Impact and Regulatory Control of the CGP3 Prophage in *Corynebacterium glutamicum*

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Eugen Pfeifer
aus Pawlodar

Jülich, Mai 2017

Aus dem Institut für Bio- und Geowissenschaften 1, IBG-1 des Forschungszentrums Jülichs (assoziiert mit der Heinrich-Heine-Universität Düsseldorf).

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Juniorprof. Dr. Julia Frunzke
Institut für Bio- and Geowissenschaften, IBG-1: Biotechnologie
Populationsheterogenität und Signaltransduktion
Forschungszentrum Jülich

2. Prof. Dr. Michael Feldbrügge
Institut für Mikrobiologie
Heinrich-Heine Universität Düsseldorf

Tag der mündlichen Prüfung: 31. August 2017

"In the long run the practice of solidarity proves much more advantageous to the species than the development of individuals endowed with predatory inclinations."

Pyotr Kropotkin, Mutual Aid: A Factor of Evolution

The studies presented in this dissertation have been published in the following articles:

Helfrich S.*, **Pfeifer E**.*, Krämer C*, Sachs CC., Wiechert W., Kohlheyer D, Nöh K., and Frunzke J. (2015). Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations. *Molecular Microbiology*. doi:10.1111/mmi.13147

Pfeifer E., Hünnefeld M., Popa O., Polen T., Kohlheyer D., Baumgart M., and Frunzke J. (2016) Silencing of cryptic prophages in *Corynebacterium glutamicum*. *Nucleic Acids Research*. doi: 10.1093/nar/gkw692

Pfeifer E., Gätgens C., Polen T., and Frunzke J. (2017). Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium. ***Under Review at Scientific Reports.***

Further results on a side project not discussed in this thesis have been published in:

Donovan C*, Heyer A*, **Pfeifer E**, Polen T., Wittmann A., Krämer R., Frunzke J., and Bramkamp M. (2015). A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria. *Nucleic Acids Research.* doi: 10.1093/nar/gkv374

* Authors equally contributed to this work

Abbreviations

ALE	Adaptive laboratory evolution
AHL	*N*-Acyl-homoserine lactone
ATCC	American Type Culture Collection
CRISPR	Clustered regularly interspaced short palindromic repeats
ChAP-Seq	Chromatin affinity purification and sequencing
ChIP	Chromatin immunoprecipitation
CGP1-3	*Corynebacterium glutamic*um prophage 1-3
CgpS	*C. glutamicum* prophage silencer
DNA	Deoxyribonucleic acid
DND	DNA phosphorothioation
e.g.	*exempli gratia*
et al.	*et alii*
etc.	*et cetera*
FACS	Fluorescence-activated cell sorting
HGT	Horizontal gene transfer
H-NS	Histon-like (or heat stable) and nucleoid-structuring protein
ICE	Integrative and conjugative element
LTEE	Long-term evolution experiment
MMC	Mitomycin C
NAP	Nucleoid associated protein(s)
NGS	Next generation sequencing
ORF	Open reading frame
PK	Pyruvate kinase
PPP	Pentose phosphate pathway
PSI-BLAST	Position-Specific Iterated BLAST
RM	Restriction-Modification
RNA-Seq	RNA Sequencing
ROS	Reactive oxygen species
SNPs	Single nucleotide polymorphisms
SPI	Spontaneous prophage induction
TCA	Tricarboxylic acid cycle
TA	Toxin-Antitoxin
Topo IA	Topoisomerase type IA
UV	Ultraviolet
v/v	Volume per volume
WT	Wild-type
w/v	Weight per volume
XS	Xenogeneic silencer

Further abbreviations not included in this section are according to international standards, as, for example, listed in the author guidelines of the Journal of Cell Biology.
(http://jcb.rupress.org/content/standard-abbreviations)

Contents

1. Summary 1
1.1 English summary 1
1.2 German summary 2
2. Introduction 3
2.1 Impact of foreign DNA on bacteria 3
2.2 Defective virus-like elements may harbor beneficial traits 3
2.3 Activation of prophages, the lysogenic-lytic switch 4
2.4 SOS-independent prophage induction 6
2.5 Spontaneous prophage induction (SPI) 8
2.6 Defense systems against exogenous DNA 9
2.7 An outstanding group within nucleoid-associated proteins: xenogeneic silencer 10
2.8 Xenogeneic silencing facilitates mutual adaptation 11
2.9 State of the art: CGP3 prophage in *Corynebacterium glutamicum* 12
2.10 Aims of this work 13
3. Results 15
3.1 Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations 17
3.2 Silencing of cryptic prophages in *Corynebacterium glutamicum* 35
3.3 Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium. 55
4. Discussion 71
4.1 Prophages in bacterial genomes 71
4.2 SOS-dependent and independent SPI 72
4.3 Virus-like elements are silenced by small NAPs 75
4.4 How does the molecular mechanism of CGP3 induction work? 77
4.5 Maintenance and impact of CGP3 on *C. glutamicum* populations 80
4.6 XS-derived interactions in regulatory circuits of bacteria and virus-like elements 83
4.7 A step into the unknown – CgpS orthologues on phage genomes 85
5. References 91
6. Appendix 101
6.1 A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria 101
6.2 Results of phenotypic micro array experiment 119

6.3 Supplemental material to: Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations 133
6.4 Supplemental material to: Silencing of cryptic prophages in *Corynebacterium glutamicum* 147
6.5 Supplemental material to: Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium. 167
6.6 Supplemental material to: A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria 185

1. Summary

1.1 English summary

Almost all bacterial genomes contain substantial amounts of viral DNA, which may significantly impact microbial physiology. The spontaneous induction of prophages (SPI) has been reported as a common phenomenon of lysogenic bacterial strains and occurs even in the absence of an external trigger. By sacrificing a small fraction for the good of all, SPI was shown to promote the fitness of bacterial communities and to contribute to horizontal transfer of genetic information. Generally, it is considered that SPI is triggered by sporadic DNA damage activating the host's SOS response. However, various examples demonstrated that also alternative pathways may cause the activation of prophages. One of them includes specific counteraction of xenogeneic silencers (XS). XS such as Lsr2 of *Mycobacterium tuberculosis* and H-NS of *Escherichia coli* are small nucleoid-associated proteins, which preferentially bind and silence AT-rich foreign DNA. Furthermore, it is suggested that XS mediate the stepwise acquisition of potentially useful genes and enable a mutual adaption of the host and the foreign element.

In the present thesis, the SPI of CGP3, a cryptic prophage of the industrial relevant *Corynebacterium glutamicum* strain ATCC 13032, was investigated to examine prophage-host interactions, focusing on the impact of CGP3 on host physiology and the regulatory control of this element. In a first set of experiments, the dynamics of the SOS response and CGP3 induction were monitored by using promoter reporter fusions. Live-cell imaging of reporter strains in the microfluidic environment enabled to follow the fate of SOS and phage positive cells. This approach revealed that in ~63% of the cells an SOS response preceded prophage activation, whereas >30% of the cells displayed an SOS-independent CGP3 activation.

SPI and CGP3 inducing experiments demonstrated that activation of CGP3 causes a growth arrest in all cells (or cell death), evincing that a small fraction of a *C. glutamicum* population is continuously lost due to SPI. This fitness burden was proven by a competitive growth experiment, where the prophage-free *C. glutamicum* strain MB001 showed a slight advantage in comparison to the wild type strain. Further on, comparative analysis revealed neither by phenotypic microarrays (>1100 conditions) nor by a long-term adaptive evolution experiment (>600 generations) significant disadvantages of strain MB001 regarding growth, genome stability and mutation frequencies.

In previous studies, the prophage-encoded Lsr2-like protein CgpS was fished by a DNA affinity chromatography using an early phage promoter of CGP3. This work revealed that CgpS, in fact, inherits a crucial role as a silencer of cryptic prophage elements in *C. glutamicum*. In particular, ChAP-Seq experiments and transcriptome analysis confirmed the binding of CgpS to AT-rich DNA regions and repression of phage genes. Beside the similar mode of action and the XS-typical domain organization, CgpS can complement a Δ*hns* phenotype in *E. coli*. Remarkably, bioinformatics analysis revealed that orthologues of CgpS/Lsr2 are present in several actinobacteriophages and occur even more likely in (predicted) temperate rather than in virulent phages. This finding emphasizes that XS-like systems may play important and so far overlooked roles in the interaction of bacteria and their phages. Besides maintaining a stable co-existence, they may also be exploited as weapon by phages in the arms-race of bacteria and phages.

1.2 German summary

Bakterielle Genome sind praktisch übersät von viralen zum Teil kryptischen Elementen, welche die Physiologie von Bakterien maßgeblich beeinflussen können. Vor allem die spontane Aktivierung von viraler DNA (SPI), oder viral-ähnlichen Elementen, in lysogenen Bakterien ist ein fundamentales Phänomen, welches sogar unter nicht induzierenden Bedingungen auftritt. Durch SPI wird zum Wohle der gesamten Population ein kleiner Anteil der Population „geopfert“, was die Fitness von pathogenen oder Biofilm-bildenden Stämmen fördern kann und gleichzeitig auch zum horizontalen Transfer von genetischer Information beiträgt. Als Ursache für SPI werden spontan auftretende DNA Schäden, die zur Aktivierung der wirtsspezifischen SOS-Antwort führen, angenommen. Diverse Studien zeigten jedoch, dass auch alternative Signalwege die Prophagen aktivieren können. Einer dieser alternativen Wege basiert auf der Gegenwirkung von xenogenen *silencer* (XS) Proteinen. XS zu welchen Lsr2 aus *Mycobacterium tuberculosis* und H-NS aus *Escherichia coli* gehören, binden und reprimieren AT-reiche, fremde DNA Regionen. Interessenterweise können XS Proteine auch eine sukzessive Adaptation zwischen dem Wirt und dem extern angeeigneten Element ermöglichen.

In dieser Arbeit wurde die SPI von CGP3, einem kryptischen Prophagen des industriell relevanten Stammes *Corynebacterium glutamicum* ATCC 13032, untersucht, mit einem besonderen Fokus auf den Einfluss von CGP3 auf die Wirtsphysiologie und den molekularen Mechanismus hinter der Prophageninduktion. In ersten Experimenten wurde die Dynamik der SOS-Antwort und der Prophagenaktivierung mittels Promoterfusionen in Reporterstämmen quantitativ durch Durchflusszytometrie untersucht. Des Weiteren konnte durch Fluoreszenzmikroskopie in mikrofluidischen Systemen die weitere Entwicklung von SOS- und CGP3-induzierten Zellen verfolgt werden. Während in ~63% CGP3 durch die SOS-Antwort induziert wurde, erfolgte bemerkenswerterweise in >30% eine SOS-unabhängige Aktivierung.

Sowohl in Untersuchungen zur SPI als auch unter induzierenden Bedingungen stoppten alle Zellen, in denen CGP3 aktiviert wurde, das Wachstum und zeigten somit, dass ein geringer Anteil einer *C. glutamicum* Population durch SPI ausgelesen wird. Dieser SPI-assoziierte Nachteil konnte in einem kompetitiven Wachstumsexperiment bestätigt werden, da der Wildtyp dem Prophagen-freien Stamm MB001 im Wachstum unterlag. Auch konnten weder in einer Hochdurchsatz-Vergleichsstudie (>1100 Bedingungen wurden getestet) noch in einem Langzeit-Evolutionsexperiment (>600 Generationen) signifikante Nachteile bezüglich Wachstum, Genomstabilität und Mutationshäufigkeiten für MB001 nachgewiesen werden.

In vorangegangen Studien wurde CgpS, das CGP3-kodierte und Lsr2-ähnliche Protein, mit Hilfe eines frühen Phagenpromotors über eine DNA-Affinitätschromatographie gereinigt. In dieser Arbeit konnte die zentrale Rolle von CgpS als ein XS von kryptischen Prophagenelementen gezeigt werden. Darüber hinaus konnte die bevorzugte Bindung an AT-reiche Regionen und die Repression von Phagengenen vor allem durch DNA-Bindestudien (ChAP-Seq, EMSA) und Transkriptomanalysen gezeigt werden. Neben einer klassischen XS-Domänenorganisation zeigte CgpS auch die Fähigkeit einen Δ*hns* Phänotypen in *E. coli* zu komplementieren. Mittels bioinformatischer Analysen konnten zahlreiche Aktinobakteriophagen mit Orthologen von CgpS/Lsr2 identifiziert werden, die interessanterweise häufiger in vorhergesagten temperenten als in virulenten Phagen vorkommen. Diese Studie hebt besonders hervor, dass XS Systeme wichtige und bislang übersehene Funktionen in der Interaktion zwischen Bakterien und Phagen einnehmen. Neben der Erhaltung einer stabilen Koexistenz können sie somit auch als virale Waffe im Wettrüsten zwischen Bakterien und Phagen genutzt werden.

