on SSV2. Standards of DNA molecular weight are indicated (kb).

weak intensity for cells withdrawn at different generation stages (data not shown). To confirm this preliminary results a size reduction of the pSSVrt vector was obtained by eliminating a redundant sequence in the *E.coli* plasmid moiety not necessary for replication and ampicillin selection and producing the vector pMSSV. Moreover a smaller (2025 bp) *lacS* expression cassette provided by Dr C. Schleper (35) and containing a shorter 3′-UTR replaced the one used in the first attempt. This *lacS* expression cassette was inserted into the polycloning site of pMSSV, generating the expression vector pMSSV*lacS* (Figure 5). Both pMSSV and pMSSV*lacS* were transferred into *S.solfataricus* GθW/SSV2

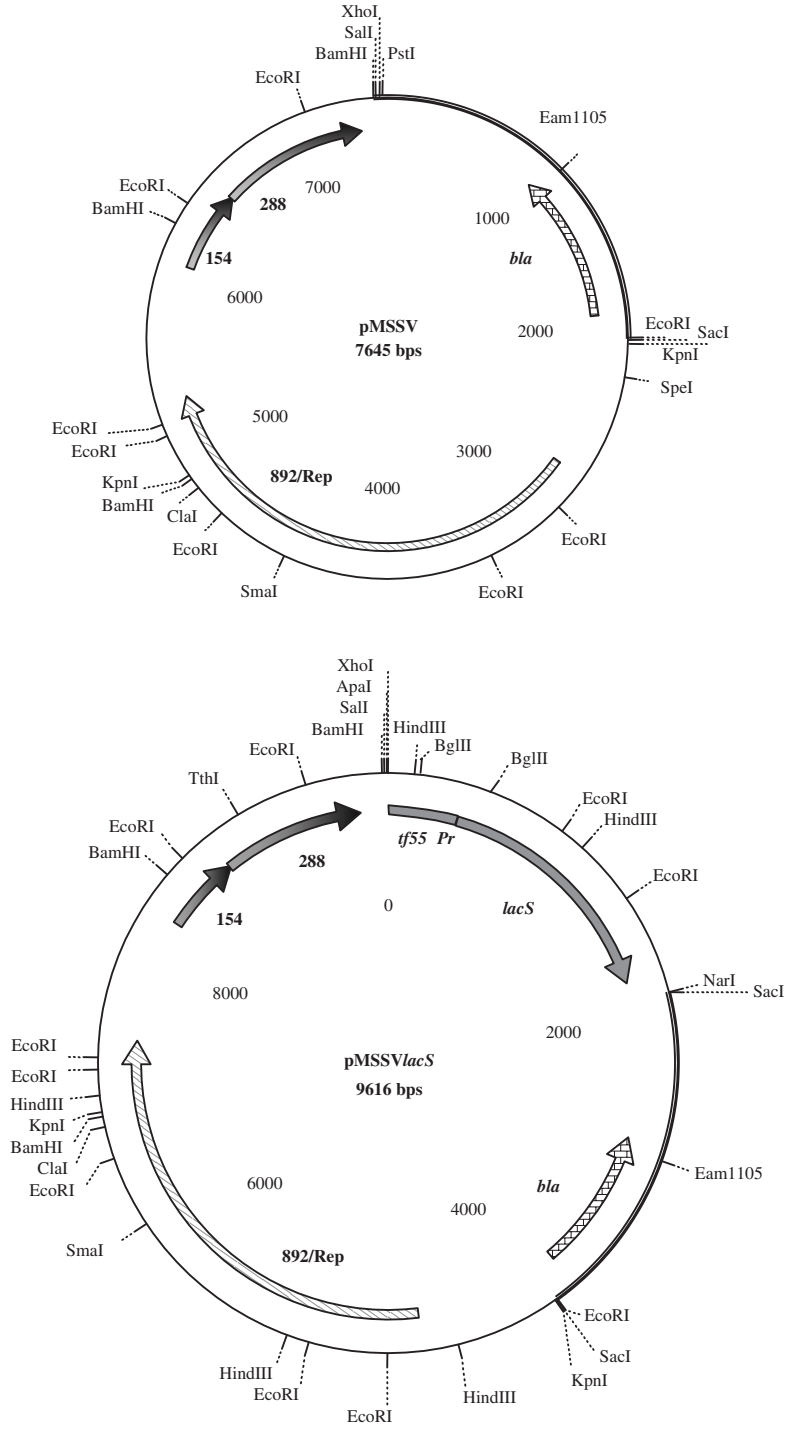

Figure 5. Plasmid maps of the minimal plasmid pMSSV and of the β-glycosidase expression vector pMSSV*lacS*. The bacterial moiety (indicated as a solid bar) of the plasmid pSSVrt was reduced in size, by eliminating every redundant sequence from the pUC19 vector and maintaining its ColE1 replicon and the ampicillin resistance marker (*bla*), to generate a new shuttle vector named pMSSV. An expression cassette containing the *lacS* gene fused to the thermosome tf55α subunit promoter was inserted into the newly constructed vector producing the expression vector pMSSV*lacS*.

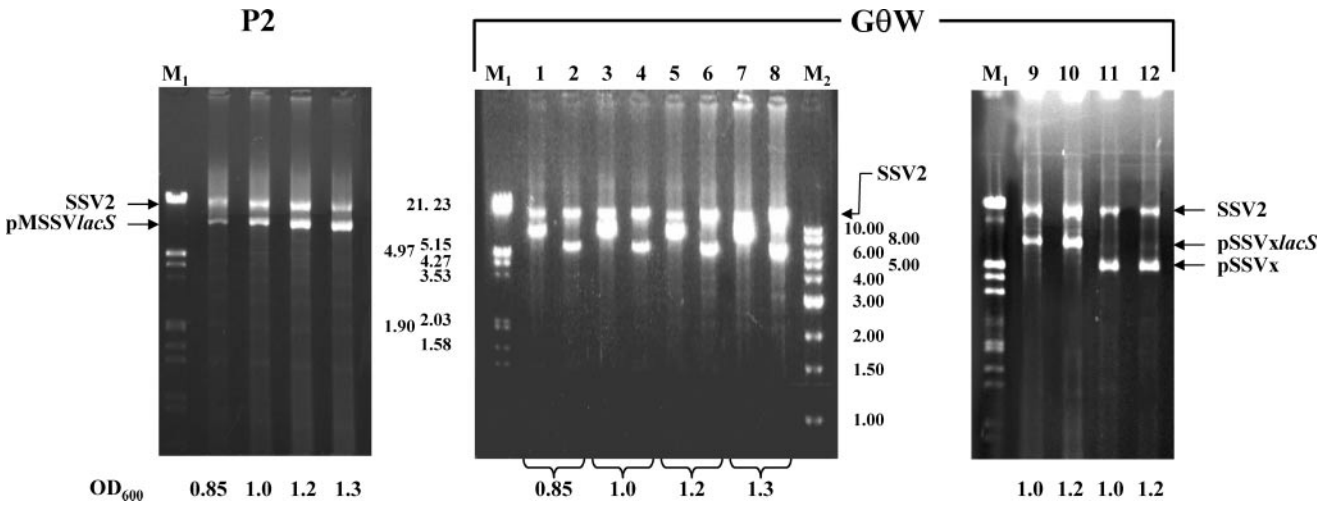

Figure 6. Transformation of *S.solfataricus* with the vectors pMSSV and pMSSV*lacS*. Competent cells of *S.solfataricus* P2 and GθW infected with the SSV2 virus were electroporated with the vector carrying the β-glycosidase gene pMSSV*lacS*. Strain GθW was also transformed with pSSVx*lacS*, a pMSSV*lacS*-derived plasmid obtained by excision of the *E.coli*-specific sequences and re-ligation. pMSSV and wild-type pSSVx were also transferred into the GθW/SSV2 lysogen for comparison of vector transfer and propagation efficiency in the presence and absence of the *lacS* gene. Extrachromosomal DNAs extracted from transformed cells after propagation of the cultures were checked by agarose electrophoresis for the presence of pMSSV*lacS* (all lanes in P2 and lanes 1, 3, 5 and 7 in GθW) and pMSSV (lanes 2, 4, 6 and 8 in GθW) at the different growth stages indicated as cell densities (OD_{600}). Similarly, recovery of the pSSVx*lacS* plasmid (SacI cut, lanes 9 and 10) was monitored in comparison to the parental pSSVx (SmaI cut, lanes 11 and 12). Mobility of the virus SSV2 is highlighted. M_1 and M_2: molecular weight markers (kb).

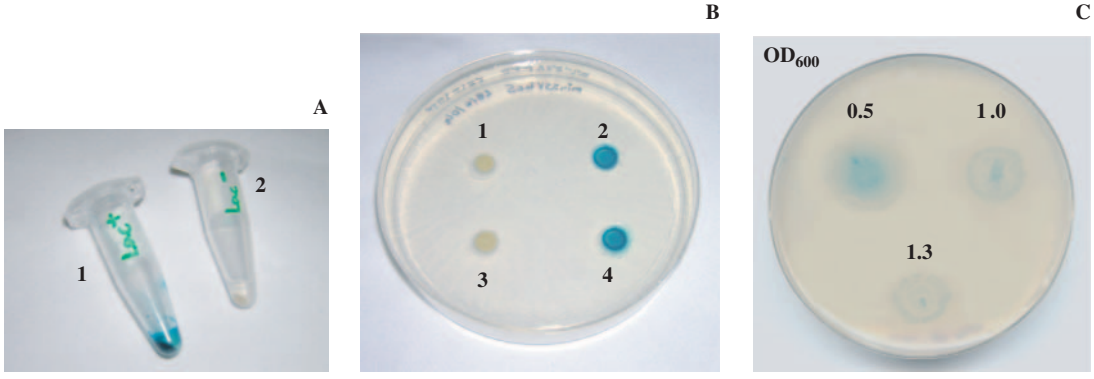

Figure 7. Complementation of the β-glycosidase mutation in *S.solfataricus* GθW and spreading of the pMSSV*lacs* vector. The expression vector pMSSV*lacS* was transferred into *S.solfataricus* GθW cells lysogens for SSV2. Successful transformation was checked after propagation for several generation by direct exposure of cell pellets from liquid cultures to X-Gal and development of the blue color [**A** (1) Lac^+, pMSSV*lacS* transformed cells; (2) Lac^-, cells transfected with pSSVx]. Maintenance of the plasmid was also confirmed for colonized areas on plates and X-Gal test [**B** (1) Strain GθW; (2) wild-type strain Gθ; (3) SSV2/ pMSSV-infected GθW; (4) SSV2/ pMSSV*lacS* -infected GθW]. **C.** 4-µl aliquots of culture supernatant of SSV2/pMSSV*lacS*-infected GθW, withdrawn at the different cell densities indicated, were spotted onto a continuous lawn of uninfected GθW cells. The spreading of the recombinant satellite particles was revealed by the blue color developed on the plaques (primary infected cells) upon X-Gal exposure. Stain intensity depended on the extent of cell growth retardation in the plaques and was progressively less intense for increasingly higher virus titers.

cells by electroporation, whereas pMSSV*lacS* was also transferred into cells of the P2/SSV2 strain. After culture propagation, extrachromosomal DNAs were prepared and analyzed by agarose electrophoresis (Figure 6). The presence and growth-dependent accumulation of both vectors in GθW/ SSV2 and of pMSSV*lacS* in P2/SSV2 revealed successful transformation and DNA replication, confirming that the plasmid size was critical for particle formation and spreading; moreover the *lacS* gene is harmless and does not induce recombination of the vector also in the P2 strain which already contains a wild-type chromosomal copy of the gene. Identical results were obtained when SSV2 lysogens of *Sulfolobus* GθW were transformed with the plasmid pSSVx*lacS* lacking the *E.coli* sequences and obtained by cleavage with *Sac*I and re-ligation (Figure 6). All pSSVx derived plasmid failed to transform *Sulfolobus* if linearized prior to transfer. In fact, X-Gal staining test was positive on *Sulfolobus* GθW/SSV2 transformed with pMSSV*lacS* (or pSSVx*lacS*, data not shown) on liquid-cultured (Figure 7A), plated (Figure 7B), and primary infected cells (Figure 7C). Plaques depicted in Figure 7C stained after incubation with the chromogenic substrate as a result of propagation of the engineered satellite virus. Since cell growth retardation in the plaques is directly proportional to the SSV2 virus titer

which is maximal in the stationary phase cultures (P. Contursi, unpublished data), turbidity and color faded progressively with increasing cell density.

Therefore, the β-galactosidase was expressed from the recombinant plasmid and the engineered satellite virus had spread efficiently throughout the culture.

pMSSV*lacS* DNA prepared from *S.solfataricus* could be re-transferred into *E.coli* without suffering recombinational rearrangements. Plasmid preparations and total DNA from *Sulfolobus* transformants were analysed by restriction analyses and Southern hybridization that confirmed the maintenance of the vector at the same levels of the parental pSSVrt. No vector integration into the host chromosome occurred since the recombinant vector exhibited the same restriction pattern as the DNA prepared from *E.coli* and no signal relative to the plasmid could be detected on genomic DNAs in the Southern blots (Figure 8). Bands corresponding to chromosomal *tf55*α promoter (in all strains tested) and *lac*S gene (absent in GθW) could also be visualized as internal controls for detection of single copy chromosomal sequences. Signal assignment of the restriction fragments of *tf55*α gene was performed on the basis of the sequence and localization on the P2 strain genome, whereas the mapping performed by Bartolucci *et al*. (32) allowed the identification of *lacS* gene specific signals. A similar Southern analysis confirmed that the pMSSV*lacS* copy number varied from 10–15 (mid-log phase) to ∼130 molecules per cell (stationary phase).

Isolation of single pMSSV*lacS* transfectants

Single colonies formed by suitably diluted mid exponential pMSSV*lacS* transformant cells (0.3 OD_{600}) on rich solid medium (M182, glucose 0.1%) resulted positive to X-Gal staining only in the fraction of $1–3 \times 10^{-3}$; below this cell density value the presence of the plasmid was undetectable in plated cells. The fraction of positive clones could be increased to 40% when stationary phase cultures were plated, namely when the plasmid had reached its maximum copy number per cell. Unfortunately, blue stained clones loosed the recombinant plasmid when suspended and directly propagated in liquid cultures as indicated by the analysis on extrachromosomal DNA and negative staining with X-Gal. This result was nevertheless expected, since also wild-type pSSVx has been demonstrated to be lost in cells of single colonies on plates and/or of progressively diluted cultures (9); indeed we confirmed and took advantage of this feature for the SSV2 lysogen selection of GθW and P2.

In order to stabilize the pMSSV*lacS* transformants, selection on minimal media containing lactose as the only nutrient source was performed (cells not complemented for *lacS* function are unable to grow because of the lack of any β-galactosidase activity) (32). Streaking (and/or seeding of suitable dilutions) revealed that the culture had 100% colony forming efficiency on lactose. Moreover, all colonies resulted positive in the X-Gal test (Figure 9A), demonstrating that they were able to retain the plasmid under selective nutrient conditions. Interestingly, cells plated as spots after resuspension and immediate re-seeding onto rich medium, maintained the β-galactosidase activity (for the analysis of 10 independent clones see Figure 9B). This procedure (resuspension and immediate re-seeding) overcame critical dilution and was successful also for the smaller fraction of positive transformants isolated from solid rich medium.

These experiments demonstrate that cultures of pMSSV*lacS* transformants before selection on plates were homogeneously infected (all cells contained the plasmid) and that in diluted cell suspensions (such as those necessary for single colony formation on plates and/or obtained by transfer of single colonies in liquid medium), actively dividing cells loosed the vector, unless a selective pressure was imposed.

β-galactosidase assays on pMSSV*lacS* transfectants

Detection of β-galactosidase activity on denaturing gels by enzyme staining revealed that the *lacS* gene product was indistinguishable when expressed as heterologous in *E.coli* and xenologous in *S.solfataricus*, namely the *Sulfolobus* expression system did not interfere with the correct polypeptide syntesis (data not shown).

β-galactosidase activity was measured using a spectrophotometric assay with the specific substrate ONPG in crude cell extracts of both transfected mutant GθW strain and wild type Gθ. The mutant recipient strain GθW has been shown to exhibit no detectable activity (32), whereas the specific activities of pMSSV*lacS* transformants in primary transformation mixtures rose from undetectable levels to 1.2 U/mg protein; this value was ∼2-fold higher than that found in wild-type cells expressing *lacS* under the control of its own promoter. The activity remained stable in diluted and propagated cells cultured as described above, when assays were performed at the same growth phase. Under identical growth conditions, the value of endogenous β-galactosidase activity in the P2 wild-type strain (i.e. the natural source of both the *lacS* gene and *tf55*α promoter sequences inserted in the pMSSV*lacS*), reached 0.1 U/mg as the highest value. Therefore, using this value as reference the expression level in GθW/pMSSV*lacS* is ∼12-fold higher. For a single culture, activity increased up to late logarithmical growth phase and then maintained approximately the same value up to late stationary phase (Figure 10). This result confirmed that the recombinant satellite viral DNA was replicated inside the cells and that virus particle formed and spread throughout the culture; the increasing activity should therefore depend only on the copy number of the plasmid and hence on the number of the *lacS* gene copies per cell.

Heat shock of the stably transfected cells, shifting the culture temperature from 75° to 88°C, induced an increase of the specific activity up to 2.5-fold (3.0 U/mg) after 3 h and remained constant in cells exposed to thermal stress for 24 h.

DISCUSSION

In this paper, we have developed a relatively small-sized and high copy number shuttle vector for *S.solfataricus* based on the satellite virus pSSVx from *S.islandicus* REY15/4.

An extended intergenic region between the still uncharacterized ORFs c68 and 288 on the pSSVx genome sequence

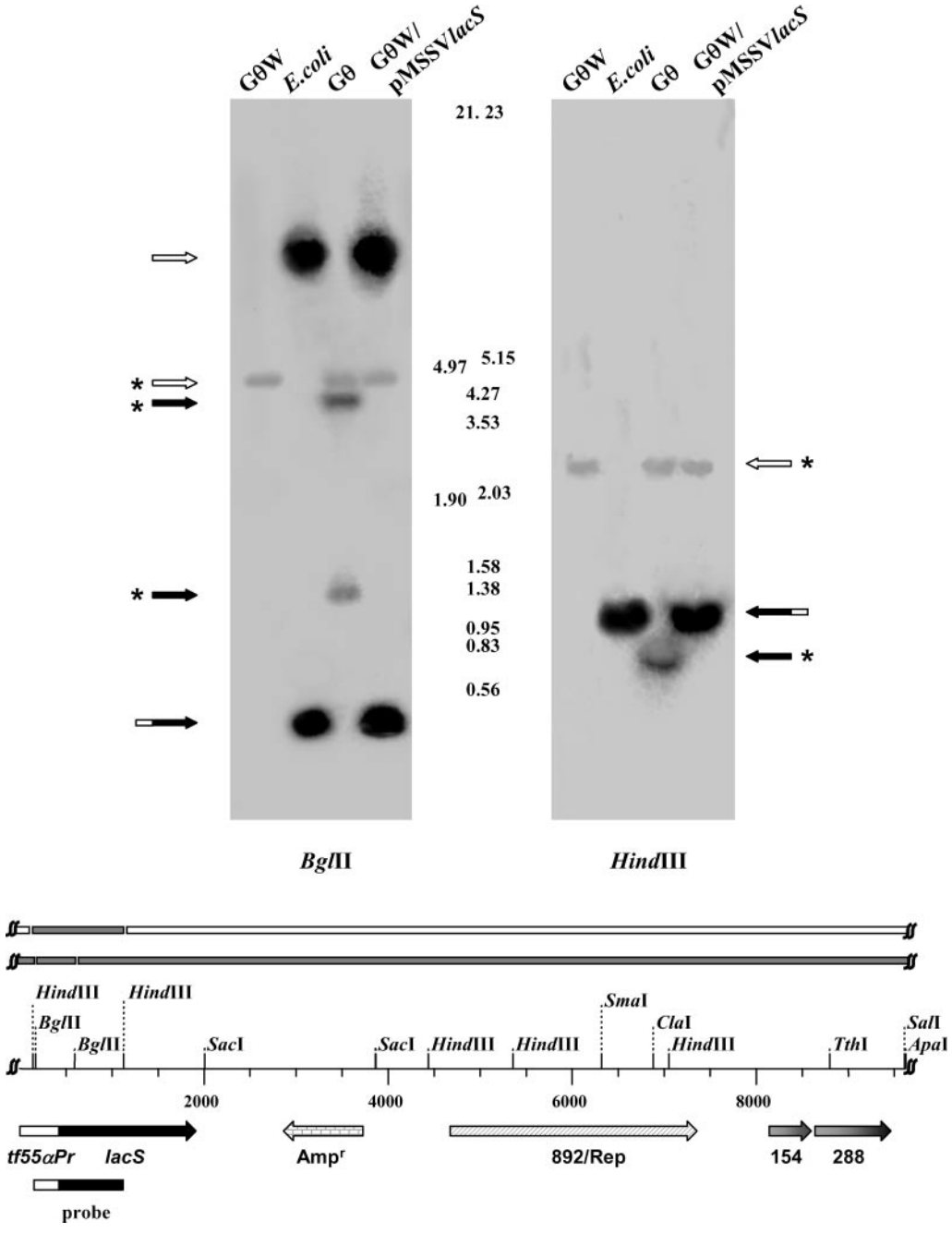

Figure 8. Southern analysis of pMSSV*lacS* transformants. For hybridization, total DNAs from transformant cells (GθW/pMSSV*lacS*) and from the recipient mutant strain GθW as well as from wild-type cells Gθ were cut with BglII and HindIII as indicated. The recombinant pMSSV*lacS* vector prepared from *E.coli* before transformation was used as a reference for correct restriction patterns (molecular weight standards are indicated). White and black arrows indicate hybridization to *tf55*α and *lacS* gene sequences, respectively, the asterisks distinguishing signals of the chromosomal copies. Signals of the *tf55*α–*lacS* gene fusion on the vector are highlighted by white/black arrows. The scheme on the bottom represents the restriction patterns for the two enzymes on the map of the linearized pMSSV*lacS* and the fragments producing positive signals for hybridization (grey bars).

appeared potentially useful for inserting foreign DNA. In fact, the vector pSSVrt clearly showed efficient replication and maintenance in SSV2 lysogens of the strains GθW and P2.

The insertion of foreign sequences into the pSSVrt vector allowed also to determine the limiting upper size (~11 kb) of the DNA to be accommodated in the virus particles. The additional sequences severely affected packaging and spreading rather than the transfer and/or the replication/maintenance of the viral DNA. These results also indicated that transport in viable virions and no other mechanism, such as conjugation (29,46–48), is responsible for cell-to-cell transfer of this genetic element.

A 'minimal' vector pMSSV was devised to accommodate inserts of a wider size range at least up to the 2.0 kb of a smaller *tf55*α/*lacS* expression cassette (35). In fact, the deletion of a region non-essential for replication and

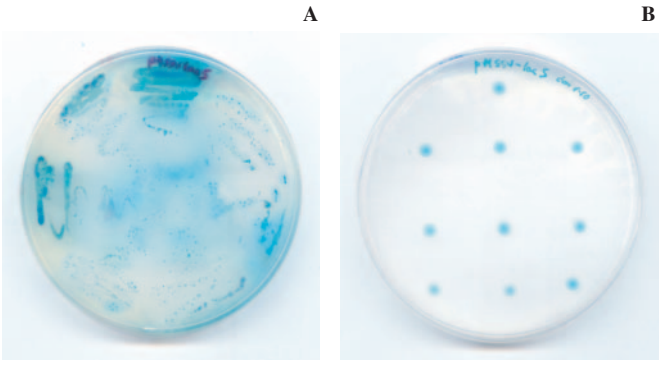

Figure 9. Selection of the pMSSV*lacs* transformants as independent clones. SSV2/GθW cells from a stationary phase culture (1.2 OD_{600}) grown in rich medium and containing pMSSV*lacS* at maximum copy number were streaked onto minimal 0.25% lactose medium to obtain single colonies. After blue color development with X-Gal at 75°C (**A**), 10 independent clones were picked, resuspended in liquid minimal medium and seeded drop wise onto rich solid medium for growth in circumscribed colonized areas. All clones maintained the plasmid as indicated by the positive X-Gal test (**B**).

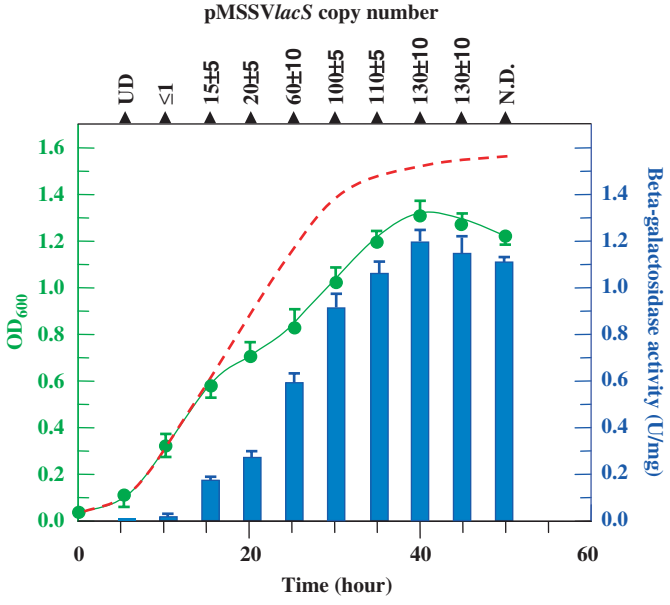

Figure 10. β-galactosidase activity and plasmid copy number in cultures of pMSSV*lacS* transformants. The β-galactosidase activity was measured in crude extracts of pMSSV*lacS* transformants harvested at different optical densities from a single culture. Specific activity (indicated by solid bars) is plotted together with the growth curve and the relative plasmid copy number. All values are the average of independent experiments performed on three different cultures and standard deviation is indicated for optical density and activity as well as for plasmid copy number. The dashed red line indicates a typical growth curve of the empty strain GθW. UD, undectectable; N.D., not determined.

selection in *E.coli* and a shortened *lacS* 3′-UTR resulted in an efficient shuttle vector able to transform *S.solfataricus* and to spread in primary cultures. The vector did not suffer either integration into the host chromosome or rearrangements and resided with undetectable to 130 copies per chromosome, the number depending on the cell growth phase. Interestingly, β-galactosidase activity in complemented mutant cells followed an increasing trend which correlated to the plasmid accumulation up to stationary phase and was not dependent on the number of the generations. *lacS* gene expression was inducible by thermal stress, namely it was confirmed to be transcriptionally regulated by the chaperonin gene promoter. However, heat shock did not produce the same β-galactosidase activity increase (∼10-fold) described by Jonuscheit *et al* (35), although maximum gene expression resulted at comparable levels in the two host/vector systems (3 versus 5 U/mg).