2. Introduction

2.1 Impact of foreign DNA on bacteria

Free-living bacteria are frequently confronted with fluctuating and challenging environments placing microbial populations under ever-changing selection pressures. Harsh conditions set them into ‘evolve-or-die’ situations, forcing single individuals to rapidly adapt to the new circumstances. Acquisition of foreign genetic material by horizontal gene transfer (HGT) facilitates microorganism a fast adaptation towards the changing environments. Furthermore, HGT is a major driver of microbial evolution enabling the transfer of large pieces of DNA between cells. Evidence for HGT can be found in bacterial genomes, that are practically littered with genetic regions acquired horizontally by phages, transposons or conjugative plasmids (Canchaya *et al.*, 2003, Juhas *et al.*, 2009, Ochman *et al.*, 2000). Remarkably, in some bacterial strains, prophages and virus-like elements may constitute up to 20% of the entire genome (Canchaya *et al.*, 2003, Casjens, 2003). Several recent studies proved a considerable contribution to the host fitness by enhancing stress tolerance, antibiotic resistance, biofilm formation or virulence (Nanda *et al.*, 2015, Wang *et al.*, 2010). Especially, various disease-causing strains such as the Shiga toxin-producing *Escherichia coli*, *Corynebacterium diphtheriae* or *Salmonella enterica* sv. Typhimurium became pathogen after viral infections (Brüssow *et al.*, 2004, Fortier & Sekulovic, 2013). In fact, virulence genes are often clustered in genomic islands, which are transmitted by HGT (Hacker & Kaper, 2000). However, HGT also harbors high risks due to the fact that most of the acquired genes will interfere with the host metabolism and regulatory networks. This, for example, is the case for selfishly replicating mobile elements and genes encoding proteins with no useful or even detrimental functions. In the worst case, infection by viral DNA often leads directly to cell lysis (Figure 1).

2.2 Defective virus-like elements may harbor beneficial traits

A recent study revealed that nearly half of the sequenced bacteria are lysogens by identifying over 2246 *bona fide* prophages in 2110 bacterial genomes (Touchon *et al.*, 2016). These numbers apparently reflect the ubiquitous presence of integrated viral elements and the close interconnections to their hosts. Furthermore, if defective prophages would be included, these numbers would definitely be higher. Cryptic prophages represent the result of genetic degradations of intact temperate prophages. The grade of their genetic decay can be classified

by the different functions affected, including prophage excision, virion formation, lysis, or infectivity (Figure 1). In a bioinformatics study the evolutionary dynamics of intact and defective prophages were examined (>600 prophages), supporting a rapid prophage inactivation followed by a slower degeneration (based on a bimodal size distribution) (Bobay *et al.*, 2014). Cryptic prophages or phage remnants are often considered as relicts or 'junk' DNA. However, several studies convinced the opposite by demonstrating that cryptic prophages deliver beneficial functions to their hosts, e.g., in disseminating virulence genes (Asadulghani *et al.*, 2009), at different stress conditions (Wang *et al.*, 2010), or to serve as active regulatory switches (Feiner *et al.*, 2015). One prominent example was given by Wang et al., who showed that the deletion of all nine cryptic prophages in the *E. coli* strain K12 BW25113 leads to less fitness and stress tolerance (antibiotic and osmotic stress) (Wang *et al.*, 2010). In addition, one of these defective prophages, CP4-57, was identified to be involved in the regulation of motility genes, which activation leads to an increase of their expression and therefore to larger biofilm communities (Wang *et al.*, 2009). However, activations of prophage-like elements not only trigger the expression of accessory genes, but, are often associated with cell death or lysis. This is also in line with the bioinformatics study of Bobay et al., who compared several hundred prophages and concluded that the bacterial host cells select for typical phage genes (cell lysis, structural components) in a similar rate as for accessory genes (Bobay *et al.*, 2014). In addition, the controversy advantages of cell lysis are nicely illustrated by the striking phenomenon dubbed 'spontaneous prophage induction' (SPI). This feature, *inter alia*, allows bacterial populations to benefit from the loss of a small fraction by enhancing the overall fitness (Nanda *et al.*, 2015). Thus, these studies emphasize the great potentials of the viral gene pool and underline the mutual benefits of incorporation of foreign elements, such as prophages, into host regulatory circuits.

2.3 Activation of prophages, the lysogenic-lytic switch

'Do not attract attention and keep discretely your silent state' are the key strategies that prophages follow during the lysogeny. On the basis of these simple assignments, viral DNA was efficiently spread vertically across microbial species, as indicated by several bioinformatics studies (Canchaya *et al.*, 2003, Hatfull, 2014, Touchon *et al.*, 2016). However, if things get really bad, it is a decisive advantage to follow a back-up plan. In the prophage life cycle this concept is described as the lysogenic-lytic switch, also termed prophage induction. It enables the reactivation of the viral element by a controlled switch to the lytic cycle, virion production and, finally, the escape by lysing its host cell (Figure 1).

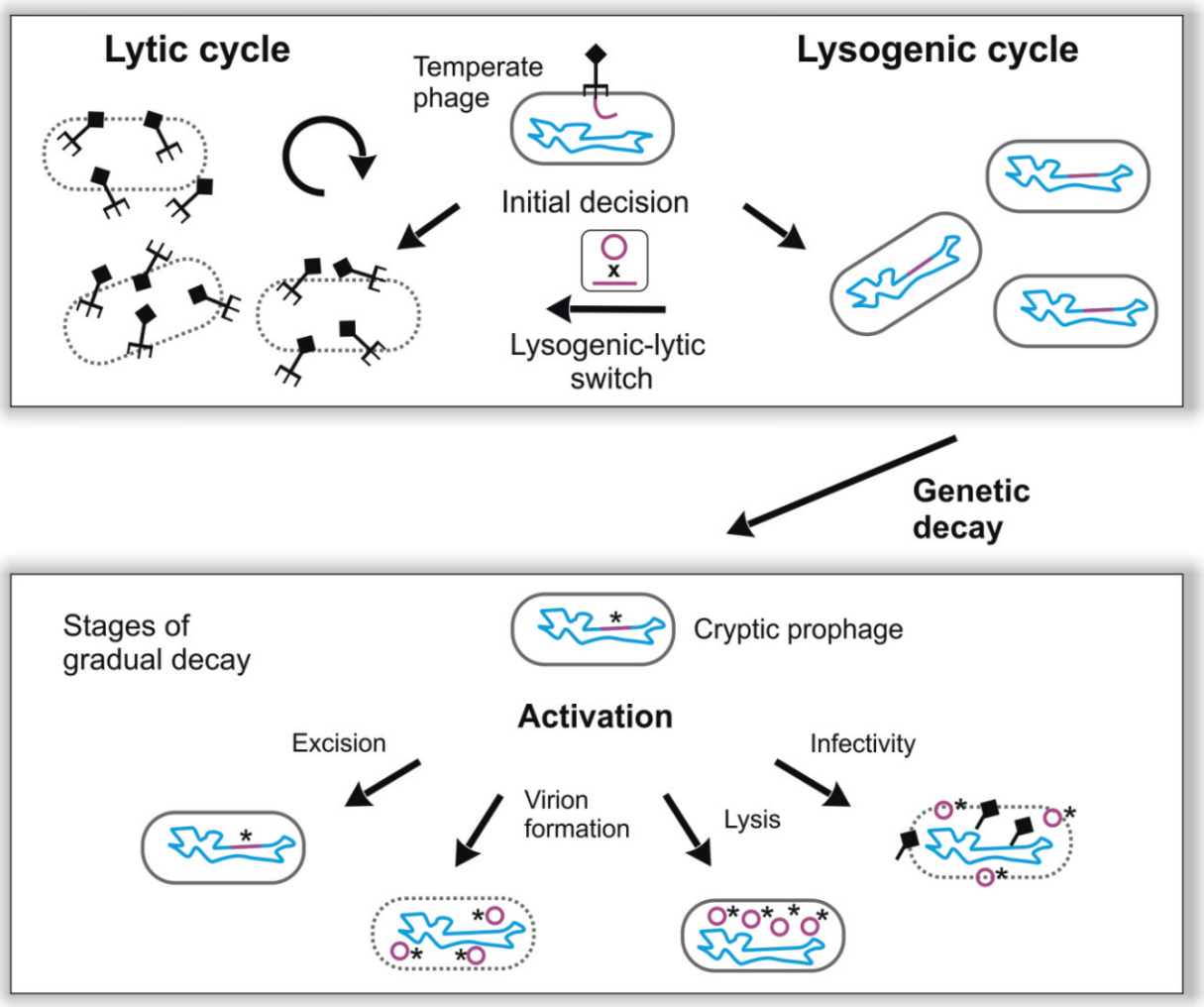


Figure 1. Emergence of degenerated (cryptic) prophages. If temperate phages take the initial decision to enter the lysogenic cycle, their DNA will be (in the most cases) integrated into the bacterial chromosome (prophage). The lysogenic-lytic switch allows prophages to re-enter the lytic cycle again. The most prominent trigger of prophage induction is represented by the bacterial SOS response (Janion, 2008, Canchaya *et al.*, 2003, Casjens, 2003). Genomic rearrangements and mutations may lead to the gradual genetic decay of the prophages, which then may lose specific functions of the lytic cycle. Different degrees of genetic degradation can generate various types of defective prophages that are degenerated in, for example, excision, virion formation, cell lysis, or in their infectivity.

In the classic *E. coli*-λ model, the SOS response was elucidated as the key trigger of the prophage induction. The SOS response is described as a DNA repair system, encoded by SOS genes involved in protection, repair, replication, metabolism, and mutagenesis of DNA, providing cells to maintain their correct DNA sequence (Janion, 2008). Key players in the regulatory control of prophage induction are the proteins LexA and RecA. Upon DNA damage, the co-protease RecA forms nucleoprotein filaments on ssDNA, thereby triggering the autocatalytic cleavage of the SOS gene repressor LexA. In addition, the ssDNA-RecA filaments also trigger self-digestion of the phage repressor CI (due to conserved cleavage residues (Mustard & Little, 2000)). Thus, RecA inhabits a crucial role in the de-repression of SOS genes as well as in

prophage induction (Casjens & Hendrix, 2015, Janion, 2008). Beside this 'classic' activation, cases were also described, where LexA is either directly used by proviral DNA to maintain the lysogeny (Fornelos *et al.*, 2011) (Figure 2), or where LexA represses a prophage encoded anti-repression system, which activation triggers prophage induction (Kim & Ryu, 2013) (Figure 2). However, it is rather remarkable that, although the *E. coli*-λ system was studied over the last 50 years, new findings are still achieved. This, for example, was nicely illustrated in a recent publication, describing an additional way how λ modulates the lysogenic-lytic switch. The authors characterized the λ-encoded protein λRexB, which main function is to prevent the degradation of the antitoxin MazE. MazE facilitates *mazEF*, a TA system, to stay in its nontoxic state (Engelberg-Kulka & Kumar, 2015, Erental *et al.*, 2014). Furthermore, the *mazEF* pathway is reported to inhibit the SOS response and, thus, by protecting MazE, λRexB allows the SOS response to correctly fulfill a proper λ induction and also provides sufficient time for virion formation (Engelberg-Kulka & Kumar, 2015). In sum, these examples illustrate the complex but finely adjusted regulatory interactions of prophages and their particular host systems.

2.4 SOS-independent prophage induction

The addition of the DNA damaging agent mitomycin C (MMC) represents a general approach for the experimental identification of prophage elements in a particular strain. MMC induces the SOS response and, hence, enables the verification whether predicted prophages are able to enter and fulfill the lytic cycle. Several studies confirmed the success of this approach by isolating and characterizing phages from various bacterial species (Canchaya *et al.*, 2002, Lorenz *et al.*, 2016, Sekulovic & Fortier, 2016, L. Brown *et al.*, 2017). However, a serious disadvantage of MMC is that only the SOS-dependent induction is triggered, but, differently regulated prophages are not affected. So far, there are only a few cases that describe the activation of viral elements in an SOS-independent manner. In a recent study, the spontaneous induction rate of closely related phages were found to be affected by RecA-dependent and independent effects (Colon *et al.*, 2016). In particular, they proved that Shiga toxin (Stx2) encoding prophages form less stable lysogens, than *stx2* negative phages even in Δ*recA* mutants, suggesting that the less availability of CI repressor proteins is the reason for the higher inducibility (Colon *et al.*, 2016). Further studies revealed that cell density depending molecules can also be involved in prophage induction (Ghosh *et al.*, 2009, Rossmann *et al.*, 2015). Moreover, the SOS-independent prophage activation was nicely demonstrated in co-cultivation experiments, in which quorum sensing signaling molecules, produced by a *Pseudomonas aeruginosa* strain, triggered λ induction in a Δ*recA E. coli* lysogen (Ghosh *et al.*, 2009) (Figure 2).