Maximum size for DNA insertion into pMSSV*lacS* was determined to be ∼1.3 kb (data not shown). In principle longer inserts up to 3.1 kb can be accommodated after cloning in *E.coli* and excision of the bacterial sequences prior to the transfer into *Sulfolobus*.

Maintenance of pMSSV*lacS* in early exponential grown and/or plated cells seemed to be the critical point for the efficacy of this system, because of both the low copy number (this work) and the reduced viral production in metabolically active cells (P. Contursi, unpublished data). In fact, segregation of the plasmid in an actively dividing cell (in which SSV2 is also maintained at low copy number and hence hardly sustains active replication and/or particle proliferation of pSSVx and derivatives) can be very asymmetrical until it is lost. We were able to overcome this bottleneck; as expected, efficient selection and stabilization of single pMSSV*lacS* transformants was indeed obtained by plating cells grown up to stationary phase (maximum copy number and highest viral production), by avoiding critical dilutions in the resuspension of single colonies and/or by imposing selective metabolic pressure in lactose minimal media.

This system presents many advantages compared with others already mentioned: (i) after DNA transfer, the vector propagates efficiently throughout the culture as a virus, overcoming the usual low transformation efficiency of *Sulfolobus* cells; (ii) the vector is stably maintained at high-copy number with no integration into the chromosome and hence no reduction in the number of episomal molecules and (iii) in stable transfectants, the β-galactosidase activity is dependent only on the copy number of the vector.

The availability of this new two-element transformation systems based on SSV2 and the engineered pSSVx will contribute to clarify the mechanisms responsible for the satellite/helper dependence as well as for replication, gene regulation and packaging of the episomal DNAs.

Further work will explore the use of this vector for the expression of both homologous and heterologous genes in *S.solfataricus* as well as for testing *Sulfolobus* regulatory sequences.

ACKNOWLEDGEMENTS

The authors are grateful to Prof. Quinxin She for providing us with the strain REY 15/4 of *S.islandicus*. This work was funded by the Ministero dell'Università e della Ricerca Scientifica (Decreto Direttoriale Prot. N. 1105/2002). Funding to pay the Open Access publication charges for this article was provided by 'MIUR-Decreto Direttoriale prot. n. 1105/2002' Project.

Conflict of interest statement. None declared.

REFERENCES

1. Weinbauer,M.G. and Rassoulzadegan,F. (2004) Are viruses driving microbial diversification and diversity? *Environ. Microbiol*., **6**, 1–11.
2. Forterre,P. (2001) New viruses for the new millennium. *Trends Microbiol*, **9**, 114.
3. Wommack,K.E. and Colwell,R.R. (2000) Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev*., **64**, 69–114.
4. Zillig,W., Prangishvilli,D., Schleper,C., Elferink,M., Albers,S., Janekovic,D. and Gotz,D. (1996) Viruses, plasmids and other genetic elements of thermofilic and hypertermofilic *Archaea*. *FEMS Microbiol. Rev*., **18**, 225–236.
5. Prangishvili,D. (2003) Evolutionary insights from studies on viruses of hyperthermophilic Archaea. *Res. Microbiol*., **154**, 289–294.
6. Snyder,J.C., Stedman,K., Rice,G., Wiedenheft,B., Spuhler,J. and Young,M.J. (2003) Viruses of hyperthermophilic Archaea. *Res. Microbiol*., **154**, 474–482.
7. Arnold,H.P., Ziese,U. and Zillig,W. (2000) SNDV, a novel virus of the extremely thermophilic and acidophilic archaeon *Sulfolobus*. *Virology*, **272**, 409–416.
8. Zillig,W., Arnold,H.P., Holz,I., Prangishvili,D., Schweier,A., Stedman,K., She,Q., Phan,H., Garrett,R. and Kristjansson,J.K. (1998) Genetic elements in the extremely thermophilic archaeon *Sulfolobus*. *Extremophiles*, **2**, 131–140.
9. Arnold,H.P., She,Q., Phan,H., Stedman,K., Prangishvilli,D., Holz,I., Kristjansson,J.K., Garret,R. and Zillig,W. (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.
10. Rice,G., Stedman,K., Snyder,J., Wiedenheft,B., Willits,D., Brumfield,S., McDermott,T. and Young,M.J. (2001) Viruses from extreme thermal environments. *Proc. Natl Acad. Sci. USA*, **98**, 13341–1345.
11. Rice,G., Tang,L., Stedman,K., Roberto,F., Spuhler,J., Gillitzer,E., Johnson,J.E., Douglas,T. and Young,M. (2004) The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc. Natl Acad. Sci. USA*, **101**, 7716–7720.
12. Wiedenheft,B., Stedman,K., Roberto,F., Willits,D., Gleske,A.K., Zoeller,L., Snyder,J., Douglas,T. and Young,M. (2004) Comparative genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses. *J. Virol*., **78**, 1954–1961.
13. Martin,A., Yeats,S., Janekovic,D., Reiter,W.D., Aicher,W. and Zillig,W. (1984) SAV1, a temperature u.v.-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius*. *EMBO J*., **3**, 2165–2168.
14. Schleper,C., Kubo,K. and Zillig,W. (1992) The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc. Natl Acad. Sci. USA*, **89**, 7645–7649.
15. Palm,P., Schleper,C., Grampp,B., Yeats,S., McWilliam,P., Reiter,W.D. and Zillig,W. (1991) Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. *Virology*, **185**, 245–250.
16. She,Q., Peng,X., Zillig,W. and Garret,R.A. (2001) Gene capture in archaeal chromosomes. *Nature*, **409**, 478.
17. Reiter,W.D., Zillig,W. and Palm,P. (1988) Archaebacterial viruses. *Adv. Virus Res*., **34**, 143–188.
18. She,Q., Brugger,K. and Chen,L. (2002) Archaeal integrative genetic elements and their impact on genome evolution. *Res Microbiol*., **153**, 325–332.
19. Stedman,K., She,Q., Phan,H., Arnold,H.P., Holtz,I., Garret,R. and Zillig,W. (2003) Relationships between fuselloviruses infecting the extremely thermofilic archaeon *Sulfolobus*: SSV1 and SSV2. *Res. Microbiol*., **154**, 295–302.
20. Kletzin,A., Lieke,A., Urich,T., Charlebois,R.L. and Sensen,C.W. (1999) Molecular analysis of pDL10 from *Acidianus ambivalens* from extremely thermophilic and acidophilic Archaea. *Genetics*, **152**, 1307–1314.
21. Peng,X., Holz,I., Zillig,W., Garrett,R.A. and She,Q. (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.
22. Zillig,W., Kletzin,A., Schlepr,C., Holz,I., Janekovic,D., Nain,J., Lanzerdorfer,M. and Kristjanson,J.K. (1994) Screening for *Sulfolobales*, their plasmids and their viruses in Icelandic solfataras. *Syst. Appl. Microbiol*., **16**, 609–628.
23. Cannio,R., Contursi,P., Rossi,M. and Bartolucci,S. (2001) Thermoadaptation of a mesophilic hygromicin B phosphotransferase by directed evolution in hyperthermophilic Archaea: selection of a stable genetic marker for DNA transfer into *Sulfolobus solfataricus*. *Extremophiles*, **5**, 153–159.
24. Contursi,P., Cannio,R., Prato,S., Fiorentino,G., Rossi,M. and Bartolucci,S. (2003) Development of a genetic system for hyperthermophilic Archaea: expression of a moderate thermophilic bacterial alcohol dehydrogenase gene in *Sulfolobus solfataricus*. *FEMS Microbiol. Lett*., **218**, 115–120.
25. She,Q., Singh,R.K., Confalonieri,F., Zivanovic,Y., Allard,G., Awayez,M.J., Chan-Weiher,C.C.Y., Clausen,I.G., Curtis,B.A., De Moors,A. *et al*. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl Acad. Sci. USA*, **98**, 7835–7840.
26. Pisani,F.M. and Rossi,M. (1994) Evidence that an archaeal alpha-like DNA polymerase has a modular organization of its associated catalytic activities. *J. Biol. Chem*., **269**, 7887–7892.
27. Chong,P.K. and Wright,P.C. (2005) Identification and characterization of the *Sulfolobus solfataricus* P2 proteome. *J. Proteome Res*., **4**, 1789–1798.
28. Snijders,A.P., de Koning,B. and Wright,P.C. (2005) Perturbation and interpretation of nitrogen isotope distribution patterns in proteomics. *J. Proteome Res*., **4**, 2185–2191.
29. Elferink,M., Schleper,C. and Zillig,W. (1996) Transformation of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus* via a self-spreading vector. *FEMS Microbiol. Lett*., **137**, 31–35.
30. Aravalli,R.N. and Garret,R. (1997) Shuttle vectors for hyperthermophilic archaea. *Extremophiles*, **1**, 183–191.
31. Cannio,R., Contursi,P., Rossi,M. and Bartolucci,S. (1998) An autonomously replicating vector for *Sulfolobus solfataricus*. *J. Bacteriol*., **180**, 3237–3240.
32. Bartolucci,S., Rossi,M. and Cannio,R. (2003) Characterization and Functional complementation of a nonlethal deletion in the cromosome of a β-glycosidase mutant of *Sulfolobus solfataricus*. *J. Bacteriol*., **185**, 3948–3957.
33. Worthington,P., Hoang,V., Perez-Pomares,F. and Blum,P. (2003) Targeted disruption of the alpha-amylase gene in the hyperthermophilic archaeon *Sulfolobus solfataricus J. Bacteriol*., **185**, 482–488.
34. Albers,S.V., Jonuscheit,M., Dinkelaker,S., Urich,T., Kletzin,A., Tampe,R., Driessen,AJ. and Schleper,C. (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Environ. Microbiol*., **72**, 102–111.
35. Jonuscheit,M., Martusewitsch,E., Stedman,K. and Schleper,C. (2003) A reporter gene system for the hypertermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Mol. Microbiol*., **48**, 1241–1252.

36. Grogan,D.W. (1989) Phenotypic characterization of the archaebacterial genus *Sulfolobus*: comparison of five wild-type strains. *J. Bacteriol.*, **171**, 6710–6719.
37. Stedman,K., Schleper,C., Rumpf,E. and Zillig,W. (1999) Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors. *Genetics*, **152**, 1397–1405.
38. Cubellis,M.V., Rozzo,C., Monteucchi,P. and Rossi,M. (1990) Isolation and sequencing of a new beta-glycosidase-encoding archaebacterial gene. *Gene*, **94**, 89–94.
39. Sambrook,J. and Russel,D.W. (2001) *Molecular Cloning: A Laboratory Manual*. 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
40. Pisani,F.M., Rella,R., Raia,C.A., Rozzo,C., Nucci,R., Gambacorta,A., De Rosa,M. and Rossi,M. (1990) Thermostable beta-galactosidase from the archaebacterium *Sulfolobus solfataricus*. Purification and properties. *Eur. J. Biochem.*, **187**, 321–328.
41. Reiter,W.D., Palm,P. and Zillig,W. (1988) Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. *Nucleic Acids Res.*, **16**, 2445–2459.
42. Ermolaeva,M.D., Khalak,H.G., White,O., Smith,H.O. and Salzberg,S.L. (2000) Prediction of transcription terminators in bacterial genomes. *J. Mol. Biol.*, **301**, 27–33.
43. de Hoon,M.J.L., Makita,Y., Nakai,K. and Miyano,S. (2005) Prediction of transcriptional terminators in *Bacillus subtilis* and related species. *PLoS Comput. Biol.*, **1**, e25.
44. Grogan,D.W. (1991) Evidence that beta-galactosidase of *Sulfolobus solfataricus* is only one of several activities of a thermostable beta-D-glycosidase. *Appl. Environ. Microbiol.*, **57**, 1644–1649.
45. Schleper,C., Roder,R., Singer,T. and Zillig,W. (1994) An insertion element of the extremely thermophilic archaeon *Sulfolobus solfataricus* transposes into the endogenous β-galactosidase gene. *Mol. Gen. Genet.*, **243**, 91–96.
46. Prangishvilli,D., Albers,S.V., Holz,I., Arnold,H.P., Stedman,K., Klein,T., Singh,H., Hiort,J., Schwiver,J., Kristjansson,J.K. *et al.* (1998) Conjugation in Archaea: frequent occurrence of conjugative plasmids in *Sulfolobus*. *Plasmid*, **40**, 190–202.
47. She,Q., Phan,H., Garret,R., Albers,S., Stedman,K. and Zillig,W. (1998) Genetic profile of pNOB8 from *Sulfolobus*: the first conjugative plasmid from an archaeon. *Extremophiles*, **2**, 417–425.
48. Stedman,K., She,Q., Phan,H., Holz,H., Singh,H., Prangishvilli,D., Garret,R. and Zillig,W. (2000) pING family of conjugative plasmids from the extremely thermophilic archaeon *Sulfolobus islandicus*: insights into recombination and conjugation in Crenarchaeota. *J. Bacteriol.*, **182**, 7014–7020.

Extremophiles (2006) 10:615–627
DOI 10.1007/s00792-006-0017-2

ORIGINAL PAPER

Patrizia Contursi · Susanne Jensen · Tiziana Aucelli
Mosé Rossi · Simonetta Bartolucci · Qunxin She

Characterization of the *Sulfolobus* host–SSV2 virus interaction

Received: 13 March 2006 / Accepted: 5 July 2006 / Published online: 3 August 2006

Abstract The *Sulfolobus* spindle virus, SSV2, encodes a tyrosine integrase which furthers provirus formation in host chromosomes. Consistently with the prediction made during sequence analysis, integration was found to occur in the downstream half of the $tRNA^{Gly}$ (CCC) gene. In this paper we report the findings of a comparative study of SSV2 physiology in the natural host, *Sulfolobus islandicus* REY15/4, versus the foreign host, *Sulfolobus solfataricus*, and provide evidence of differently regulated SSV2 life cycles in the two hosts. In fact, whereas a significant induction of SSV2 replication takes place during the growth of the natural host REY15/4, the cellular content of SSV2 DNA remains fairly low throughout the incubation of the foreign host. The accumulation of episomal DNA in the former case cannot be traced to decreased packaging activity because of a simultaneous increase in the virus titre in the medium. In addition, the interaction between SSV2 and its natural host is characterized by the concurrence of host growth inhibition and the induction of viral DNA replication. When this virus–host interaction was investigated using *S. islandicus* REY15A, a strain which is closely related to the natural host, it was found that the SSV2 replication process was induced in the same way as in the natural host REY15/4.

Keywords Archaeal virus · Integration · *Sulfolobus* · Replication induction · Growth regulation

Communicated by G. Antranikian

P. Contursi · S. Jensen · Q. She (✉)
Department of Biological Chemistry,
Institute of Molecular Biology and Physiology,
University of Copenhagen, Sølvgade 83H,
1307 Copenhagen, Denmark
E-mail: qunxin@mermaid.molbio.ku.dk
Tel.: +45-3532-2013
Fax: +45-3532-2040

T. Aucelli · M. Rossi · S. Bartolucci (✉) · P. Contursi
Dipartimento di Biologia Strutturale e Funzionale,
Universitá degli Studi di Napoli Federico II,
Complesso Universitario Monte S.Angelo,
Via Cinthia, Naples, Italy
E-mail: bartoluc@unina.it
Tel.: +39-081-679052
Fax: +39-081-679053

M. Rossi
Istituto di Biochimica delle Proteine,
CNR, via Pietro Castellino 111, 80131 Naples, Italy

Introduction

The first spindle-shaped virus was discovered in *Sulfolobus shibatae* B12, a crenarchaeon isolated from a sulphurous hot spring in Beppu, Japan (Grogan et al. 1990; Martin et al. 1984; Yeats et al. 1982). The virus was named SSV1 and has been characterized in greater detail. SSV1 virus production can be induced in both the natural and foreign hosts by UV irradiation or treatment with mitomycin C (Reiter et al. 1987a, b; Schleper et al. 1992). All transcripts in the viral genome have been mapped and their promoters and terminators have been determined (Palm et al. 1991; Reiter et al. 1987b, 1988). The SSV1 genome is stable in three forms: (1) the viral genome packaged in virus particles which is positively supercoiled (Nadal et al. 1986); (2) the episomal genome in the *Sulfolobus* cells which includes positively and negatively super-coiled and relaxed double-stranded DNAs (Nadal et al. 1986; Prangishvili et al. 2001; Snyder et al. 2003); (3) a provirus is present in the chromosomes of both the natural host, *Sulfolobus shibatae* B12, and the foreign host, *Sulfolobus solfataricus* P1 (Reiter and Palm 1990; Reiter et al. 1989; Schleper et al. 1992).

The functions of four SSV1 open reading frames (ORFs) have so far been identified based on functional analysis. The ORF codings of the three structural proteins VP1, VP2 and VP3 have been identified by sequencing the purified coat proteins (Reiter et al. 1987a). The fourth ORF encodes a tyrosine integrase

(Esposito and Scocca 1997; Nunes-Duby et al. 1998) of the SSV-type, which facilitates the recombination between the virus and archaeal host attachment sites, *attP* and *attA* (Muskhelishvili et al. 1993; Serre et al. 2002) whereas the remaining SSV ORFs show little or no significant sequence similarity to any sequence with a known function in public databases.

A few more SSV viruses have been isolated from different locations in the world, forming a novel *Fuselloviridae* virus family (Arnold et al.1999b). These include: (1) SSV2 and pSSVx, i.e. the helper and satellite viruses carried by *Sulfolobus islandicus* REY15/4, which were originally isolated from a solfataric hot spring in Reykjanes, Iceland (Arnold et al. 1999a; Stedman et al. 2003); (2) SSV RH, which was obtained from an enrichment culture sampled in the Ragged Hills region of the Norris Geyser basin in Yellowstone National Park, United States (Snyder et al. 2003; Wiedenheft et al. 2004) and (3) SSV K1, which was isolated in the Geyser Valley region of the Uzhno-Kamchatsky National Park in the Kamchatka peninsula, Russia (Snyder et al. 2003; Wiedenheft et al. 2004). Genome comparison has revealed that these four SSV viruses are similar in size and gene organization (Wiedenheft et al. 2004). Sequence and orientation similarity has been observed in the ORFs of one half of the genome, but not in those of the other half (Palm et al. 1991; Stedman et al. 2003; Wiedenheft et al. 2004).

Comparative genomics analysis with a selected set of archaeal viruses revealed that these viruses are interrelated albeit they exhibit very different morphotypes and belong to different virus families (Prangishvili et al. 2006a). Among SSV viruses, only SSV1 was included for that analysis. It has been revealed that SSV1 encodes ORFs orthologous to those present in the genomes of *Acidianus* ATV (Prangishvili et al. 2006b) and AFV1 (Bettstetter et al. 2003), as well as *Sulfolobus* SIRV1 (Blum et al. 2001; Peng et al. 2001) and STSV1 (Xiang et al. 2005) viruses, all of which belong to crenarchaeal viruses. This allows that Prangishvili et al. (2006a) postulated that these viruses share a common origin and their morphotypic diversity reflects their fast evolution. Moreover, archaeal viral genomes usually contain genes that are likely acquired from their hosts, and vice versa (Prangishvili et al. 2006a; She et al. 2004).

Another striking finding among *Sulfolobus* virus studies has been the identification of the first archaeal helper and satellite virus system: *Sulfolobus* viruses SSV2 and pSSVx. In this system, SSV2 acts as an ordinary virus and a helper virus to pSSVx, while pSSVx is a satellite virus assumed to generate virus particles thanks to the packaging mechanisms of SSV2 (Arnold et al. 1999a; Stedman et al. 2003). At sequence level, pSSVx genome contains the putative minimal replicon of the *Sulfolobus* pRN plasmid family (pRN1, pRN2, pHEN7), the *Acidianus ambivalens* plasmid pDL10 (Garrett et al. 2004), and several integrated plasmids which are found in *Sulfolobus* genomes (She et al. 2001a, 2002, 2004).

Here we report the characterization of the specific interactions between the SSV2 virus and its *Sulfolobus* hosts and provide evidence that SSV2 replication activity is reversely related to the growth rate of the natural host.

Materials and methods

Enzymes and chemicals

Restriction and modification enzymes were purchased from New England Biolabs or Amersham Biosciences. Synthetic oligonucleotides were purchased from TAG Copenhagen. Radioactive materials were obtained from Amersham Biosciences.

Strains used in this study

Sulfolobus islandicus REY15/4 and related *Sulfolobus* enrichment cultures were kindly provided by Dr. Wolfram Zillig (Zillig et al. 1998, 1994).

Sulfolobus islandicus REY31A is a pSSVx-cured strain isolated from REY15/4. The plasmid curation was done by growing REY15/4 continuously for approx. 100 generation times at an optical density below 0.4 (OD_{600}). Thereafter, six single colonies were isolated via Gelrite plating and purified by re-streaking on plate for three times. Since pSSVx was hardly detected at low optical density values in REY15/4 cells, the REY31 isolates, were carefully checked for the absence of pSSVx by polymerase chain reactions (PCRs) and Southern analysis over all the growth phases.

Sulfolobus islandicus REY15A is virus-free isolate obtained by Gelrite plating of the same enrichment culture from which the REY15/4 strain had been isolated. Six single clones (REY15A-F) were isolated and purified as single colonies as described for REY31A. They were then tested for the absence of extrachromosomal DNA and for the incapability in conferring growth inhibition to *S. solfataricus* P2 strain.

Sulfolobus solfataricus P2 (DSM1617) was purchased from the German Collections of Microorganisms and Cell Cultures.

Virus-harbouring *S. solfataricus* P2 (SSV2-P2) and *S. islandicus* REY15A (SSV2-REY15A) strains were obtained as following: 1–2 µl of the REY15/4 supernatant containing SSV2 and pSSVx virons were spotted onto the soft layer of a Gelrite plate seeded with *S. solfataricus* P2 or *S. islandicus* REY15A; the plates were incubated at 80°C for 2–3 days and turbid halos (plaques) resulting from the inhibition of host growth by SSV2 appeared; the virus-infected cells present in the plaques were extracted and revitalized in liquid medium. Finally, single colonies of SSV2-P2 and SSV2-REY15A were isolated and purified as described for REY31A.

Characterization of *S. islandicus* strains

Comparison of the *S. islandicus* REY strains/isolates were conducted by cloning and sequencing the gene codings of 16S RNA, MCM (mini-chromosomal maintenance protein) and the orotate phosphoribosyltransferase and orotidine-5′-monophosphate decarboxylase (*pyrEF*). The primers designed for PCR cloning of the above genes were listed in Table 1. PCR reaction was conducted as described for detecting SSV2 integration (see below) except that the annealing temperature was set at 55°C. The amplified PCR products were first purified by means of the QIAGen PCR Purification Kit and then sequenced using PCR primers and additional internal sequencing primers. DNA sequencing was carried out via dye terminator chemistry and MegaBACE 1000 (Molecular Dynamics/Amersham). The sequences obtained were analysed using Sequencher 4.2 (Gene Codes Corp).

Media and growth conditions

Sulfolobus were colonized on 0.8% gelrite plates as described by Zillig et al. (1994). For liquid cultivation, the basal salt solution was supplemented with 0.1% tryptone, 0.05% yeast extract and 0.2% sucrose (w/v); its pH was adjusted to 3.5 using concentrated H_2SO_4. Incubation was conducted in a 100 or 250 ml Erlenmeyer flask with a long neck, using the Innova 3100 water bath shaker (New Brunwick Scientific Corp). The incubation temperature was 80°C, and the shaking rate 150 rpm.

To determine the number of viable cells in *Sulfolobus* cultures, samples were taken at different OD_{600} values and used for making tenfold dilution series. The diluted samples were then mixed with 5 ml of the above medium containing 0.2% Gelrite and poured onto prewarmed Gelrite plates. Colonies were counted up after 5 days of incubation at 80°C.

Table 1 Primers used in this study

Name	Sequence
16SArc21F	5′-TTCCGGTTGATCCYGCCGG-3′
16SUni1492R	5′-GGTTACCTTGTTACGACTT-3′
SulMCMfwd	5′-TAGAAMTTAGAAAAMTAAGAAGTA-3′
SulMCMrev	5′-TCTCATKATATTTATTGCTCTTTC-3′
Sulpyr-fwd	5′-GTACTWGGCTCAAAGAATGCTA-3′
Sulpyr-rev	5′-GATATAGCTGARGCTACAATAGAT-3′
Sul16Sp03	5′-GTAAGATTCCAGGCGTTGACTCCA-3′
Sul16Sp04	5′-CCTGGGGAGTACGGTCGCAAGA-3′
SulMCM-p03	5′-MTTGAAGAMGATTTAGTTGAT-3′
SulMCM-p04	5′-GATAGAGTWGCMATTCATGARGCA-3′
SulMCM-p05	5′-GTYTGYTGTTCCATTGCYTCATGW-3′
Sulpyr-p03	5′-TCGTACATTAGAACTCGATGT-3′
Sulpyr-p04	5′-ATAAGTTAAATCTGTTGTTGGA-3′
Sulpyr-p05	5′-GAAACTACGTGTCTTAATCTCACA-3′
Sulpyr-p06	5′-TACTCTACTTTTCAACATTCTTCA-3′
SSV2attAp01	5′-GTGTTCTACCTTTTCCACAGTC-3′
SSV2attA'p02	5′-TGGGTACGTCATTTATTGATCTT-3′
SSV2attP'p03	5′-GCTTTTATGCAGTTATTGCTTT-3′
SSV2attPp04	5′-GGGAAATGATTAGATGGTCTTC-3′
SSV2Seq2p05	5′-ACTCTTAATATACCCACTGTTGGG-3′
SSV2Seq2'p06	5′-GTATATAACCCTTGGCTTCTTGTT-3′

SSV plaque assay

The virus titre of a culture was determined by means of plaque assay using *S. solfataricus* P2 as an indicator strain according to Schleper et al. (1992). The plates were incubated for 2–3 days at 80°C and examined for the appearance of the turbid halos which the inhibition of host growth by SSV2 virus generates.

DNA isolation and characterization of the SSV2 integration into the genome of *S. islandicus* REY15/4

The Plasmid DNAs of SSV2 and pSSVx were prepared from *Sulfolobus* cells using the QIAGen Spin Miniprep kit. Total DNAs were isolated from 10–15 ml of *Sulfolobus* cultures. Summing up, cell pellets were re-suspended in 400 µl of the lysis buffer (50 mM Tris–HCl, pH 7.5, 50 mM EDTA and 0.2% N-Lauroyl Sarcosine) and incubated at 55°C in the presence of proteinase K (380 µg/ml) for 2 h. To extract the samples we used phenol twice, phenol–chloroform once and, finally, chloroform. Thereafter, the DNAs were precipitated using ethanol.