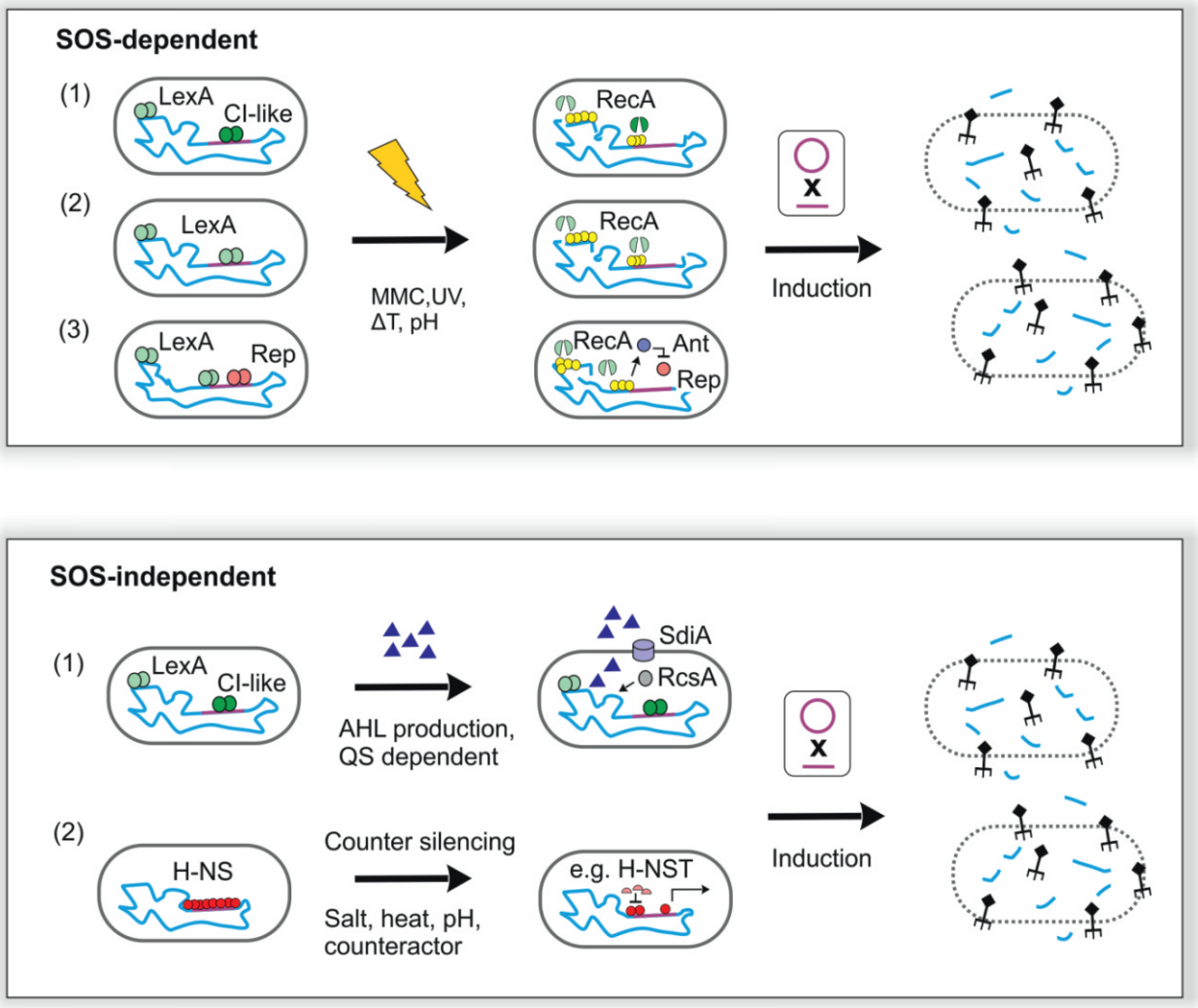


Figure 2: SOS dependent and independent induction of prophages. SOS dependent: According to the classic model prophage induction is triggered at once by the SOS response due to the autocatalytic cleavages of the SOS (LexA) and prophage repressor (CI) proteins caused by the co-protease RecA. In special cases, temperate phages use directly LexA to preserve their lysogeny (Fornelos *et al.*, 2011) or their induction is based on an anti-repression system, where a LexA-negatively regulated antirepressor (here Ant) facilitates induction by inactivating a phage repressor (here Rep) (Kim & Ryu, 2013). **SOS-independent:** Based on the quorum sensing principle, the signaling molecules acyl-homoserine lactones are produced and trigger RecA-independently the induction of λ. The AHL receptor SdiA and a transcriptional regulator RcsA are described to be linked to the induction process (Ghosh *et al.*, 2009). The XS protein H-NS was also identified to be involved in the lysogeny of λ (Rozanov *et al.*, 1998) or as a direct repressor of the cryptic prophage Rac (Hong *et al.*, 2010). It was shown that counter action of H-NS activity caused induction of the respective prophages (Hong *et al.*, 2010, Rozanov *et al.*, 1998).

In a screening approach with *E. coli* Δ*recA* λ lysogens, two host-encoded genes were found (prior known to be involved in capsular polysaccharide synthesis) to affect the induction of λ (Rozanov *et al.*, 1998). Interestingly, one of these genes is usually negatively regulated by the global silencer H-NS (described in detail in section 2.7), which was also identified to repress the cryptic prophage Rac in *E. coli* (Hong *et al.*, 2010). Thus, the H-NS silencing system represents a further SOS-independent strategy to modulate the activation of viral-like elements. A further

recent study reported on the influence of the transcription termination factor Rho on the maintenance of prophage-like elements. Menouni et al. demonstrated that inactivation of Rho by the selective antibiotic bicyclomycin triggers the expression of prophage genes (even in a Δ*recA* strain) (Menouni *et al.*, 2013). Furthermore, it was shown that Rho is also involved in silencing foreign genetic elements, especially to protect the host cell against the toxic effects (Cardinale *et al.*, 2008). These examples highlight that nature goes far beyond the classic *E. coli* λ system and the diversity of mechanisms behind prophage induction.

2.5 Spontaneous prophage induction (SPI)

As firstly described by the famous virologist Lwoff in the early 1950s, activation of prophages can occur spontaneously in a population of lysogenic cells (Lwoff, 1953). In several independent studies, SPI was shown to have multifaceted physiological impacts on bacterial communities (Nanda *et al.*, 2015). Due to the phage's predatory nature, the impetuous activation of viral elements in single individuals is often associated with cell lysis of the particular cell. Interestingly, several studies highlighted that SPI may increase the overall fitness of the population by improving biofilm formation (Carrolo *et al.*, 2010) or facilitating the toxin releases of pathogenic species (Livny & Friedman, 2004) (Figure 3). A rather cell-'friendly' consequence of spontaneous induction is more frequently distributed among integrative and conjugative elements (ICE), that spread over conjugational bridges and usually do not cause cell lysis (Johnson & Grossman, 2015). ICEs harbor high similarities in transmission and regulatory mechanisms to those of plasmids, transposons, and viral elements and, therefore, in an ICE-bearing population, spontaneous induction may enable the action of donating cells at any given minute (Johnson & Grossman, 2015) (Figure 3). Sporadic occuring DNA damage was described as a predominant trigger of SPI and the activation of other mobile elements (Nanda *et al.*, 2015, Johnson & Grossman, 2015, Nanda *et al.*, 2014, Pennington & Rosenberg, 2007). Stalled replication forks, mutations, especially in the repair system, or external triggers (such as reactive oxygen species (ROS), UV light or DNA damaging-agents) can cause severe DNA lesions, inducing the SOS response (Kuzminov, 1999, Pennington & Rosenberg, 2007) (Figure 2). However, only little is known about the SOS-independent induction of viral like elements (see section 2.5). For certain, given by the tremendous different types of phages and the great progresses in *omic* technologies, such as next-generation sequencing (NGS), future studies will definitely reveal more examples of SOS dependent and -independent SPI, allowing to study their multifaceted impacts on the bacterial physiology.

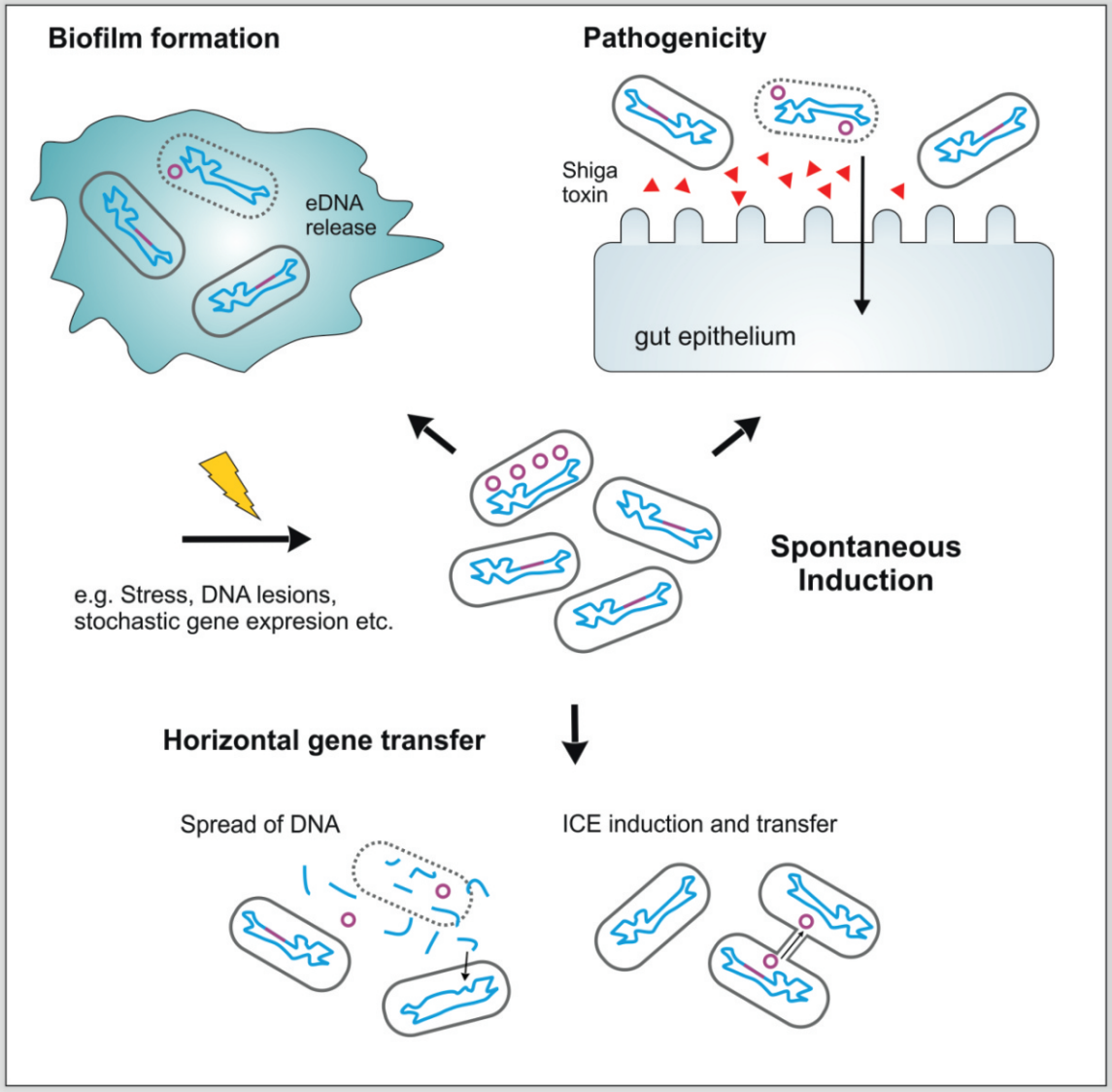


Figure 3: Impact of spontaneous prophage induction (SPI) on bacterial populations. Several cases emphasized beneficial features of SPI. Biofilm formation can be enhanced due to the release of eDNA caused by cell lysis. The propagation and virulence of pathogenic strains are facilitated since prophage encoded toxins, such as Shiga toxin, were released by the lysis of a small fraction of cells. Furthermore, SPI contributes to lateral gene transfer by spreading DNA into the environment and, as a special case, ICE use the spontaneous activation to dispense their genetic material over conjugational bridges. Figure taken and adapted from (Nanda *et al.*, 2015).

2.6 Defense systems against exogenous DNA

It is very likely that acquisition of foreign DNA, especially of temperate phages (or in general by HGT), will cause in the most cases detrimental damages to the host organism due to toxic effects and/ or incompatibility to regulatory circuits. For their own protection, bacterial cells

possess several defense strategies, which can be classified in different groups, depending on their mode of actions. The first group takes the advantage of the self- and non-self-discrimination principle. The CRISPR-Cas pathway (Clustered Regularly Interspaced Palindromic Repeats-CRISPR-associated protein) belongs to this group and is referred as the 'bacterial immunity'. It allows serial acquisition of small pieces of viral and plasmid DNA, termed spacers, which are utilized to target and cleave viral and plasmid DNA in subsequent genomic invasions. Hence, this system provides critical immunological memory (Weinberger *et al.*, 2012). Further members of this group are the Restriction-Modification (RM) and the DNA phosphorothioation (DND) systems, which label self-genomic DNA, recognize and cleave unmodified foreign DNA (Makarova *et al.*, 2013). The second group consists of numerous Toxin-Antitoxin (TA) systems. Their various mechanisms are based on the poison and the corresponding antidote principle. Strategies like abortive infection or phage exclusion, which also apply the mechanisms of cell death or dormancy, also belong to this class (Makarova *et al.*, 2013). The third group is represented by small nucleoid-associated proteins (NAP) that are specified as xenogeneic silencers (XS). In contrast to the drastic actions of the first two groups, XS operate in a more 'smooth' way since they do not cleave the invading DNA element. One of their main functions is to target and silence foreign AT-rich DNA by preventing its transcription (Navarre *et al.*, 2007).