To characterize the integration of SSV2 into the *Sulfolobus* genome, total DNA was digested using *Sal*I, *Eco*RV, *Bgl*II, *Hin*dIII or *Bam*HI as restriction enzymes and the resulting DNA fragments were analysed by Southern blot and hybridization, according to the standard procedures (Sambrook and Russell 2001). Two adjacent similar-size fragments (*Bgl*II–*Bgl*II and *Bgl*II–*Sal*I) encompassing each a 4,689 bp sequence of the SSV2 genome were obtained by digesting with *Sal*I and *Bgl*II and used as probe. Thereafter, the Random Prime DNA Labelling Kit (Roche Applied Science) was used to label these DNA fragments with radioactive [α-^{32}P]dATP as described by the manufacturer. Unincorporated [α-^{32}P]dATP was removed from the labelled probe by gel filtration, using the Nick Columns (Amersham Biosciences).

Identification of the SSV2 integration site by PCR

The primers designed for the PCR identification of SSV2 integration are listed in Table 1. The PCR reactions contained ca. 100 ng of the template DNA, 1 U of Taq DNA polymerase, 1 µM of each primer, 1 µM of dNTP and the buffer needed (final concentration of Mg^{2+} was 1.5 mM) for 30 cycles of amplification (60 s at 95°C, 60 s at the appropriate annealing temperature and 60 s at 72°C) with initial denaturation and final elongation for 5 min at 95 and 72°C, respectively. The amplified

products were analysed by means of Agarose gel electrophoresis and photographed.

Estimation of cellular SSV2 and pSSVx content during host growth

To visualize SSV2 and pSSVx DNA accumulation during the growth of host cells, total DNAs were extracted from the samples collected at different incubation stages (as indicated in the individual experiments) from the cultures of *S. islandicus* REY15/4, REY31A, SSV2-P2 and SSV2-REY15A. Total DNAs were isolated from the same amount of the cells (normalized by OD_{600}) using the procedure described above. Similar amount of DNAs (estimated by OD_{260} and ethidium bromide fluorescence) was used for restriction digestion and Southern blot analysis.

Normalization of the loaded DNAs was performed using the *alba* gene (SSO0962, one copy per genome) or the correspondent PCR product amplified from the genome of *S. islandicus* for the DNAs of infected *S. solfataricus* and *S. islandicus* REY15/4, respectively.

To determine the SSV2 and pSSVX copies per chromosome, three different sequences, i.e. the *alba* gene, the above mentioned 4.69-kb SSV2 (or the coding region of *vp3* gene) and a *Hin*dIII 1.7-kb pSSVx fragments, were ^{32}P-radiolabelled by random priming, yielding to a specific activity of ca. 10^9 cpm/μg DNA. The same amounts of the radiolabelled probes (1×10^6 cpm/ml for each) were mixed and used for Southern hybridization, according to the standard procedure (Sambrook and Russell 2001).

For the experiments illustrated in Figs. 4 and 7, the host–SSV2 hybrid fragment derived from SSV2 integration into the host genome (one copy per genome) was used for normalization in place of *alba* gene.

Since only one copy of the *alba* gene (or the host–SSV2 hybrid fragment) is present in *Sulfolobus* genomes, SSV2 and pSSVx copy numbers were determined by the ratios of extrachromosomal-derived hybridization signals to chromosomal-derived ones which were quantified using a phosphoroimager and the Quantity One 4.2 software (Bio-Rad).

SSV2 copy number has been determined for the samples of REY31A, REY15/4 and SSV2-harbouring *S. solfataricus* from at least three independent growth experiments.

Results

Site-specific SSV2 integration into host chromosomes

The *Sulfolobus* SSV2 virus encodes a tyrosine integrase (Esposito and Scocca 1997; Nunes-Duby et al. 1998) of the SSV1 integrase family (She et al. 2002; Stedman et al. 2003). To establish if SSV2 generates provirus in the natural host, the episomal and the genomic DNAs of the REY15/4 strain were prepared from an exponentially growing culture and analysed by Southern hybridization using the 4.7 kb SSV2 fragment as a probe (see Materials and methods). As expected, in REY15/4 cells the SSV2 genome was found to be present in an episomal form. However, the fact that the SSV probe detected an additional host chromosome fragment (see arrowheads in Fig. 1a) indicated that SSV2 generates provirus in the natural host. The host–virus hybrid fragments which correspond to the provirus and were obtained by digestion with *Eco*RI, *Eco*RV and *Sal*I contain the left attachment site, *attL*, while the only host–virus hybrid fragment from the digestion of *Bgl*II contains the right attachment site, *attR* (Fig. 1b). Because of the probe used, *Hin*dIII digestion does not allow the detection of the host–virus hybrid fragments. When the membrane used was hybridized with a pSSVx-specific probe we found that pSSVx was only present in an episomal form and did not integrate into the host chromosome. This result is consistent with the fact that pSSVx does not encode an integrase.

It has been reported that the putative SSV2 *attP* site is located within its *int* gene (Stedman et al. 2003) as for the SSV1 virus. While searching the complete genome of the *S. solfataricus* P2 (She et al. 2001a, b) with the sequence of SSV2 *attP* region using the BlastN tool (Altschul et al. 1990), we found two sequence matches denoted as *attA* and seq2, each overlapping a tRNA^{Gly} gene in the downstream halves of the tRNA genes. The only mismatch was in the anticodons (Fig. 2a). Hence we were interested in establishing if the two tRNA^{Gly} genes would act as *attA* sites for SSV2 in the *S. solfataricus* genome.

A PCR approach was employed to establish the integration site of SSV2 into *S. solfataricus* P2 chromosome. Three pairs of PCR primers were designed for use in identifying the putative host–virus hybrid fragments of the SSV2 provirus in each tRNA^{Gly} gene (Materials and methods, Table 1). When genomic DNA prepared from two single clones of SSV2–P2 cells was used as a template, host–virus hybrid attachment sites (*attL* and *attR)* of the predicted sizes were amplified at the tRNA^{Gly} (CCC), but not at the "seq2" site (Fig. 3). Identical results were obtained with the other six clones analysed (not shown). These results suggest that the tRNA^{Gly} (CCC) gene serves as the only *attA* for SSV2 in the *S. solfataricus* chromosome. SSV2 integration into the *S. solfataricus* genome was also investigated using the Southern hybridization procedure and the same probe, as described for SSV2 integration into the natural host. Our finding was that the sizes of the integrated fragments were those predicted assuming the tRNA^{Gly} (CCC) gene to be the attachment site (Fig. 2c), thus confirming that SSV2 integrates only at one site in the *S. solfataricus* genome. The SSV2 integration model is shown in Fig. 2b.

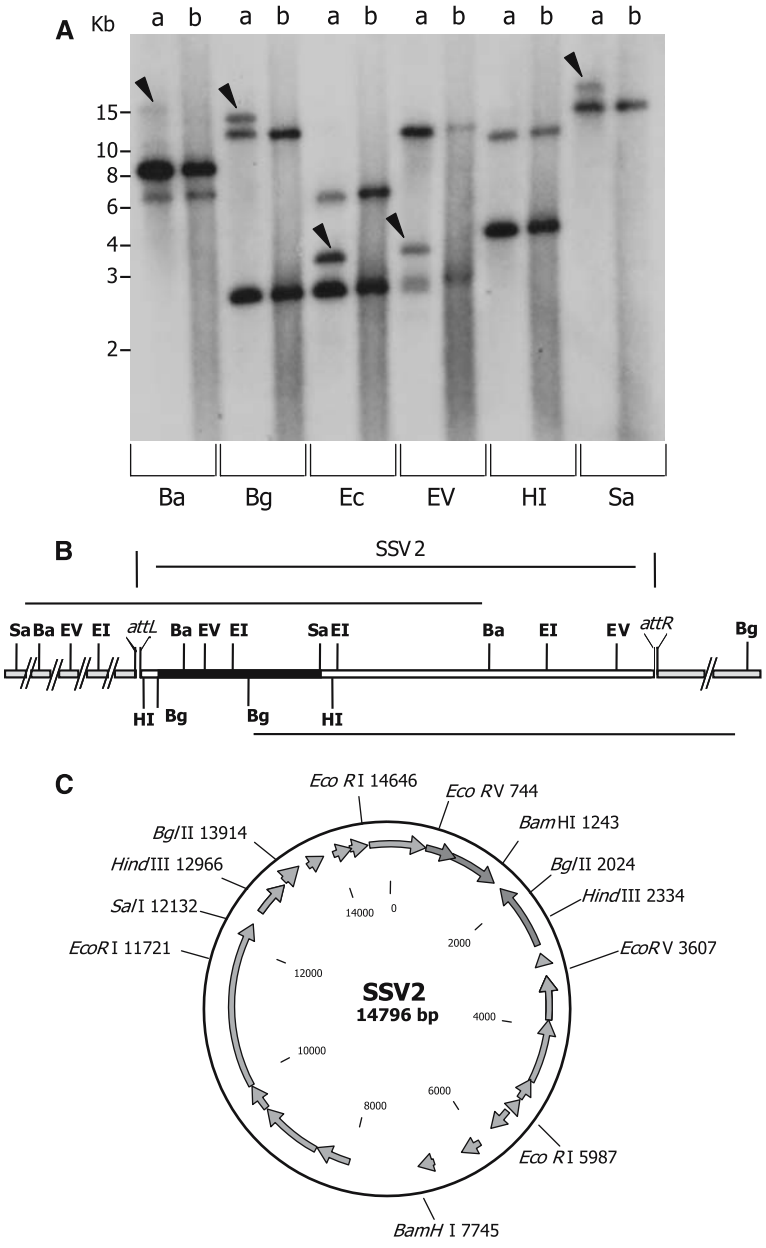

Fig. 1 Identification of the SSV2 provirus in the natural host. **a** Southern analysis of the SSV2 integration. Genomic DNA ("*a*" lanes) and plasmid DNA ("*b*" lanes) prepared from an exponentially growing culture of REY15/4 were digested using the restriction enzymes indicated *below*. The sizes of DNA molecular markers (kb) are reported on the *left*. *Arrowheads* appear beside host–virus hybrid fragments of the SSV2 provirus. **b** An illustration of host–virus hybrid fragments. Host–virus hybrid fragments containing the attachment site *attL* or *attR* are illustrated by *lines*. The 4.69 kb SSV2 fragment used as a probe for Southern hybridization is filled *black*. Restriction sites are schematically indicated on both SSV2 and host genomes. *attL* and *attR*: left and right attachment sites of the provirus. Restriction enzymes used: Bg, *Bgl*II; EI, *Eco*RI, EV, *Eco*RV; HI, *Hin*dIII; Sa, *Sal*I. **c** Restriction map of SSV2 genome. The positions of the restriction enzymes used for the analysis of the SSV2 integration into the genome of REY15/4 are indicated

Accumulation of SSV2 and pSSVx DNA during the growth of *S. islandicus* REY15/4

At an early stage of our experiment, we prepared episomal SSV2 DNA using approximately equal amounts of *S. islandicus* REY15/4 cells from different cultivation batches. In some experiments, we obtained a large yield of episomal SSV2 and pSSVx DNAs; in others, a fairly small amount of plasmid DNAs. To account for this finding, we prepared episomal and total DNA using REY15/4 cells at different optical densities. We used Southern hybridization and the 4.69-kb SSV2 as well as the *Hin*dIII 1.7-kb pSSVx fragments as probes (see Material and methods), to identify the episomal SSV2 and pSSVx. Figure 4 shows that whereas in early growth phases (0.5–0.9, OD_{600}) episomal SSV2 and pSSVx DNA was produced at a low rate, large amounts of SSV2 and pSSVx DNA accumulated at maximum cell density (1.25–1.3 OD_{600}). This suggests that there was an induction of SSV2 and pSSVx replication during the growth of the natural host.

Furthermore, multiple topological forms of the SSV2 and pSSVx episomal genomes were observed after the induction of viral replication (Fig. 4, 1.25–1.3 OD_{600}), consistent to the results obtained from the SSV1 studies (Nadal et al. 1986).

Characterizing the SSV2 and pSSVx replication induction process

A typical growth curve of *S. islandicus* REY15/4 is reported in Fig. 5a and sample taking indicated. In a basal medium which contained tryptone, yeast extracts and

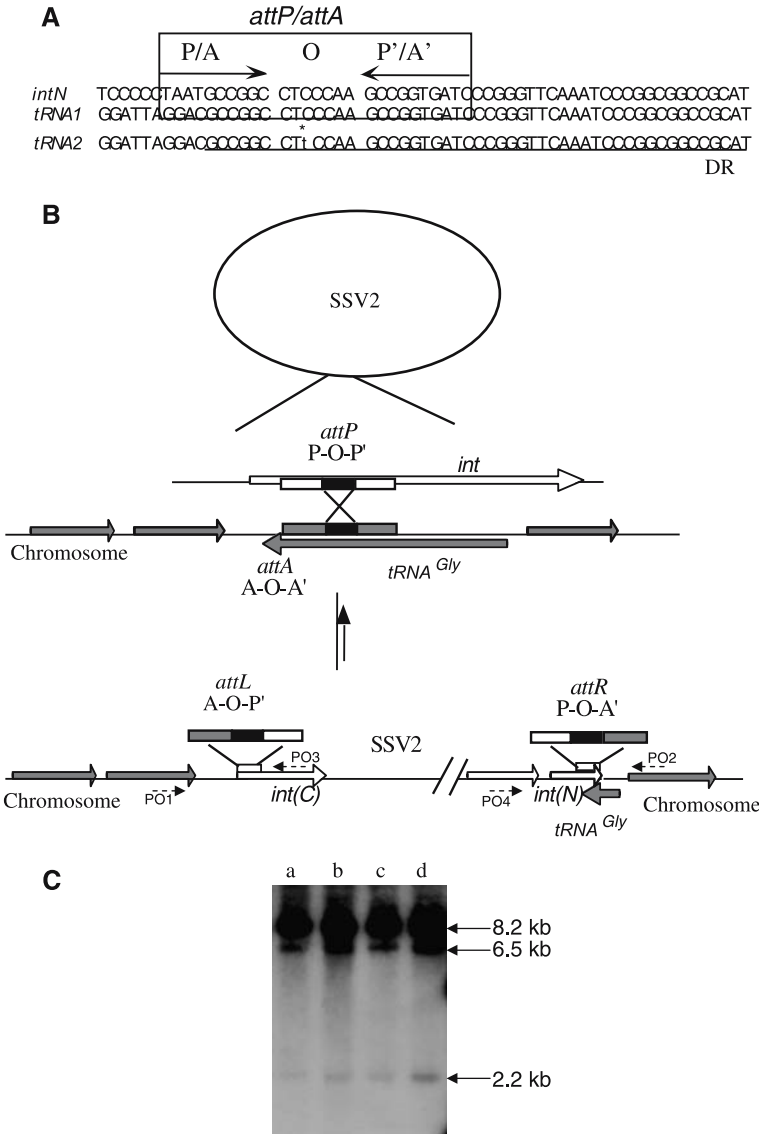

Fig. 2 SSV2 integration in the foreign host. **a** Composition of the SSV2 integration sites. The boxed motifs correspond to the minimal attachment sites identified for the SSV1 integrase in an in vitro assay (25). *attP* (P-O-P′): virus integration attachment site; *attA* (A-O-A′): host integration attachment site. *Lines with arrows* indicate imperfect inverted repeats (A, A′, P and P′) which constitute putative binding sites for the integrase. "O" indicates the overlap region where DNA recombination occurs. The tRNA2 sequence with only one mismatch to the *attA* site, $tRNA^{Gly}$ (UCC), has been included for comparison. In the report of the PCR assay this site is named "seq2" (see Materials and methods). DR bordering the provirus is *underlined*. *DR* direct repeat, *tRNA1* $tRNA^{Gly}$ (CCC), *tRNA2* $tRNA^{Gly}$ (UCC). **b** A graphic presentation of SSV2 integration/excision. *Unfilled* ORFs are those of the SSV2 genome, *shaded* ORFs are those of the *S. solfataricus* P2 genome. The composition of all SSV2 attachment sites is illustrated. The *non-filled box* represents P or P′; the inverted repeat of the *attP*, *box filled black* indicates the overlap region "O", and the *shaded box* represents A or A′, the inverted repeat of the *attA; int* is the integrase gene, *attL* (A-O-P′) and *attR* (P-O-A′) are the left and right host–virus hybrid attachment sites, respectively. The positions of the PO1 and PO2 primers on the host chromosome and of PO3 and PO4 primers on the SSV2 genome are schematically indicated. **c** Southern analysis of the SSV2 integration into the genome of *S. solfataricus*. Genomic DNAs prepared from four independent SSV2-infected *S. solfataricus* clones were digested with *Bam*HI. The 8.2 and 6.5 kb fragments correspond to the episomal SSV2, whereas the 2.2 kb sizes is the host–virus hybrid fragment of the SSV2 provirus

sucrose, REY15/4 grew in a generation time of about 4.5 h as long as the OD_{600} value was below 0.4. This growth rate slowed down slightly at higher OD_{600} values and stopped altogether when the OD_{600} value rose to 1.3 approx. Interestingly, the optical density of the cultures remained fairly constant for at least 20 h and no apparent cell lysis occurred in that period. In contrast, the pH of the culture changed very little at cell densities (OD_{600}) below 1.3; thereafter it gradually increased to pH 5.5 in 20 h (Fig. 5a).

To shed further light on the increase in SSV2 and pSSVx observed during the growth of the natural REY

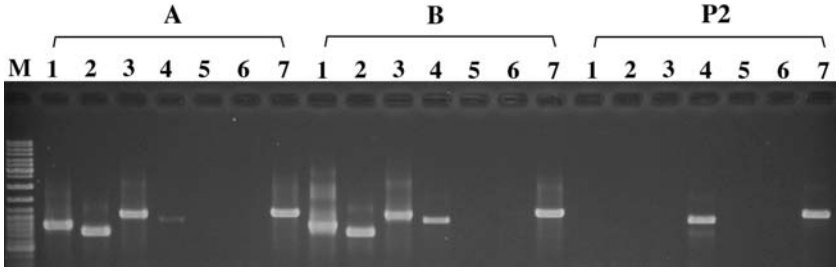

Fig. 3 PCR amplification of SSV2 integration attachment sites. Total DNAs were prepared from two independent SSV2-containing *S. solfataricus* clones (*A* and *B*) and an uninfected *S. solfataricus* P2 (*P2*) and used as templates for PCR. For each sample, PCR products containing *attP* (the SSV2 viral integration site, 912-bp) and the host integration site *attA* (968-bp) or a sequence motif which is very similar to *attA* (named seq2) were loaded in *lanes 1, 4 and 7*, respectively; *lanes 2 and 3* are the PCR amplifications of the virus–host hybrid fragments of *attL* and *attR*, with the sizes of 722 and 1,158-bp, respectively; *Lanes 5, 6* are PCR amplifications for the virus–host hybrid fragments at the seq2 site (predicted to be 1,195 and 843 bp, respectively). *M*: DNA kb ladder. Identical results were obtained with the other six SSV2-containing *S. solfataricus* clones analysed

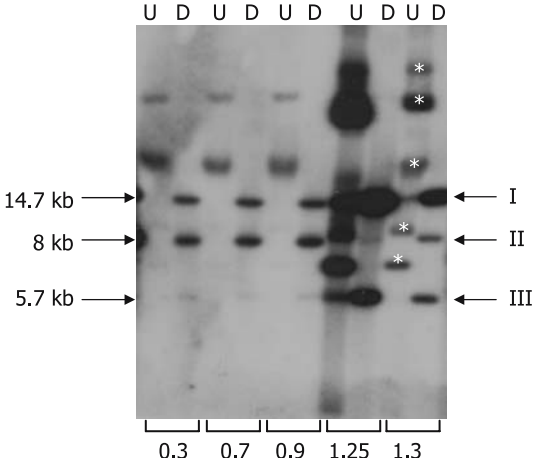

Fig. 4 Content of the virus DNAs in REY15/4 cells of different optical densities. The OD_{600}values of the *S. islandicus* REY15/4 cultures used for preparing the genomic DNA are reported at the bottom. Lanes labelled with "*D*" indicate the DNA samples digested with *Cla*I restriction enzyme (one cut for both SSV2 and pSSVx genomes). Lanes marked with "*U*" indicate undigested total DNAs. The linear SSV2 and pSSVx are indicated. "*I*" indicates the linearized episomal SSV2 (14.7 kb), "*II*" indicates the integrated SSV2 (8 kb) and "*III*" indicates the linearized episomal pSSVx (5.7 kb). The amount of total DNAs at 1.25 and 1.3 OD_{600} values was respectively ten and fivefolds lower, than that used for samples at 0.3–0.9 OD_{600}. The *asterisks* indicate the multiple topological forms of the SSV2 and pSSVx episomal genomes observed after the induction of viral replication

host we collected REY15/4 cell samples at 2-h intervals during incubation (see Materials and methods). Over the first 24 incubation hours, no episomal SSV2 and pSSVx bands were visible in the total DNAs prepared from rapidly growing REY15/4 cells; thereafter, episomal SSV2 and pSSVx DNA accumulated rapidly reaching a maximum within 4 h, i.e. between the 26th and 30th incubation hours (Fig. 5b). This indicates that the increase of SSV2 and pSSVx must be traced to the induction of SSV2 and pSSVx replication which sets in 4 h after the growth of host cells came to a stop.

The extent of SSV2 and pSSVx replication induction was determined via Southern blot from three independent cultures and the results of a typical experiment are shown in Fig. 6. Since there is only one copy of *alba* gene in *S. islandicus* REY15/4 genome, the amount of SSV2 was measured at every growth stage as the ratio episomal/chromosomal radioactive signals via a phosphorimager (see Materials and methods). The episomal SSV2 was produced at a rate of just one copy per chromosome approx during the exponential growth phase and a 50-fold increase (on average) occurred within 4 h upon the growth stop (Fig. 6b).

The amount of pSSVx was under the limit of Southern blot detection between the 16th and 26th hours, therefore a precise evaluation of pSSVx increase was not possible. However, the pSSVx genome was induced in a growth-dependence fashion, similarly to SSV2. At the latest sample-taking points analysed (30th–32nd hours, Fig. 6) the estimated copy number was ca. 2–5 copies per cell. Moreover, after repeating this experiment several times, we found that the copy number of pSSVx in a rapidly growing REY15/4 varied from undetectable to detectable by Southern analysis. Nevertheless, there is a general trend of increase of pSSVx copies upon viral replication induction (Figs. 4, 5b, 6).

It was previously shown (Arnold et al. 1999a; Stedman et al. 2003) that SSV2 replication is not dependent on pSSVx in *Sulfolobus solfataricus* P1 cells. We were interested in establishing if pSSVx contributes to the SSV2 replication induction observed in the natural host REY15/4. To obtain a pSSVx-free strain for investigation, a REY15/4 culture was left to grow continuously for approx. 100 generation times at an optical density below OD_{600} 0.4. Thereafter, we obtained single colonies via Gelrite plating and one of these, REY31A, was used to investigate the physiology of SSV2 replication, after checking carefully the absence of pSSVx (see Materials and methods).

Figure 7a shows that the growth characteristics of REY31A were similar to REY15/4. Moreover the analysis of SSV2 DNA content performed by quantification of the radioactive signals after Southern blot, revealed that the induction of SSV2 takes place in

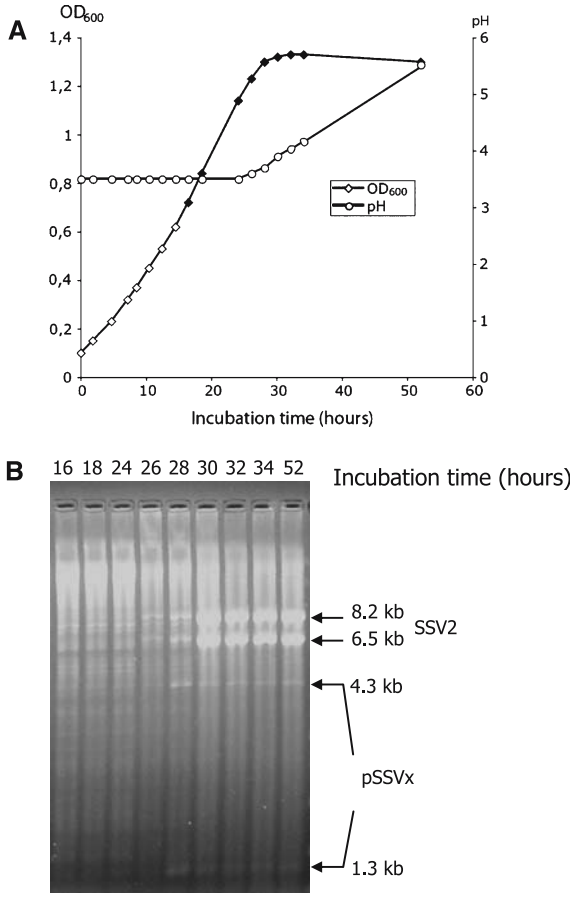

Fig. 5 Induction kinetics of SSV2 and pSSVx replication. **a** Growth characteristics of the *S. islandicus* REY15/4. During incubation, the OD_{600} (on the first *Y* axis) and the pH (on the second *Y* axis) of the culture were measured and plotted versus the incubation time. At the same time, samples were also taken for preparing total DNA. **b** Relative content of the episomal SSV2 and pSSVx genomes in the REY15/4 cells during incubation. The genomic DNA was prepared from REY15/4 samples indicated by *solid cycles*, digested with *Bam*HI, and analysed by agarose gel electrophoresis. The agarose gel was stained with ethidium bromide and a picture was taken using a Kodak system. Two *Bam*HI fragments of the episomal SSV2 and pSSVx are indicated with their sizes

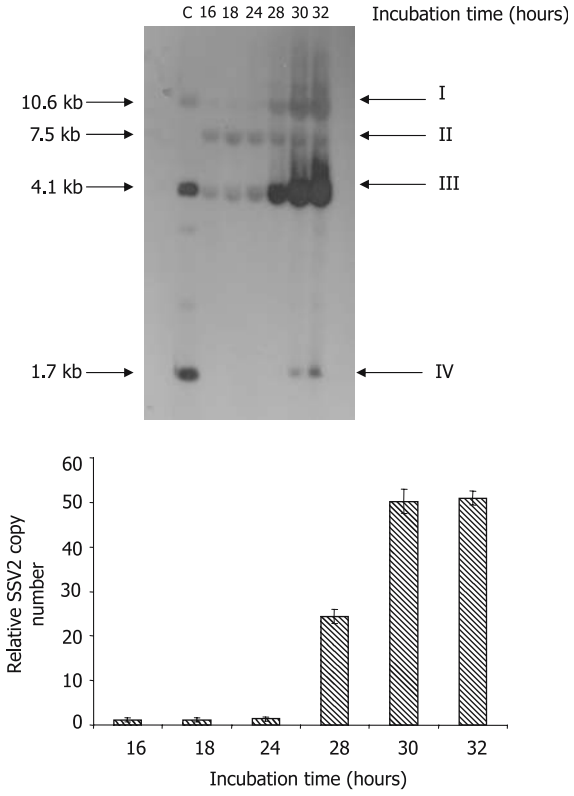

Fig. 6 Estimate of the SSV2 induction level in REY15/4 cells. **a** Total DNAs of REY15/4 *Sulfolobus* cultures were prepared at the indicated incubation times and analysed by *Hind*III restriction digestion and Southern hybridization. *Lane C* was loaded with *Hind*III digests of SSV2 and pSSX episomal DNAs, as positive controls. A mix composed of ^{32}P-radiolabelled *alba* gene (SSO0962 amplified from the *Sulfolobus* REY15/4 genome), the *Bgl*II and *Sal*I fragments of the SSV2 genome and a *Hind*III-pSSVx 1,7 kb fragment (see Materials and methods), were used as probes. The sizes of the radioactive signals of SSV2 (*I and III*), pSSVx (*IV*) and the chromosomal restriction fragment containing the *alba* gene (*II*) are reported on the *left*. The SSV2 probe used does not allow the detection of the integrated SSV2 host–virus fragment (see Fig. 1a, b). **b** The SSV2 copy number in REY 15/4 cells was calculated as the ratio of the radioactivity counts (cpm) of the SSV2 hybridized signal at a specific growth stage to that of the *alba* gene (see Materials and methods). The *graphic bars* show the increase of SSV2 copy number and the *error-bars* represent the standard deviations for five independent experiments

REY31A cells with the same modality and at the same extent described for REY15/4 cells (Fig. 7b). These finding indicate that the presence of pSSVx is dispensable to the SSV2 replication induction process.