2.7 An outstanding group within nucleoid-associated proteins: xenogeneic silencer

Bacterial proteins with the ability to alter the shape of DNA, making it more compact and influencing transcription are referred to as NAPs. Through bending, wrapping, or bridging these proteins are able to change the trajectory of DNA molecules, thereby affecting its transcription (Browning *et al.*, 2010, Dillon & Dorman, 2010). Interestingly, one subclass of NAPs was reported to recognize exogenous, AT-rich DNA, specifically bind and silence it. Members of this group are termed as xenogeneic silencers (XS). The histon-like and nucleoid-structuring protein H-NS, which can be found in nearly all Gram-negative bacteria (Dillon & Dorman, 2010), is the best-characterized XS. Further members are represented by Lsr2-like proteins of actinomycetes (Gordon *et al.*, 2008), MvaT of *Pseudomonas* (Ding *et al.*, 2015) and Rok-like proteins of *Bacillus* species (Smits & Grossman, 2010). Common characteristics of XS are their typically small size, high intracellular protein levels, similar domain organizations, and a preference to bind AT-rich DNA. Since foreign genetic material tends to be AT-richer than their new host genome (Daubin *et al.*, 2003, Rocha & Danchin, 2002), these regions are therefore targeted by XS. The domain organization is separated into a C-terminal DNA binding part and an N-terminal domain that is responsible for the oligomerization. For H-NS, it is suggested that after an initial

binding of an H-NS monomer to an AT-rich DNA region, the oligomerization of additional H-NS molecules is triggered at this site, polymeric chains are formed and cause the structural change (Singh *et al.*, 2016). Due to the constrained DNA strands, transcription of this region is strongly inhibited and therefore silenced (Navarre *et al.*, 2007).

2.8 Xenogeneic silencing facilitates mutual adaptation

In contrast to other defense systems, XS allow the host to integrate and access potentially beneficial traits that are encoded by the newly acquired DNA pieces (Figure 4). However, genes that might be useful under defined circumstances, nevertheless, require a strict regulation since uncontrolled (over-) expression, especially of phage genes, might be detrimental. This hypothesis of mutual adaptation was suggested as the first XS were characterized (Navarre *et al.*, 2007) and recent studies provide first evidence supporting this assumption.

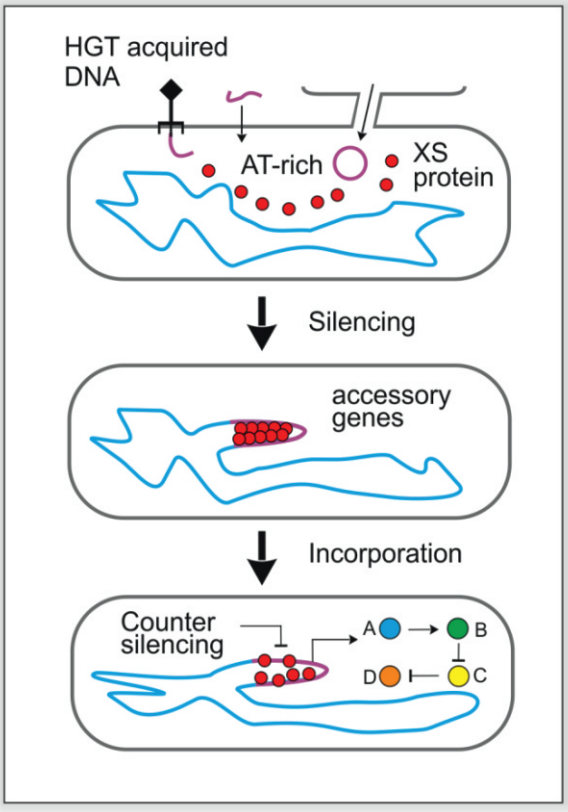


Figure 4: Insertion of alien DNA into the regulatory network of the host organism. Foreign genetic material may enter the cell via transduction, transformation, or conjugation. If these genetic elements are AT-rich and not destroyed by other defense mechanism, XS proteins will bind and silence their gene expression. During the silent state, mutual adaptation can take place by stepwise integrating foreign, beneficial proteins into the regulatory network of the host cell.

Firstly, comparative ChAP-seq analysis suggested that H-NS bound regions are more likely to mutate, allowing a tailored integration into the host regulation network (Higashi *et al.*, 2016). Secondly, time course transcriptome analysis revealed an uncoupling of mRNA level from its expression time for HGT-acquired and H-NS associated genes (Zwir *et al.*, 2014). This altered regulation mechanism illustrates how the expression level of foreign DNA can be further adjusted by XS. Lastly - and this point may well be more important than the previous two - evolution experiments conducted with *Salmonella* Typhimurium strains emphasized the indispensability of H-NS, since deregulated expression of its targets caused significant fitness losses (Huttener *et al.*, 2015).

2.9 State of the art: CGP3 prophage in *Corynebacterium glutamicum*

In this work, *Corynebacterium glutamicum* was used to investigate the impact of viral elements on bacterial physiology. *C. glutamicum* is a Gram positive, non-pathogenic soil bacterium, firstly isolated due its natural high production of glutamate (Kinoshita *et al.*, 2004). Beside its high importance for industrial processes and biotechnology, it is a well-established model organism of the Corynebacteriales, also including the closely related pathogenic strains *C. diphtheriae* and *Mycobacterium tuberculosis*. In the *C. glutamicum* strain ATCC 13032, three cryptic prophages (CGP1 ~13.5 kb, CGP2 ~3.9 kb, CGP3 ~187 kb) were predicted and experimentally verified (Frunzke *et al.*, 2008, Kalinowski *et al.*, 2003). As common for prophage elements, most predicted open reading frames (ORFs) are annotated as hypothetical or putative genes. However, several ORFs encode for typical viral proteins like phage lysins, integrases, capsids, a resolvase, or a phage primase and, therefore, these genetic regions are clearly classified as *bona fide* prophages (Helfrich *et al.*, 2015, Kalinowski *et al.*, 2003). Furthermore, in another genome sequence of the ATCC 13032 strain, published by Ikeda and Nakagawa, a fourth prophage (CGP4) was identified, which is located in the ~190 kb large CGP3 region (Ikeda & Nakagawa, 2003). For this reason, it is noteworthy that in this work all experiments were conducted with the strain harboring three prophages, which was sequenced in Bielefeld in 2003 (Kalinowski *et al.*, 2003). Furthermore, neither active phage particles nor CGP3-linked cell lysis were experimentally verified, so far (Pfeifer, 2013).

Nonetheless, studies under non-inducing conditions or conducted with MMC demonstrated that the large CGP3 prophage can excise from the genome and forms circular DNA molecules (Frunzke *et al.*, 2008, Donovan *et al.*, 2015). Moreover, it was shown that spontaneous CGP3 induction occurs in less than 1% of a *C. glutamicum* population (Frunzke *et al.*, 2008), which leaves open the question of the purpose behind the SPI.

In a genome-reduction approach, the deletions of all three cryptic prophages were achieved that resulted in the prophage-free strain MB001 (Baumgart *et al.*, 2013). Interestingly, deletion of CGP3 was only possible if prior the CGP3-encoded RM system was disrupted (Baumgart *et al.*, 2013). First studies revealed no significant differences between the wild type ATCC 13032 strain and MB001 under non-inducing conditions (Baumgart *et al.*, 2013). Based on these findings, several questions arise, especially regarding the impact of CGP3.

From first single-cell experiments, the SPI of CGP3 was positively correlated to spontaneous induction of the SOS response. The SOS-dependent CGP3 activation is in line with the classic model supported by a predicted CI-like repressor (cg2040), which is encoded in the CGP3 region. However, a deletion of this repressor gene is possible without causing impacts on CGP3 and the overexpression resulted in a minor increase in expression of adjacent genes (Heyer, 2013). Since these studies did not clarify the molecular mechanism behind CGP3 induction, another strategy was applied to identify potential regulatory proteins. By a previous time course transcriptome analysis under MMC inducing conditions (Donovan *et al.*, 2015) early phage promoters were identified. A DNA affinity chromatography was performed using the early promoter of *alpAC* (cg1890-cg1891) and resulted in the finding of a CGP3-encoded small protein, we named CgpS (cg1966) (Pfeifer, 2013). BLAST analysis revealed that CgpS is a homolog of Lsr2 from *M. tuberculosis,* which is assigned to the group of XS proteins (see section 2.7). Hence, due their homology, CgpS represents a potential repressor protein of CGP3.

2.10 Aims of this work

The major goals of this doctoral thesis will be addressed by the three following subprojects: (i) quantitative evaluation of the CGP3 SPI on the single-cell level applying live-cell imaging of reporter strains; (ii) elucidation of the regulatory role of the NAP CgpS in the control of CGP3 activity; and (iii) examination of the impact of the CGP3 island on the physiology and evolvability of *C. glutamicum* ATCC13032.

The population heterogeneity caused by the spontaneous induction of CGP3 is not fully understood yet and numerous questions arose in previous studies concerning the physiological relevance and the molecular triggers of SPI. The observation of SPI has been reported in a number of studies, but no quantitative analysis describing the dynamics of cellular stress responses and prophage induction has been performed so far. In this thesis, single-cell approaches, in particular, live-cell imaging and fluorescence-activated cell sorting (FACS) will be applied to investigate the single-cell dynamics of the cellular SOS response and its quantitative correlation to the induction of the CGP3 prophage. This study will also be important to answer

the question whether there is an SOS-independent SPI and for the identification of further potential triggers of SPI.

In recent studies, the small NAP CgpS was identified by a DNA-affinity chromatography using an early phage promoter. A major goal of this thesis will be the description of the role of CgpS in the control of CGP3 gene expression. Genome-wide binding profiling and transcriptome analysis will be performed to identify the direct target genes of CgpS. By modulating the CgpS activity level in a prophage reporter strain, its influence on the induction rate of CGP3 will be investigated.

In a third subproject, a comprehensive analysis of the impact of this large CGP3 island on *C. glutamicum* physiology will be conducted. Phenotypic microarrays and a long-term evolutionary approach will provide insights into physiological effects. Hence they will be applied in a comparative study using the wild type strain and its prophage-free variant MB001 to analyze the robustness of proviral-free *C. glutamicum* strains for metabolic engineering.

3. Results

The main purpose of the dissertation was to elucidate how spontaneous induction of CGP3 influences *C. glutamicum* populations, what types of trigger are behind the induction, and what is the impact of CGP3 on *C. glutamicum* physiology, especially on a long-term scale. Results obtained while addressing these questions were published in three articles (as first author). One other publications where I contributed as co-author in a side project dealing with the characterization of a CGP3-encoded actin-like protein (Donovan *et al.*, 2015) will not be discussed in this thesis, but are attached in the appendix section.

In the first publication “Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations” the SPI of CGP3 was quantitatively investigated using live-cell imaging of reporter strains. Fluorescence-based reporter systems, visualizing SOS and prophage activity, were introduced into *C. glutamicum* strains that were examined under non-inducing conditions by flow cytometry and in the microfluidic environment. These two complementary single-cell approaches revealed a positive correlation of the SOS response to CGP3 induction, but also uncovered a significant fraction of SOS-independent CGP3 induced cells. In the absence of *recA* SPI of CGP3 was still detectable, highlighting the fact that alternative triggers feed into the induction of CGP3. Furthermore, activation of CGP3 led in all cases to cell death or senescence, indicating to be used as a strict selection system provided by this prophage.

In a second project, the regulatory role of the NAP CgpS on CGP3 activation was elucidated and the results are summarized in “Silencing of cryptic prophages in *Corynebacterium glutamicum*”. Here, we could show that the phage-encoded homolog of Lsr2, named CgpS, is a key player in the maintenance of CGP3. Genome-wide binding profiles revealed CgpS to bind preferentially AT-rich, foreign DNA, especially its main target CGP3, thereby functioning as a typical XS protein. Furthermore, we demonstrated that CgpS is able to complement a Δ*hns* phenotype in *E. coli,* underlining an analog function of these two XS proteins, which belong to different XS classes. A bioinformatics analysis revealed several hundred orthologues mainly present in Actinobacteria but, remarkably, also in viral genomes and predicted prophage elements. Whereas multiple sequence alignments evinced a low identity match (~20%), secondary structure predictions displayed a strong conservation, emphasizing similar functions of these proteins. Based on these data, it can be assumed that viral-encoded XS proteins play a central role in (pro-)phage-bacteria interaction, which has been underestimated so far.

The studies concerning the impact of CGP3 on the physiology of *C. glutamicum* are divided into two approaches: (i) application of a high throughput comparison method (phenotypic microarrays) with the wild type strain and its prophage-free variant MB001 and (ii) investigation

of CGP3 stability and its influence on *C. glutamicum* cells in a long-term adaptive evolution experiment. Results of the first part are presented in the appendix part (see 6.2). In these experiments respiration rates (correlated to growth curves) of the wild type strain and MB001 were compared using the Phenotypic Microarray™ technology of Biolog. A comparison of the two strains, under more than 1100 different conditions, did not reveal any reproducible and significant differences between the two strains. Achievements regarding the second (evolutionary) approach are summarized in the manuscript “Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium”, which is currently under review in *Metabolic engineering*. In this study, a comparative adaptive evolution approach was conducted to examine the impact of CGP3 on the growth, genome stability and evolvability of *C. glutamicum* on glucose minimal medium.

3.1 Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations

Helfrich S.•, Pfeifer E.•, Krämer C.•, Sachs CC., Wiechert W., Kohlheyer D, Nöh K., and Frunzke J.*.

IBG-1: Biotechnology, Forschungszentrum Jülich, Jülich, Germany

•These authors contributed equally to this work.

*Corresponding author

Name of Journal: *Molecular Microbiology*

Impact factor: *3.761*

Author contributions

Own contribution to the work: 35%

Contribution to project design and manuscript

	Name	Contribution
Project planning	Helfrich, S.	15%
	Pfeifer, E.	15%
	Krämer, C.	15%
	Sachs, CC.	5%
	Wiechert, W.	5%
	Kohlheyer, D.	5%
	Nöh, K.	10%
	Frunzke, J.	30%
Writing	Helfrich, S.	20%
	Pfeifer, E.	20%
	Krämer, C.	5%
	Sachs, CC.	5%
	Nöh, K.	5%
	Wiechert, W.	5%
	Frunzke, J.	40%

Contribution to experiments, their evaluation and illustration

	Name	Experimental work	Evaluation	Processing & Preparation of figure/ table
Figure 1	**Pfeifer, E**.	100%	70%	100%
	Frunzke, J.	-	30%	-
Figure 2	Helfrich, S.	-	60%	60%
	Pfeifer, E.	-	15%	40%
	Frunzke, J.	-	25%	-
	Krämer, C.	100%	-	-
Figure 3	Helfrich, S.	-	40%	50%
	Pfeifer, E.	-	20%	50%
	Frunzke, J.	-	30%	-
	Krämer, C.	100%	-	-
	Nöh, K.	-	20%	-
Table 1	Helfrich, S.	based on Fig. 3	100%	100%

	Name	Experimental work	Evaluation	Processing & Preparation of figure/ table
Figure 4	Helfrich, S. Frunzke, J	based on Fig. 3 -	50% 50%	100% -
Figure 5	Helfrich, S. Frunzke, J	based on Fig. 3 -	60% 40%	100% -
Figure 6	**Pfeifer, E**. Krämer, C. Frunzke, J.	- 100% -	50% - 50%	100% - -
Figure 7	Sachs, CC. Frunzke, J.	100% -	40% 60%	100% -
Table 2	**Pfeifer, E.**	100%	-	100%
Table S1	**Pfeifer, E.**	-	-	100%
Table S2	**Pfeifer, E.**	-	-	100%
Figure S1	**Pfeifer, E.** Frunzke, J.	100% -	50% 50%	100% -
Figure S2	Helfrich, S. Krämer, C. Frunzke, J.	- 100%	25% 25% 50%	100% - -
Figure S3	Helfrich, S.	based on Fig. 3	100%	100%
Figure S4	Helfrich, S.	-	100%	100%
Figure S5	Helfrich, S.	based on Fig. 3	100%	100%
Figure S6	Helfrich, S.	based on Fig. 3	100%	100%
VideoS1-S3	Krämer, C.	100%	-	100%
VideoS4	**Pfeifer, E.**	100%	100%	100%
VideoS5	Krämer, C.	100%	-	100%
VideoS6	Sachs, CC.	100%	-	100%
VideoS7	Helfrich, S.	-	100%	100%

Molecular Microbiology (2015) ■ doi:10.1111/mmi.13147

Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations

Stefan Helfrich,† Eugen Pfeifer,† Christina Kr mer,† Christian Carsten Sachs, Wolfgang Wiechert, Dietrich Kohlheyer,* Katharina N h* and Julia Frunzke*

Institut f r Bio- und Geowissenschaften, IBG-1: Biotechnologie, Forschungszentrum J lich, 52425, J lich, Germany

Summary

Almost all bacterial genomes contain DNA of viral origin, including functional prophages or degenerated phage elements. A frequent but often unnoted phenomenon is the spontaneous induction of prophage elements (SPI) even in the absence of an external stimulus. In this study, we have analyzed SPI of the large, degenerated prophage CGP3 (187 kbp), which is integrated into the genome of the Gram-positive *Corynebacterium glutamicum* ATCC 13032. Time-lapse fluorescence microscopy of fluorescent reporter strains grown in microfluidic chips revealed the sporadic induction of the SOS response as a prominent trigger of CGP3 SPI but also displayed a considerable fraction (~30%) of RecA-independent SPI. Whereas approx. 20% of SOS-induced cells recovered from this stress and resumed growth, the spontaneous induction of CGP3 always led to a stop of growth and likely cell death. A carbon source starvation experiment clearly emphasized that SPI only occurs in actively proliferating cells, whereas sporadic SOS induction was still observed in resting cells. These data highlight the impact of sporadic DNA damage on the activity of prophage elements and provide a time-resolved, quantitative description of SPI as general phenomenon of bacterial populations.

Introduction

DNA of viral origin is a prevalent element of bacterial genomes and can account for up to 20% of the whole genome (Casjens, 2003; Canchaya *et al.*, 2004). Genomically integrated bacteriophage DNA not only comprises fully functional prophages that are able to undergo a lytic life cycle, but also cryptic prophages or single phage genes that were trapped in the genome due to genomic rearrangements and gradual decay. Due to the integration into the genetic circuitry of the bacterial host, these elements may have a significant impact on host fitness by equipping their host with genes for virulence factors or toxins (Barondess and Beckwith, 1990; Waldor and Mekalanos, 1996; Neely and Friedman, 1998), stress resistance (Wang *et al.*, 2010b) or metabolic traits (Edlin *et al.*, 1975).

Typically, the lysogenic, dormant state of temperate phages is very stable and maintained by action of a central phage repressor protein (Oppenheim *et al.*, 2005). However, early in the last century, free phage particles were found in cultures of lysogenic bacteria in the absence of an external trigger, leading to the term 'spontaneous prophage induction' (SPI) (Lwoff, 1953). SPI was long considered as a potentially detrimental process, as a certain fraction of cells is continuously lost by the activation of lysogenic phages. Remarkably, several recent studies revealed the beneficial impact of SPI on the fitness of bacterial populations [for a recent review see (Nanda *et al.*, 2015)] as it, for instance, contributes to the release of extracellular DNA (eDNA), which represents an important component of several microbial biofilms (Rice *et al.*, 2008; Carrolo *et al.*, 2010), to the release of toxins or adhesion factors (Livny and Friedman, 2004; Mitchell *et al.*, 2007) and has an important impact on horizontal gene transfer (Fitzgerald *et al.*, 2001; Molin and Tolker-Nielsen, 2003; Nanda *et al.*, 2015).

DNA damage, causing the induction of the cellular SOS response, represents the best-studied trigger for the switch of lysogenic phages to the lytic development (Oppenheim *et al.*, 2005). As a result of DNA damage, the occurrence of single-stranded DNA (ssDNA) is sensed by the protein RecA, which binds to ssDNA, oligomerizes and enters an active state, RecA*, in which it triggers the autoproteolysis of the SOS repressor protein LexA (Friedberg *et al.*, 2006). Derepression of the SOS genes results in the expression of more than 40 genes involved in DNA repair, recombination and inhibition of cell division. This mechanism is also exploited by lambdoid phages where the phage repressor

Accepted 23 July, 2015. *For correspondence. E-mail j.frunzke@fz-juelich.de, k.noeh@fz-juelich.de, d.kohlheyer@fz-juelich.de; Tel. +49 2461 615430, +49 2461 619294, +49 2461 612875; Fax +49(0)2461-612710. †Authors contributed equally to this work.

© 2015 John Wiley & Sons Ltd

protein CI mimics the catalytic center of LexA and, thus, is also cleaved upon induction of the SOS response leading to the expression of lytic phage genes and the irreversible switch to lytic development (Oppenheim *et al.*, 2005). In previous studies, single-cell analysis of reporter strains revealed a spontaneous activation of the SOS response caused by the sporadic occurrence of DNA damage (McCool *et al.*, 2004; Pennington and Rosenberg, 2007). It can be assumed that a prolonged induction of the SOS response by sporadic DNA damage will also impact the activation of lysogenic phages in a certain fraction of cells.

The Gram-positive soil bacterium *Corynebacterium glutamicum* represents an important platform organism of microbial biotechnology as it was successfully engineered for the production of a broad variety of value-added compounds, most importantly for the industrial production of amino acids and proteins (Eggeling and Bott, 2005). The genome of strain *C. glutamicum* ATCC 13032 contains three cryptic prophage elements (CGP1-3). The largest prophage CGP3 (187 kbp) was shown to be induced in a small fraction of cells under standard cultivation conditions (Frunzke *et al.*, 2008a). This element is integrated into a tRNA locus and is flanked by two conserved attachment sites (Kalinowski, 2005). Flow cytometry studies suggested spontaneous SOS induction as an important trigger of CGP3 SPI (Nanda *et al.*, 2014). Recently, a novel, CGP3-encoded actin-like protein was reported to play an important role in phage DNA replication upon induction of this element (Donovan *et al.*, 2015). However, the formation of infectious phage particle has not been observed so far. An update of the annotation of CGP3 genes is provided in Table S1.

Whereas the study of Friedman *et al.* already revealed a temporal modulation of the SOS response in single cells after UV radiation (Friedman *et al.*, 2005), only a snap-shot view of the spontaneous SOS induction in non-stressed bacterial populations was provided by previous studies (McCool *et al.*, 2004; Pennington and Rosenberg, 2007; Simmons *et al.*, 2009; Shee *et al.*, 2013; Nanda *et al.*, 2014). Recent advances in single-cell technologies are revolutionizing our understanding of biological systems by enabling insights into the dynamics and interaction of single microbial cells (Locke and Elowitz, 2009; Dusny and Schmid, 2014; Gr nberger *et al.*, 2014; Schallmey *et al.*, 2014; Weaver *et al.*, 2014). The fabrication and improvement of user-friendly microfluidic chip devices for microbial cultivation provide an important platform for state-of-the-art live cell imaging approaches. In previous reports, we have described the design of a disposable microfluidic chip device that enables the analysis of isogenic bacterial populations under defined environmental conditions and allows single-cell analysis by time-lapse fluorescence microscopy at spatiotemporal resolution (Gruenberger *et al.*, 2013).

In the present study, we have applied an improved setup to monitor the dynamics of the SOS response and prophage induction in single cells of *C. glutamicum* populations. This image-based single-cell analysis of reporter strains supported the hypothesis that the sporadic induction of the SOS response is an important trigger of CGP3 SPI, but also disclosed a considerable fraction of SOS-independent SPI. Whereas a transient pattern was observed with respect to the SOS reporter output, the induction of the cryptic prophage CGP3 led – in all cases – to a stop of growth and likely death of the affected cells. In combination with an automated image analysis workflow, this live cell imaging approach provides novel insights into the complex phenotypic structure of populations that is observed even under well-defined laboratory conditions.

Results

Construction of a dual SOS and phage reporter strain

Previous studies disclosed the spontaneous induction of the CGP3 prophage in *C. glutamicum* populations and revealed the SOS response as a prominent trigger of CGP3 SPI (Frunzke *et al.*, 2008a; Nanda *et al.*, 2014). The flow cytometry-based analysis of reporter strains, however, only provided a snap-shot view of the ongoing phenotypic dynamics. Therefore, we set out to monitor SOS response and prophage induction in single bacterial cells at spatiotemporal resolution using time-lapse microscopy. For the design of a suitable reporter strain, the promoter of *recA* (P_{recA}) and the promoter of a putative phage lysin (P_{lys}, cg1974) were fused to a gene encoding a fluorescent protein (Crimson, a red-fluorescent protein). The promoter for the design of the phage reporter was chosen on the basis of a previous transcriptome analysis showing the induction of CGP3 promoters upon prophage activation (Donovan *et al.*, 2015). First proof of principle experiments verified the functionality of the plasmid-based reporter constructs in *C. glutamicum* cultivations with or without the DNA damage-inducing antibiotic Mitomycin C (MMC) (Fig. 1, Fig. S1A). In the following, we used a *C. glutamicum* ATCC 13032 strain that contains a genomically integrated P_{recA}-*venus* fusion and a plasmid-encoded phage reporter where P_{lys} was fused to *crimson* (Fig. 1).