To establish if SSV2-determined growth inhibition is a reversible process, we carried out a dilution experiment, i.e. we released the REY15/4 host cells into fresh medium at a 10–20:1 ratio after the virus induction process had set in. The growth rates and quantities of SSV2 DNA in the diluted cultures were estimated as described in Materials and methods. Our results showed that host cells grow exponentially after a lag phase of varying length (Fig. 8a) and that thereafter the average SSV2 cellular content reverts to ca. one copy per chromosome, as shown in the Southern blot of Fig. 8b.

Virus packaging capacity during host growth

To decide if SSV2 DNA accumulation was caused by the lower virus packaging capacity of the host, we also measured the SSV2 virus titre of the REY15/4 culture. The virus particle titre was found to be fairly low (2–4×10^3 cfu/ml) in the exponential growth phase but rose tenfold after the host had entered the late exponential growth-phase (see below). Thereafter, it remained approx. 1×10^5 until the latest sample-taking points. These results indicate that when SSV2 replication is induced a concurrent, consistent packaging and extrusion of the virus particles occurs and that the fact that free episomal SSV2 DNA accumulates in the cytoplasm must be traced back, not to any reduced viral DNA packaging

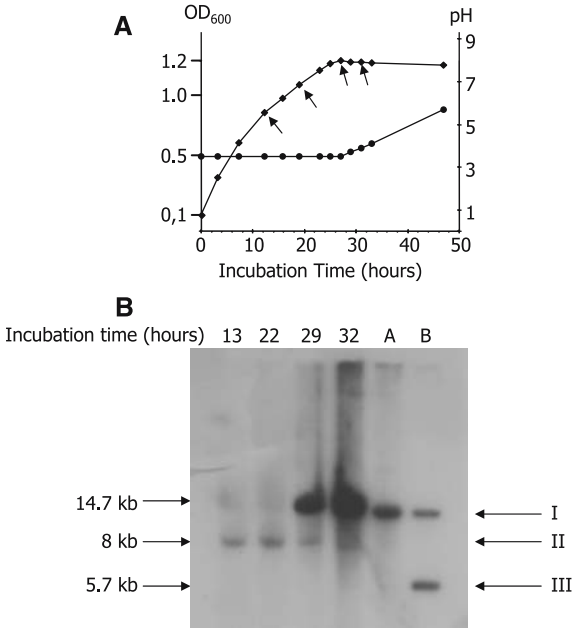

Fig. 7 Analysis of the growth characteristics of REY31A and evaluation of SSV2 induction level. **a** Growth characteristics of the *S. islandicus* REY31A. OD_{600} and pH values of the culture were measured and plotted versus the incubation time. **b** Total DNAs from equal amounts of REY31A cells (normalized by OD_{600}) were prepared at the incubation times reported at the *top* and indicated by *arrows* on the growth curve of panel **a**. DNAs were digested with *Cla* I (one single cut for both SSV2 and pSSVx genomes) and the probes used are the SSV2 4,69 kb and a pSSVx-*Hin*dIII 1,7 kb restriction fragments, this latter included to demonstrate the absence of pSSVx. The size of SSV2 (*I*), pSSVx (*III*) and the host–virus hybrid fragment (*II*) are indicated on the *left*. *Cla*I digests of purified SSV2 (*lane A*) and SSV2 together with pSSVx (*lane B*) were loaded onto the same membrane as positive control for hybridization. The radioactive signal of the host–virus hybrid fragment was used to normalise the amount of the DNA loaded and to determine the SSV2 amount at every growth stage (see Materials and methods)

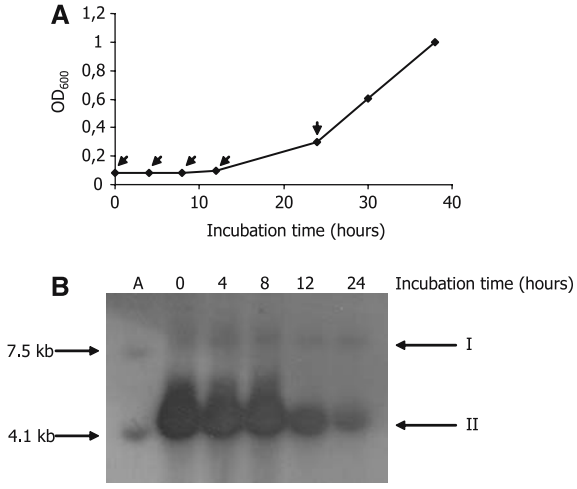

Fig. 8 The SSV2 induction process is reversible. **a** *S. islandicus* REY15/4 cells were diluted in fresh medium at a 10–20:1 ratio after the virus induction process had set in. OD_{600} values were plotted versus the incubation time. Samples taken for preparing total DNAs are indicated by *arrows*. **b** Total DNAs from equal amounts of REY15/4 cells (normalized by OD_{600}) were digested with *Hin*dIII and the membrane hybridized with the coding region of the *vp3* gene (13,652–13,930 bp on the SSV2 genome) and the REY 15/4 *alba* gene, as probes. The sizes of the radioactive signals of the chromosomal restriction fragment containing the *alba* gene (*I*) and the episomal SSV2 (*II*) are reported on the *left*. *Lane A*: Total DNA from REY15/4 cells, before the SSV2 replication had set in. *Lanes 1–5*: Total DNAs from REY15/4 cells once SSV2 replication induction had set in and harvested at different incubation hours after dilution into fresh medium (see *arrows* in **a**). A progressive decrease of episomal SSV2 content occurs within 24 h after dilution of SSV2-induced REY15/4 cells into fresh medium

and/or exportation capacity, but rather to a higher replication rate of SSV2.

SSV2 replication in *S. solfataricus* P2

To investigate host–virus interaction between SSV2 and a foreign host, an SSV2 infected *S. solfataricus* P2 culture was revitalized in liquid medium from areas of inhibition growth plaques (see Materials and methods). As it is shown in Fig. 9a cells grew up with a generation time of about 9–10 h after a lag of ca. 20 h and reach the maximum value of ca. 0.6 OD_{600}. However, when single colonies obtained by this liquid culture after two subsequent cycles of streaking and tooth-picking were made to grow in parallel with uninfected cultures, the virus-containing and virus-free strains showed the same growth characteristics throughout cultivation, indicating that once established in a foreign host SSV2 does not affect the growth of this host. SSV2-P2 single clones isolated from the plate, did not contain the pSSVx (not shown).

Southern blot analysis performed on total DNAs of a SSV2-P2 isolate, using the *vp3* coding region of SSV2 genome and the *alba* gene as probes, showed that the SSV2 content in infected *S. solfataricus* cells remained fairly constant throughout the growth. This demonstrates that SSV2 replication correlates with the growth rate of *S. solfataricus* and that the replication induction observed in the natural host does not occur in the foreign host (Fig. 9b).

The virus particles production, determined throughout SSV2-P2 cultivation, starts during the exponential growth, reaching its maximum peak (1.1×10^5 cfu/ml) before cells approach to the stationary phase and eventually decreases.

The mechanisms governing the induction of SSV2 replication

Compared to virus-free *Sulfolobus* cultures, for REY15/4 cells (Fig. 5a) we always observed a sudden halt in growth and a substantial difference in final optical density values (typically 1.3 vs. 2.5, OD_{600}), i.e. a sign pointing to growth inhibition. Since growth inhibition

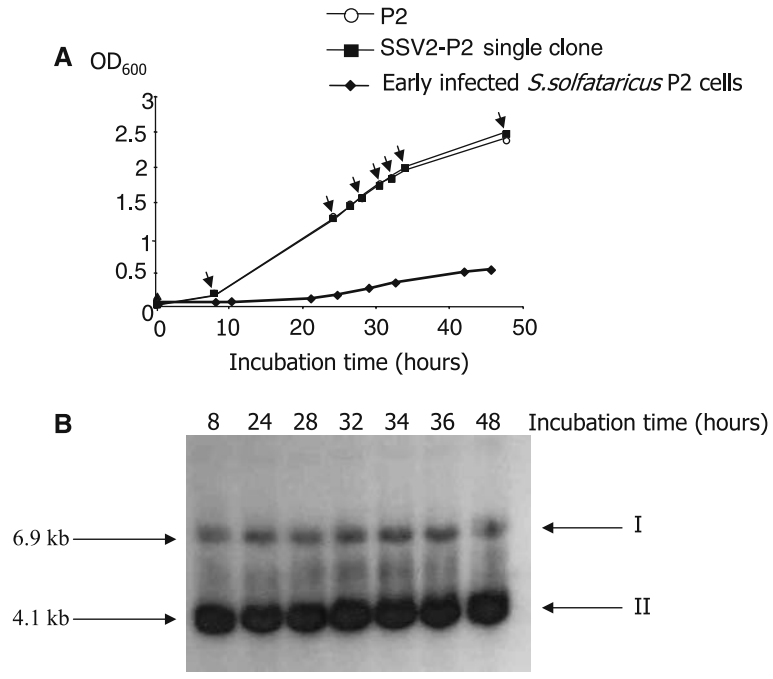

Fig. 9 Characterising *S. sofaltaricus* P2 cells containing SSV2 virus. **a** Comparison among growth rates of uninfected *S. solfataricus* P2 (*filled circle*), early infected *S. solfataricus* P2 cells (*filled diamond*) and a single clone of SSV2-harbouring P2 (*filled square*). **b** Southern blot analysis of total DNAs extracted from a SSV2-containing *S. solfataricus* single clone at the incubation hours indicated in (**a**). DNAs were digested with *Hin*dIII and the membrane hybridized with the coding region of the *vp3* gene (13,652–13,930 bp on SSV2 genome) and the *alba* gene (SSO0962), as probes. The size of the host genome fragment containing the *alba* gene (*I*) and SSV2 radioactive signal (*II*) are indicated on the *left*. The *Hin*dIII digestion and the probe used do not allow the detection of the integrated SSV2 host–virus fragment (see Fig. 1a, b)

coincided with the onset of the viral DNA replication induction, we resolved to investigate the mechanisms governing these two processes and the relations between them. For this purpose we isolated a *S. islandicus* strain (which is closely related to the REY15/4 strain) and obtained virus-harbouring strain (SSV2-REY15A) as described in Materials and methods. Figure 10 shows that the three SSV2-REY15A cultures grew in the same manner as the REY15/4 strain. However, in the infected cultures, growth came to a stop at 1.3 (OD_{600}), whereas the uninfected REY15A culture continued to grow for 20 h reaching cell density of 2.0–2.5 (OD_{600}). Analysing the total DNA in the SSV2-infected REY15A cells during growth, we found that SSV2 replication sets in within 4 h after the end of cell growth, as in the natural host of SSV2.

To establish whether the uninfected REY15A was actively growing beyond 1.3 OD_{600}, when we observed the sudden halt of SSV2-REY15A cells, we took nine cell samples between 1.1 and 2.2 OD_{600}, determined the colony formation units (CFU) by Gelrite plating (see Materials and methods) and found a linear relationship between the CFUs and OD_{600} values of REY15A cells (not shown). Moreover, the CFU/OD_{600} fell into the range of $7.29 \pm 0.66 \times 10^8$ indicating that the *Sulfolobus* cells were still able to divide when the OD_{600} of a REY15A culture increased up to 2.2 OD_{600}. Thus, the sudden halt of the optical density increase of SSV2-REY15A cultures reaching 1.3 OD_{600}, reflected growth inhibition occurring in actively dividing culture at a growth stage corresponding to an exponential growth phase.

Since REY15A and REY15/4 are very closely related to each other, we inferred that the same would be applicable to the natural host REY15/4, i.e. growth inhibition occurred at a late exponential growth phase and REY15/4 cells would continue to grow and reach a higher cell density level (as for REY15A) if it was cured of SSV2.