While both reporter constructs showed high sensitivity for MMC, the responsiveness was significantly decreased in a Δ*recA* mutant background (Fig. 1C and Fig. S1A). However, an increase in reporter signal is still observed in a *recA* mutant. This accumulation of reporter protein upon MMC treatment might be a result of the strongly reduced growth rate of this *recA* mutant, which is, in addition, also impaired in recombination as a prominent DNA repair strategy.

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

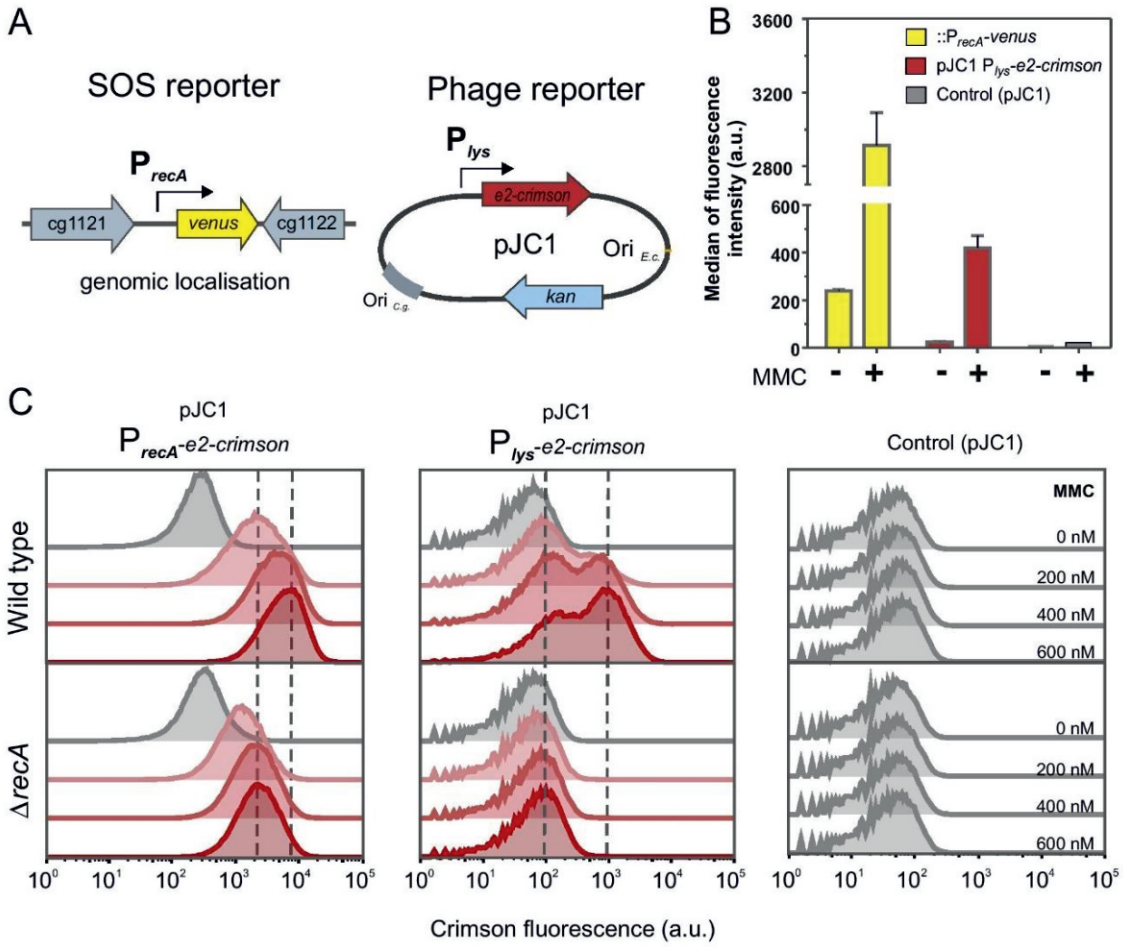


Fig. 1. Construction and verification of the reporter strain.
A. The reporter strain used in this study carries a genomically integrated SOS (P_{recA}-*venus*) and a plasmid-based phage reporter (P_{lys}-*e2-crimson*) construct.
B. The dual reporter strain shows a strongly increased activity of both reporters after 6 h cultivation in glucose minimal medium upon addition of 600 nM Mitomycin C (MMC) in comparison with the cells without MMC.
C. Comparison of the output of the SOS and phage reporter in wild type and Δ*recA* cells containing the plasmid pJC1-P_{recA}-*e2-crimson*, pJC1-P_{lys}-*e2-crimson* or the empty vector control after 6 h cultivation in glucose minimal medium upon addition of different concentrations of MMC. For a detailed analysis of reporter strains, see also Fig. S1.

Overall, this analysis of a *recA* mutant by flow cytometry and time-lapse microscopy already suggested a certain fraction of SOS-independent CGP3 induction (Fig. S1B, Video S4). Treatment with MMC resulted in a significant induction of the SOS reporter showing log-normally distributed single-cell fluorescence intensities. In contrast, the output of the phage reporter P_{lys}-*crimson* showed two peaks of reporter fluorescence reflecting the bimodal nature of the decision between the lysogenic and lytic state of CGP3 (Fig. 1C).

Spontaneous induction of SOS and prophage CGP3

To make time-lapse fluorescence microscopy under stable micro-environmental conditions possible, we performed growth experiments of *C. glutamicum* in microfluidic chip devices. The design of our in-house developed chip platform is shown in Fig. 2A (for a detailed description of the setup, see *Experimental procedures*). Single cells of the dual reporter strain *C. glutamicum* ATCC 13032::P_{recA}-*venus*/pJC1-P_{lys}-*crimson* were seeded into

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

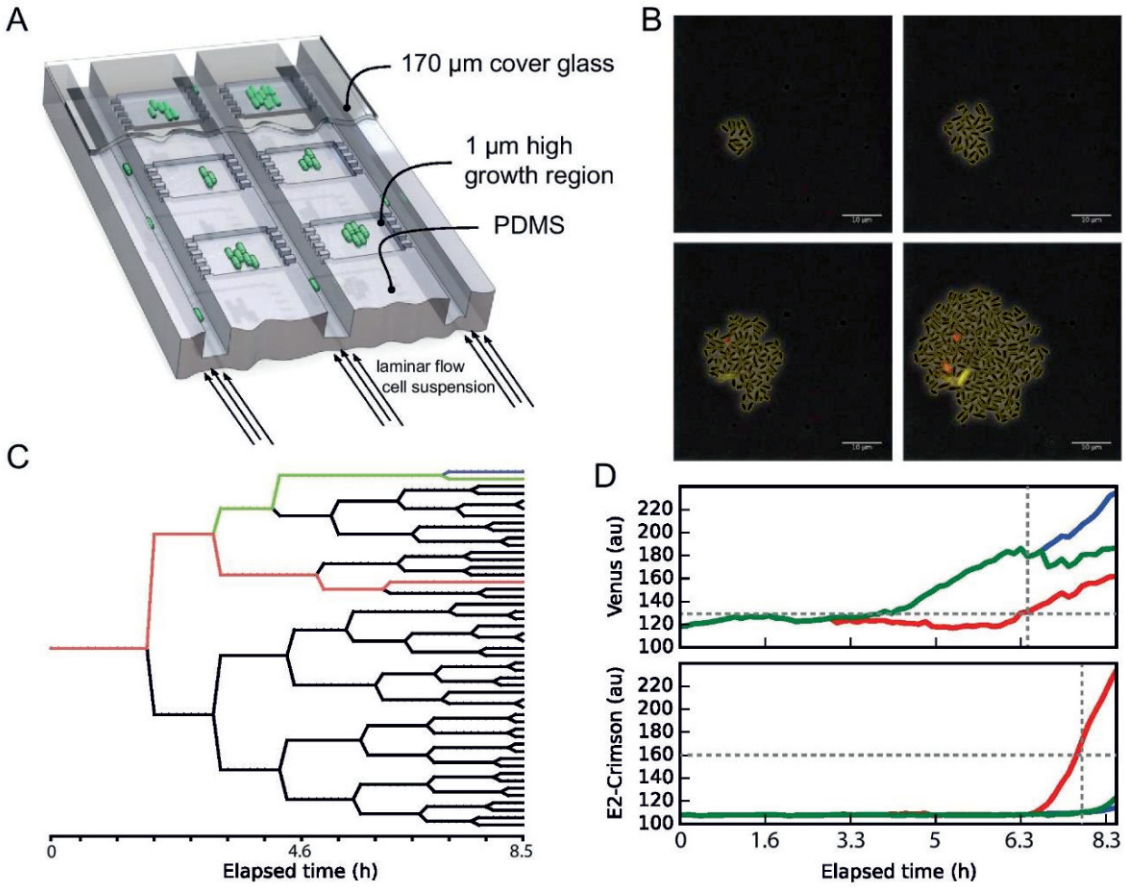


Fig. 2. Live Cell Imaging of SOS and prophage induction.
A. A schematic illustration of the cultivation chambers used in microfluidic lab-on-a-chip devices.
B. A series of processed images of a selected microcolony (MGC-16, see Video S1). Yellow contours denote the detected cell boundaries that were used for further processing.
C. Lineage tree of the colony show in B with highlighted traces of selected cellular events.
D. Fluorescence traces of Venus and Crimson reporters of traces highlighted in C. Dashed lines show time points where the reporter output passed the threshold (details on the computation are shown in Figs S4 and S5).

the chambers and growth and fluorescence in standard glucose minimal medium (CGXII + 4% glucose) was monitored for ~8 h (Fig. 2, Fig. S3 and Video S1–3). Lineage information (Fig. 2C) and single-cell fluorescence traces (Fig. 2D) were extracted for 20 independent microcolonies (4248 cells) with our semi-automated image analysis tool. Under the chosen conditions, an average growth rate of 0.47 h^{-1} was observed that is in the range of the growth rate of *C. glutamicum* determined in previous studies (Frunzke *et al.*, 2008b). To visualize the transient nature of the cellular SOS response, we applied a relatively low threshold for the Venus signal (T_{SOS} ~mean + 3 standard deviations; mean: 117.31 a.u., standard deviation: 3.85 a.u; Fig. S4). Using this experimental setup, 5.7% of the cells exhibited spontaneous induction of the SOS response and 0.7% displayed CGP3 SPI.

From continuous signal to discrete cellular state

Based on the (continuous) fluorescence output of both reporters (P_{recA}-*venus* and P_{lys}-*crimson*) and the ability of the cells to generate offspring in the course of the live cell imaging experiment, we have defined eight discrete cellular states (Fig. 3). This approach allowed us to investigate the fate of cells being in a particular state (e.g. SOS+ or Phage+), the transition frequency between states and frames as well as its impact on future growth (Offspring+/–). We constructed a model consisting of cel-

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

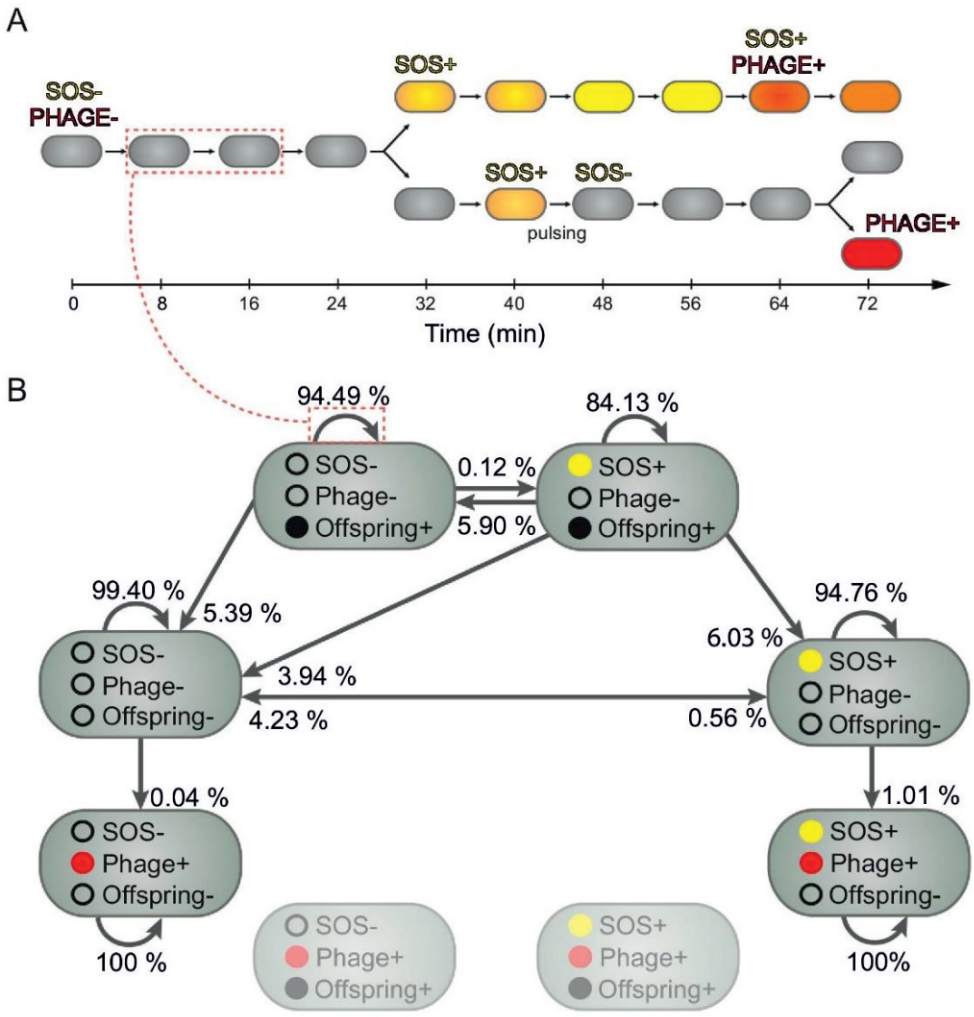


Fig. 3. Frequency of state transitions.
A. Illustration of the transitions between the four states based on the fluorescence reporter intensities (SOS–/PHAGE–, SOS+/PHAGE–, SOS–/PHAGE+, SOS+/PHAGE+). Every 8 min fluorescence images of all cultivation chambers were taken and compared with previous image frame. The red box shows the matching transition in the model.
B. Model of cellular states and transitions. Rounded rectangles denote states, while arrows denote possible transitions between the states. Arrows point from the initial state to the destination state. Values on the arrows show the conditional probability that a cell that was in the initial state at t_n is in the destination state at t_{n+1}. For details on data evaluation and the calculation of state transitions see Experimental procedures and Fig. S4, respectively.

lular states and possible transitions, in which transition probabilities are extracted from experimental data (Figs S4 and S5).

Transient induction of the SOS response

Previous studies have shown that the spontaneous induction of SOS-responsive promoters is likely due to the sporadic occurrence of DNA damage, including DNA double-strand breaks or stalled replication forks (Cox *et al.*, 2000; Pennington and Rosenberg, 2007; Nanda *et al.*, 2014). However, the time-resolved measurement of reporter output (e.g. using flow cytometry) obscures the transient nature of this DNA damage response. The frequency of state transitions from SOS+/Phage– to SOS–/Phage– (Fig. 3, Table 1) nicely illustrates the transient expression of the SOS reporter in a subset of SOS+ cells. Whereas the majority of SOS+ cells exhibits a prolonged

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

Table 1. Counted state transitions between SOS induction and SPI (between two subsequent image frames). A cell's overall states is encoded by three individual states, where '0' denotes the OFF state and '1' the ON state (order of positions: SOS, Phage, Offspring). (0,0,1), for instance, encodes the cellular state SOS–/Phage–/Offspring+.

		To							
		(0,0,0)	(0,0,1)	(0,1,0)	(0,1,1)	(1,0,0)	(1,0,1)	(1,1,0)	(1,1,1)
From	(0,0,0)	33515	0	15	0	188	0	0	0
	(0,0,1)	2433	42638	0	0	0	55	0	0
	(0,1,0)	0	0	227	0	0	0	0	0
	(0,1,1)	0	0	0	0	0	0	0	0
	(1,0,0)	71	0	0	0	1590	0	17	0
	(1,0,1)	16	24	0	0	25	342	0	0
	(1,1,0)	0	0	0	0	0	0	205	0
	(1,1,1)	0	0	0	0	0	0	0	0

induction, staying in an SOS+ state for more than one imaging interval, 18.6% of those cells recover from stress (re-entering the non-induced state and generating progeny) (Fig. 4B and C). Fig. 4C presents the ratio of SOS – cells originating from formerly SOS+ ancestor cells. This graph illustrates that, while a large fraction of SOS+ cells stops growth (Offspring-), a considerable number of SOS+ cells will resume growth and produce SOS– offspring thereby contributing to the overall growth and fitness of the population. Considering the short time-scale of the overall experiment, this percentage (18.6%) rather represents a lower limit of recovering cells. Whereas cells that recover from SOS induction show an overall reduced growth rate their growth rate is not significantly altered before induction in comparison to uninduced cells (Fig. S6).

The capability of cells to recover from SOS+ is also reflected by the fact that a measurable number (> 1) of offspring originate from SOS+ cells (Fig. 4B and C). This is particularly interesting considering that formerly SOS+ cells have suffered from an increased mutation rate, but resume growth and thereby significantly contribute to the overall mutation rate of the whole population. On the other hand, prolonged SOS induction is followed by the induction of the CGP3 prophage (SOS+/Phage+) in a considerable fraction of SOS+ cells (8.2% versus 0.7% in the overall experiment, and Fig. 5B).

CGP3 SPI leads to cell death/senescence

In contrast to the transient nature of the SOS induction, activation of the phage reporter coincides – in all cases – with a stop of cellular growth and likely cell death. This is in agreement with the cellular state and transition model, where we do not observe transitions into one of the Phage+/Offspring+ states (Fig. 3). Furthermore, the model topology (as extracted from the experimental data) shows that cells entering a Phage+ state remain in this state throughout the experiment (Figs 3 and 4D). Thus, this live cell imaging approach demonstrates that induction of the cryptic prophage CGP3 results in cell death/senescence of the affected cells.

Correlation of SOS and prophage induction

Previous flow cytometry-based studies suggested that spontaneous SOS induction is a prominent trigger of CGP3 SPI (Nanda *et al.*, 2014). In the present study, we analyzed the correlation of SOS and phage reporter output at spatiotemporal resolution in order to visualize how often these two responses correlate at the single-cell level. Therefore, we evaluated the temporal development of all single-cell traces exhibiting SOS and/or prophage induction over the course of the experiment and analyzed their correlation (Fig. 5). In agreement with previous findings, these data revealed that a significant fraction of SOS+ cells (> 8%) also exhibited prophage induction over the course of the experiment (Fig. 5B). Furthermore, SOS induction preceded CGP3 activation in > 60% of Phage+ cells, supporting a causal connection between these rare single-cell events (Fig. 5C). Remarkably, a considerable fraction of SOS-independent SPI (> 30%) was observed under the chosen experimental conditions (Fig. 5C). This is confirmed by the fact that SPI is also observed in about 0.5% ($n = 1551$) of the cells of a *recA* mutant strain containing a genomically integrated P_{lys}-*yfp* reporter (Video S4). Thus, these data clearly indicate the existence of further factors, besides the SOS response, that are involved in the activation of the cryptic prophage CGP3.

Generation-dependency of spontaneous SOS and prophage induction

Recent studies suggested that the majority of sporadic SOS inductions occurs in a generation-dependent manner, e.g. due to polymerases stalled at replication forks (Kuzminov, 2001). In the following, we set out to compare SPI and SOS induction in exponentially growing versus non-growing cells.

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

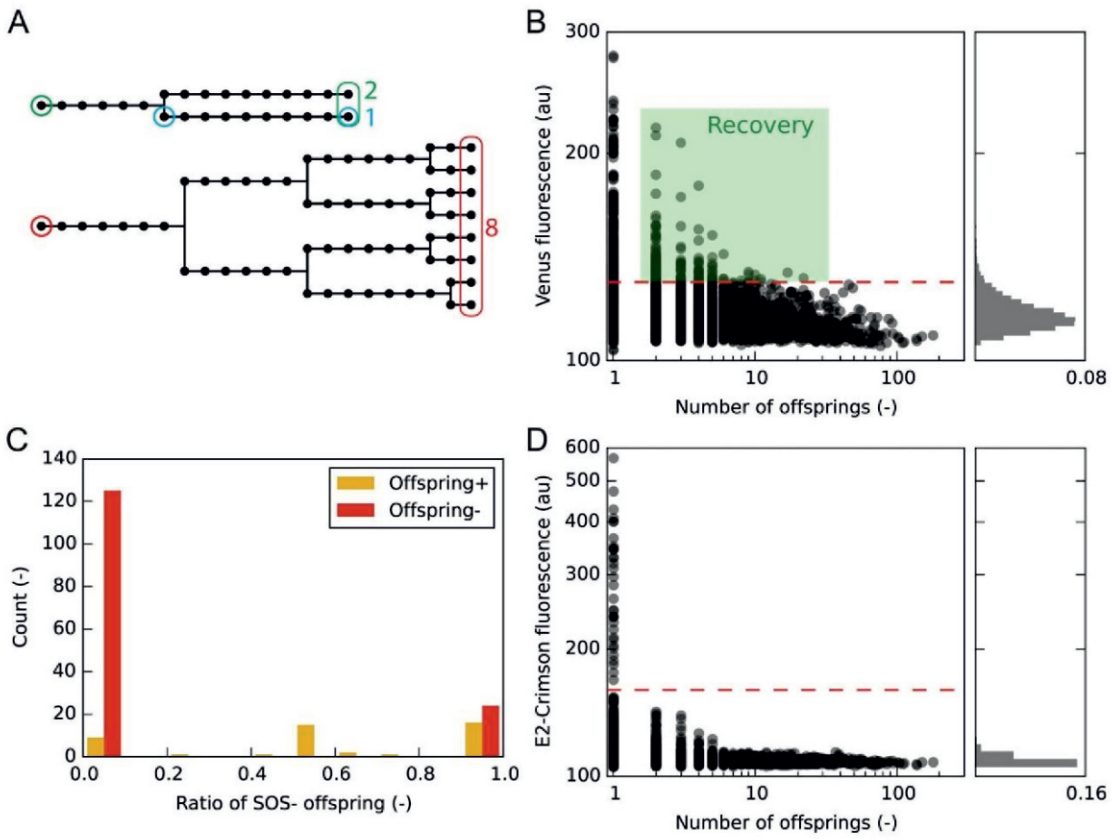


Fig. 4. Recovery of SOS+ and Phage+ cells.
A. Color-coded scheme for the computation of offspring of SOS+ or Phage+ cells. To compute the number of offspring of a cell *c*, the number of leaves of the subtree rooted at *c* is computed. Thus, cells that do not divide until the end of an experiment are denoted by an offspring of 1.
B. Scatter plot of the number of offspring versus the maximum Venus fluorescence (P_{recA}) of each generation. The histogram shows the distribution of the Venus fluorescence of all evaluated cells (normalized to bin width).
C. Histogram of the fraction of recovering offspring cells divided into cells with (yellow bars) and without offspring (red bars). Counted was the fraction of recovering (SOS–) cells originating from an SOS+ ancestor cell. 0.0: No SOS–offspring; 1.0: all cells originating from a SOS+ cell are SOS–.
D. Scatter plot of the number of offspring versus the maximum Crimson fluorescence (P_{lys}) per generation. The histogram shows the distribution of the Crimson fluorescence of all evaluated cells (normalized to bin width).

In order to discriminate between time- and generation-dependent induction of SOS and prophage, respectively, we performed a starvation experiment, where the supplied medium was switched after 8 h of exponential growth to minimal medium without carbon source. After ~24 h of starvation, the medium supply was switched back to standard glucose minimal medium and cells resumed growth. Whereas cells cultivated in the reference channels (+ carbon source) continued exponential growth, carbon-starved cells showed a significant drop in the growth rate but exhibited residual growth for the next ~12 h (Fig. 6A, Video S5). After approximately 12 h of starvation, the total cell number reached a plateau and no significant increase in population area was observed for the following ~10 h. Remarkably, even in this time span spontaneous SOS induction occurred at a lower but still measurable frequency (1.35 h^{-1}) suggesting the significant contribution of generation-independent mechanisms (Fig. 6C and D). In contrast, no SPI was observed in the starvation phase. This is in line with the finding that prophage induction is only triggered in actively proliferating cells (Pearl *et al.*, 2008).