Discussion

SSV provirus formation is furthered by virus-encoded integrases of the tyrosine integrase family (Esposito and Scocca 1997; Nunes-Duby et al. 1998). Like most integrative viruses, bacteriophages and plasmids, the *Sulfolobus* virus SSV2 integrates into host chromosomes at the host integration attachment sites located at the 3′ end of the tRNA (CCC) gene (Campbell 1992; Reiter et al. 1989; She et al. 2002; Williams 2002). SSV2 integration into host chromosomes is highly sequence-specific: integration occurs only at the sequence-identical *attA* sites of *S. islandicus* REY15/4 and *S. solfataricus* P2, and it does not occur at the "seq2" site of the $tRNA^{Gly}$ (UCC) gene of the latter because of a sequence mismatch. In comparison, SSV1 integration into the genome of *S. solfataricus* did occur

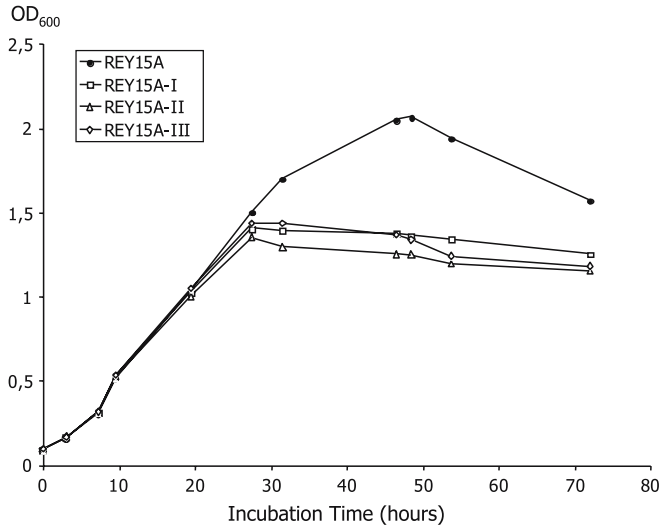

Fig. 10 Coupling of host growth inhibition and SSV2 replication induction in REY15A. *S. islandicus* REY15A and its three SSV2-containing strains, REY15A-I, II, III were grown in a TYS medium. The results were the same in three experiments carried out in succession

at the sequence of the foreign host that exhibited one mismatch versus *attA* of the natural host (Schleper et al. 1992) although it turned out that a sequence-identical SSV1 *attA* site is absent from the complete genome of *S. solfataricus* P2 (She et al. 2001b). Apparently, it remains to be studied whether SSV2 would integrate into the "seq2" site when the perfectly matched target is removed from the foreign host. Thus, further comparative studies of all known SSV integration systems will help shed light on the factors underlying the site-specificity of this viral integration.

Infecting *Sulfolbus* species with SSV viruses always retards host growth during which a large amount of virus particles has been produced, leading to the formation of the inhibition spots (plaques) on solid medium. *Sulfolobus* SSV2 virus is no exception to this rule (Stedman et al. 2003). However, once settle in the foreign *S. solfataricus* host, SSV2 no longer inhibits host growth and this is in strict contrast to what is seen in early infected cells. Neither does SSV2 genome exceed over a few copies per cell throughout cultivation. Consequently, there is a major difference in growth rates between the early SSV2-infecting culture revitalized from a plaque and SSV2-harbouring strain of *S. solfataricus*: the former showed a greatly retarded growth and the latter exhibited a similar rate of growth as an uninfected control. This suggests that a series of interactions between SSV2 and its foreign host will eventually lead to a co-existence harmony between them as seen for the SSV2-carrying *S. solfataricus* strain. It is thus very interesting to investigate whether this conversion of host response is a general scheme for the host defence of *S. solfataricus* to viral infection and to study the molecular mechanisms governing it.

Another interesting feature of SSV2–host interaction lies on the fact that SSV2 DNA replication is characterized by a physiological induction in the natural host. Viral genome copy number increases up to 50-fold within 4 h. We have shown that growth inhibition of the host concurs with the virus replication induction and the inhibition takes place at a late exponential growth phase. Moreover, the inhibition effect is a reversible. Upon being released into a fresh medium, the SSV2-induced *Sulfolobus* culture resumes their exponential growth and host regains control over SSV2 replication. Whereas it appears that the replication induction must have been turned on by an unknown internal signal(s), either originated from virus or host, it is not clear if the growth inhibition constitutes part of host self-defence mechanism or it is conferred by a signal related to the SSV2 replication induction and/or another SSV2 signal. Nevertheless we suggest that SSV2–host interaction may perform a dual function: (1) 'freezing' host metabolism into an inactive state under unfavourable or detrimental environmental conditions; (2) enabling the cells to react more quickly to changes in their ecological niches and, thus, to 'resume' their metabolic activity when favourable conditions are restored in the environment. This points to a physiological role of the SSV2 virus in host survival and to a mutually beneficial interaction (like those of some bacteriophage-host systems), which increase the host's metabolic activity and/or environmental fitness (Ackermann and DuBow 1987; Weinbauer 2004). Moreover, this kind of interaction is specific to REY strains but not to *S. solfataricus*, indicating that the specific interaction of SSV2 with its natural host is a prerequisite to initiate the growth-phase regulated virus replication process. Clearly, there is interesting molecular biology behind these phenomena and further investigation should reveal mechanisms governing host–virus interactions as well as co-evolution of viruses and their hosts in archaea.

In addition to the SSV viruses, another temperate archaeal virus, *Acidianus* two-tailed virus (ATV), has recently been described (Prangishvili et al. 2006b). As for the SSV1 virus, ATV infection of the cells growing under optimal conditions produced lysogenization and the lysogeny can be interrupted by adding mitomycin C or exposing to ultraviolet radiation. Furthermore, ATV infection can also be induced by cold-shock (Prangishvili et al. 2006b). Nevertheless, the induction of SSV1 and ATV involves only external factor(s), and by contrast SSV2 replication induction is induced by endogenous signal(s). Like most crenachaeal viruses, the induction of SSV2 replication does not lead to the lysis of host cells. By contrast, the ATV virus causes host lysis under environmental stress, constituting the first example of lytic virus ever described for crenarchaea (Prangishvili et al. 2006b). Thus, these viruses provide models for studying mechanisms and evolution of virus infection in archaea.

Sulfolobus virus study was initiated to identify suitable genetic elements for constructing host-vector systems and SSV1 virus has been successfully used for this

purpose. Several genetic systems have been reported including most recent developments in gene reporter, protein expression and gene expression systems (Jonuscheit et al. 2003; Albers et al. 2006; Lubelska et al. 2006). Thus, one of the next targets in archaeal virus study is to conduct systematic characterization on selected host–virus systems using genome microarray and genetic knockout as well as genetic complementation technologies to reveal principles of archaeal molecular biology.

Acknowledgements We thank Roger Garrett for encouraging discussion, Wolfram Zillig for providing a number of *Sulfolobus* strains and enrichment cultures, Raffaele Cannio and Kenneth Stedman for helpful discussion. We also acknowledge support from a Young Research Leader project of the Danish Research Council of Technology and Production (FTP), from the "Archaea Centre: Diversity and Evolution" of the Danish Natural Science Research Council (FNU) to the Copenhagen laboratory, from the National Competence Centre (CRdc ATIBB, Campania Region, Naples) and the European Union (Contract QLK3-CT-2000-00649) to the Naples laboratory.

References

Ackermann HW, DuBow MS (1987) Viruses of prokaryotes: general properties of bacteriophages. CRC Press, Inc., Boca Raton

Albers SV, Jonuscheit M, Dinkelaker S, Urich T, Kletzin A, Tampe R, Driessen AJ, Schleper C (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Appl Environ Microbiol 72:102–111

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Arnold HP, She Q, Phan H, Stedman KM, Prangishvili D, Holz I, Kristjansson JK, Garrett RA, Zillig W (1999a) The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. Mol Microbiol 34:217–226

Arnold HP, Stedman KM, Zillig W (1999b) Archaeal phages. In: Webster RG, Granoff A (eds) Encyclopedia of virology. Academic, London, pp 76–89

Bettstetter M, Peng X, Garrett RA, Prangishvili D (2003) AFV1, a novel virus infecting hyperthermophilic archaea of the genus *Acidianus*. Virology 315:68–79

Blum H, Zillig W, Mallock S, Domday H, Prangishvili D (2001) The genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses. Virology 281:6–9

Campbell AM (1992) Chromosomal insertion sites for phages and plasmids. J Bacteriol 174:7495–7499

Esposito D, Scocca JJ (1997) The integrase family of tyrosine recombinases: evolution of a conserved active site domain. Nucleic Acids Res 25:3605–3614

Garrett RA, Redder P, Greve B, Brügger K, Chen L, She Q (2004) Archaeal plasmids. In: Funnell BE, Phillips GJ (eds) Plasmid Biology, ASM Press, Washington DC, pp 377–392

Grogan D, Palm P, Zillig W (1990) Isolate B12, which harbours a virus-like element, represents a new species of the archaebacterial genus *Sulfolobus*, *Sulfolobus shibatae*, sp. *nov*. Arch Microbiol 154:594–599

Jonuscheit M, Martusewitsch E, Stedman KM, Schleper C (2003) A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. Mol Microbiol 48:1241–1252

Lubelska JM, Jonuscheit M, Schleper C, Albers SV, Driessen AJ (2006) Regulation of expression of the arabinose and glucose transporter genes in the thermophilic archaeon *Sulfolobus solfataricus*. Extremophiles April 8 (Epub ahead of print)

Martin A, Yeats S, Janekovic D, Reiter WD, Aicher W, Zillig W (1984) SAV1, a temperate U. V. inducible DNA virus like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12. EMBO J 3:2165–2168

Muskhelishvili G, Palm P, Zillig W (1993) SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. Mol Gen Genet 237:334–342

Nadal M, Mirambeau G, Forterre P, Reiter WD, Duguet M (1986) Positively supercoiled DNA in a virus-like particle of an archaebacterium. Nature 321:256–258

Nunes-Duby SE, Kwon HJ, Tirumalai RS, Ellenberger T, Landy A (1998) Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucleic Acids Res 26:391–406

Palm P, Schleper C, Grampp B, Yeats S, McWilliam P, Reiter WD, Zillig W (1991) Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. Virology 185:242–250

Peng X, Blum H, She Q, Mallok S, Brügger K, Garrett RA, Zillig W, Prangishvili D (2001) Sequences and replication of genomes of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. Virology 291:226–234

Prangishvili D, Stedman K, Zillig W (2001) Viruses of the extremely thermophilic archaeon *Sulfolobus*. Trends Microbiol 9:39–43

Prangishvili D, Garrett RA, Koonin EV (2006a) Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. Virus Res 117:52–67

Prangishvili D, Vestergaard G, Haring M, Aramayo R, Basta T, Rachel R, Garrett RA (2006b) Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. J Mol Biol April 27 (pub ahead)

Reiter WD, Palm P (1990) Identification and characterization of a defective SSV1 genome integrated into a tRNA gene in the archaebacterium *Sulfolobus* sp. B12. Mol Gen Genet 221:65–71

Reiter WD, Palm P, Henschen A, Lottspeich F, Zillig W, Grampp B (1987a) Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 206:144–153

Reiter WD, Palm P, Yeats S, Zillig W (1987b) Gene expression in archaebacteria: physical mapping of constitutive and UV-inducible transcripts from the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 209:270–275

Reiter WD, Palm P, Zillig W (1988) Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. Nucleic Acids Res 16:1–19

Reiter WD, Palm P, Yeats S (1989) Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res 17:1907–1914

Sambrook J, Russell DW (2001) Molecular cloning, a laboratory manual. 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Schleper C, Kubo K, Zillig W (1992) The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. Proc Natl Acad Sci USA 89:7645–7649

Serre MC, Letzelter C, Garel JR, Duguet M (2002) Cleavage properties of an archaeal site-specific recombinase, the SSV1 integrase. J Biol Chem 277:16758–16767

She Q, Peng X, Zillig W, Garrett RA (2001a) Gene capture in archaeal chromosomes. Nature 409:478

She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC-Y, Clausen IG, Curtis BA, Moors AD, Erauso G, Fletcher C, Gordon PMK, Jong IH-D, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan M A, Sensen CW, van der Oost J (2001b) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci USA 98:7835–7840

She Q, Brugger K, Chen L (2002) Archaeal integrative elements and their impact on evolution. Res Microbiol 153:325–332

She Q, Shen B, Chen L (2004) Archaeal integrases and mechanisms of gene capture. Biochem Soc Trans 32:222–226

Snyder JC, Stedman K, Rice G, Wiedenheft B, Spuhler J, Young MJ (2003) Viruses of hyperthermophilic *Archaea*. Res Microbiol 154:474–482

Stedman KM, She Q, Phan H, Arnold HP, Holz I, Garrett RA, Zillig W (2003) Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. Res Microbiol 154:295–302

Wiedenheft B, Stedman KM, Roberto F, Willits D, Gleske AK, Zoeller L, Snyder J, Douglas T, Young M (2004) Comparative genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses. J Virol 78:1954–1961

Weinbauer MG (2004) Ecology of prokaryotic viruses. FEMS Microbiol Rev 28:127–181

Williams KP (2002) Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. Nucleic Acids Res 30:866–875

Xiang X, Chen L, Huang X, Luo Y, She Q, Huang L (2005) *Sulfolobus tengchongensis* spindle-shaped virus STSV1: virus–host interactions and genomic features. J Virol 79:8677–8686

Yeats S, McWilliam P, Zillig W (1982) A plasmid in the archaebacterium *Sulfolobus acidocaldarius*. EMBO J 1:1035–1038

Zillig W, Kletzin A, Schleper C, Holz I, Janekovi PD, Hain J, Lanzendorfer M, Kristjansson J (1994) Screening for S*ulfolobales*, their plasmids, and their viruses in Icelandic sulfataras. Syst Appl Microbiol 16:609–628

Zillig W, Arnold HP, Holz I, Prangishvili D, Schweier A, Stedman KM, She Q, Phan H, Garrett RA, Kristjansson JK (1998) Genetic elements in the extremely thermophilic archaeon *Sulfolobus*. Extremophiles 2:131–140