As a reference experiment, we set out to conduct a long-term monitoring of SOS and prophage induction in exponentially growing cells. Therefore, we performed an experiment in a comb-like microfluidic structure (mother machine; see Wang *et al.*, 2010a). In this device, the progenies of single mother cells are continuously

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

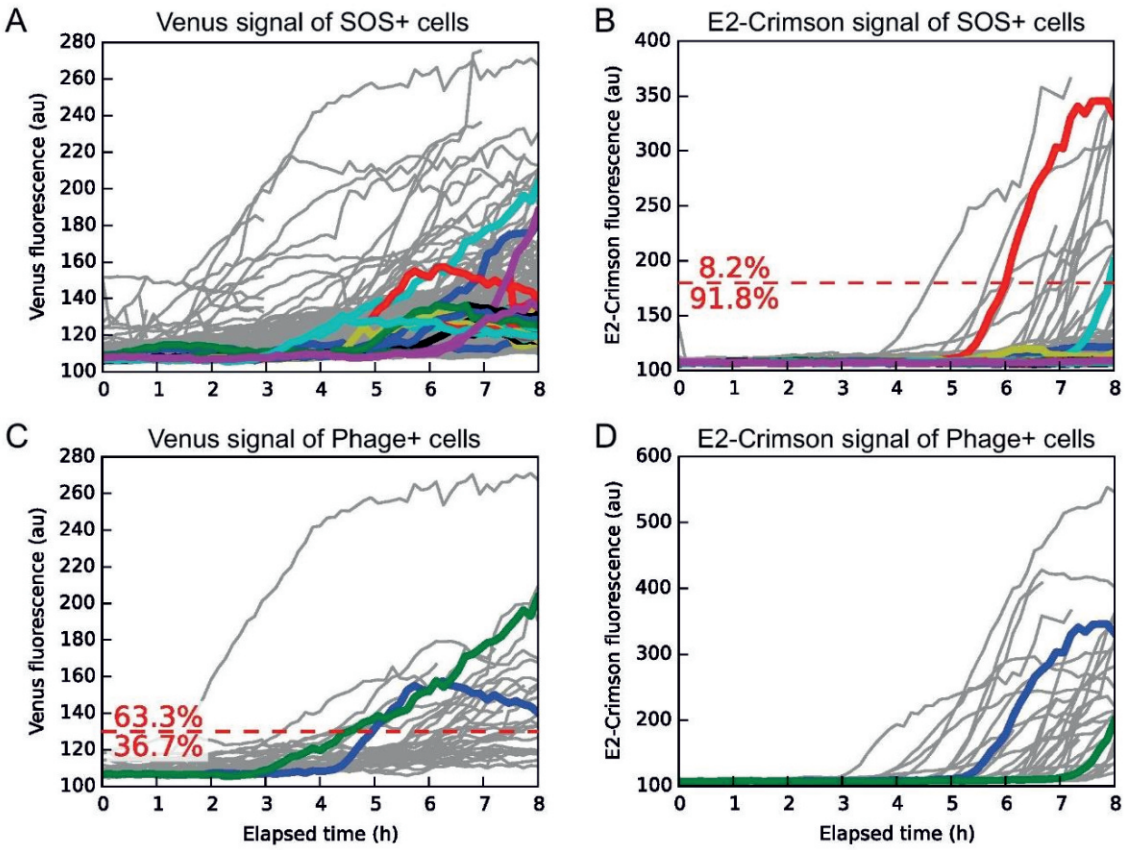


Fig. 5. Correlation between spontaneous SOS and prophage induction. Colored traces are highlighted as examples from the complete set of traces (gray lines).
A. Fluorescence traces (Venus) of all cells that have been identified as SOS+ during the experiment.
B. SOS-dependent prophage induction: Fluorescence traces (Crimson) of all cells identified as SOS+ during the experiment.
C. SOS-dependent and -independent SPI: Fluorescence traces (Venus) of all Phage+ cells.
D. Fluorescence traces (Crimson) of all Phage+ cells.

removed and thereby allow the long-term observation of exponential growth of the particular mother cells (Fig. 7, Video S6). The two fluorescence reporter signals were monitored for 24 h (Fig. 7). After the lag phase (~first 5 h) where an initial rise of reporter signals was observed, the number of SOS (~1.1%) and CGP3 (~0.7%) positive cells, respectively, stayed relatively constant throughout the measurement. These data emphasize that sporadic DNA damage and SPI occurs at a constant rate in actively proliferating cells.

Discussion

In the present study, we used a state of the art microfluidic chip/imaging platform to monitor SPI in *C. glutamicum* microcolonies at spatiotemporal resolution. These experiments provide detailed insights into the dynamics of prophage activation and SOS response at the single-cell level and highlight the value of single-cell approaches in comparison to standard bulk assays.

The host SOS response represents the by-far best-characterized trigger of lambdoid prophages (Friedberg *et al.*, 2006). Interestingly, previous single-cell studies showed that a small fraction of cells exhibits a spontaneous induction of this stress response – likely caused by the sporadic occurrence of DNA damage (McCool *et al.*, 2004; Pennington and Rosenberg, 2007; Simmons *et al.*, 2009; Shee *et al.*, 2013). Using a flow cytometry and microscopy-based analysis of reporter strains, the spontaneous induction of SOS-responsive promoters (*E. coli*: *sulA*, *umuCD*, *lexA* and *recA*, *C. glutamicum*: *recA*, *B. subtilis:* TagC-GFP) was measured to range between 0.09% and 3.1% (Pennington and Rosenberg, 2007; Simmons *et al.*, 2009; Kamensek *et al.*, 2010; Nanda

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

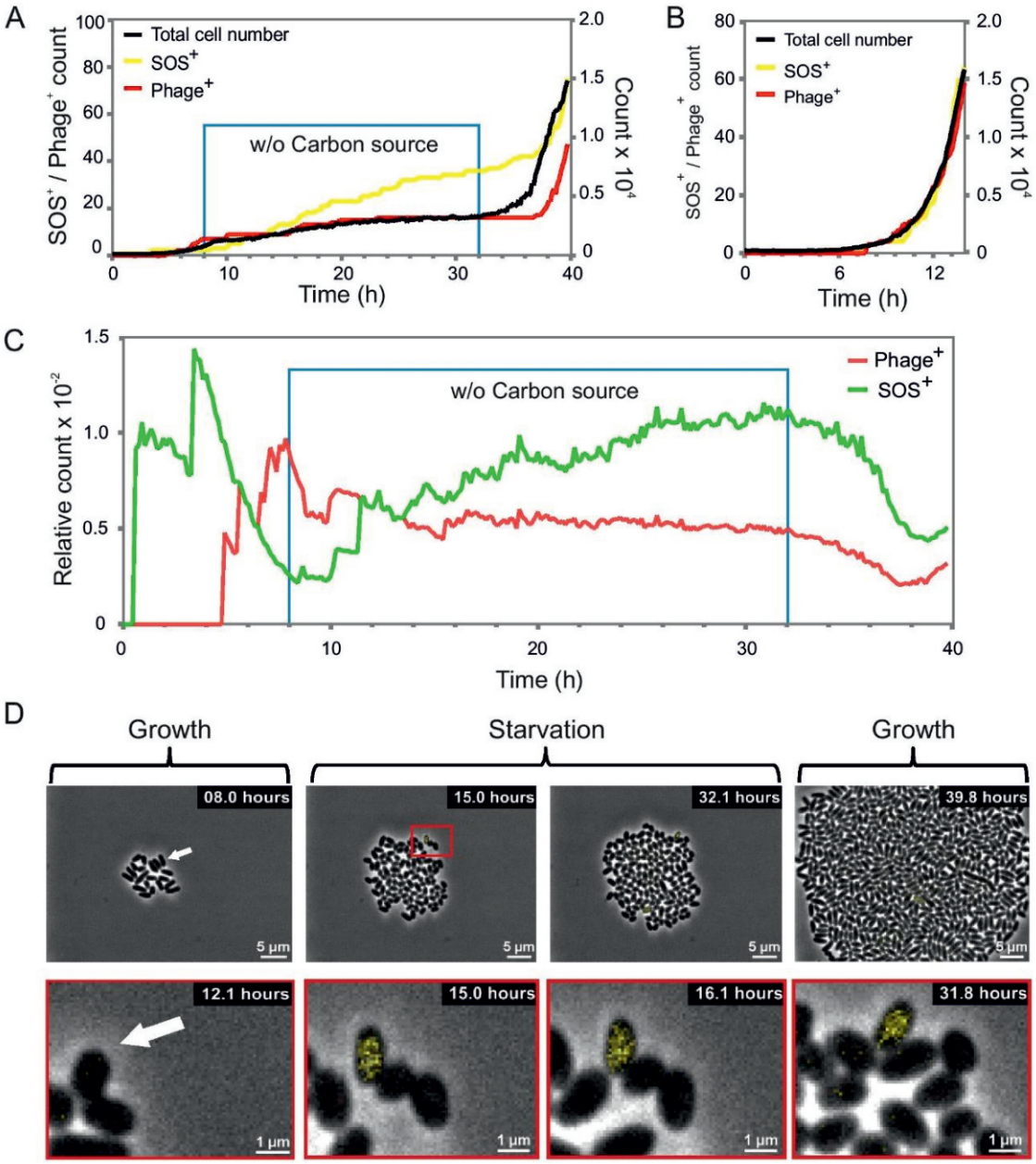


Fig. 6. Time-dependent SOS and prophage induction. Cells of the dual reporter strain ATCC 13032::P_{recA}-*venus*/pJC1-P_{lys}-*e2-crimson* were cultivated in CGXII with 4% glucose in the microfluidic device for 8 h. At this time point the medium was switched to minimal medium lacking any carbon source (w/o carbon source). After 24 h of starvation the medium was switched back to CGXII with 4% glucose (see also Video S5).

A. Counts of the total cell number, SOS positive and phage positive cells were plotted against the time. The starvation phase is indicated by the blue box.

B. Control experiment where the cells were grown in the presence of a carbon source in standard CGXII minimal medium.

C. The number of SOS and phage positive cells divided by the total cell number plotted against the time. During the starvation phase (blue box), SOS induction is still increasing, whereas no additional SPI events were observed.

D. Picture series of a selected microcolony cultivated as described in A. One cell exhibiting SOS induction during the starvation phase is highlighted (white arrow).

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

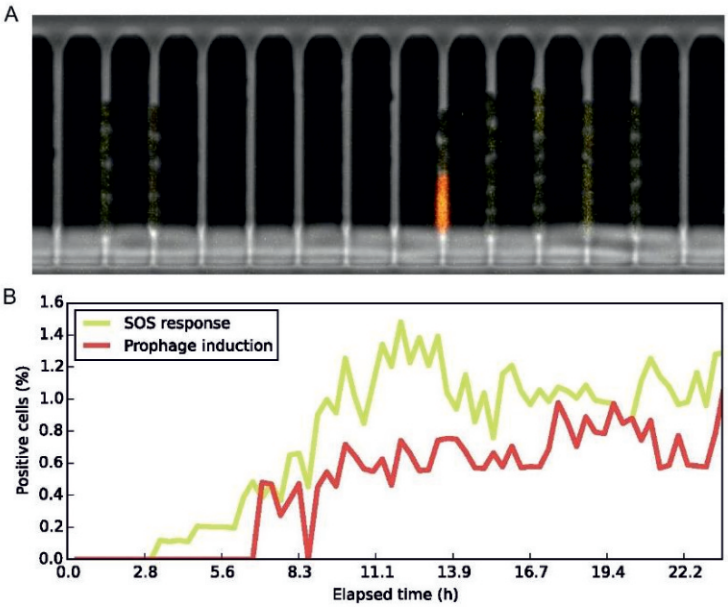


Fig. 7. Long-term study of SOS induction and SPI in exponentially growing cells.
A. An exemplary image frame of the long-term mother machine experiment conducted using the reporter strain *C. glutamicum* ATCC 13032::P_{recA}-*venus*/pJC1-P_{lys}-*e2-crimson*.
B. A time-resolved statistics of the percentage of cells that are SOS+ and Phage+, respectively. While the percentage increases in the first hours of the experiments, where several cells still resume growth from the lag phase, it stays constant from ~9 h until the end of the experiment.

et al., 2014). These numbers reflect the high dependency of absolute values of these kinds of measurements on the species, the experimental procedures and the data evaluation. More importantly, such values were generated to describe a stress response exhibiting a highly dynamic nature. A snap-shot measurement (e.g. flow cytometry analysis) of single cells at a certain time cannot reveal the fate, i.e. development of particular cells over time. In order to visualize these transient changes of the SOS response, we applied a rather low threshold (T_{SOS} ~mean + 3 standard deviations; mean: 117.31 a.u., standard deviation: 3.85 a.u.). This setting resulted in a high fraction of SOS+ cells in the present live cell imaging study (5.7%) in comparison with previous flow cytometry measurements (0.2%) where a rather stringent cutoff was applied (Nanda *et al.*, 2014). The setting of this threshold is, however, supported by the fact that SOS+ cells displayed in almost all cases a significant drop in growth and showed an elongated morphology typical for an induction of SOS genes encoding cell division inhibitor proteins (Huisman *et al.*, 1984; Ogino *et al.*, 2008).

Sporadic DNA damage, e.g. the spontaneous occurrence of DNA double-strand breaks, may for instance occur by the collapse of replication forks (Sassanfar and Roberts, 1990; Kuzminov, 2001) and is a frequent cause for genomic rearrangements or mutation hot spots (Drake, 2007a,b; Martincorena *et al.*, 2012). Under stress conditions this may significantly accelerate evolution of bacterial strains or communities (Al Mamun *et al.*, 2012; Rosenberg *et al.*, 2012). However, we also have to keep in mind that significant differences exist between species with respect to the contribution of the SOS response to the repair of double-strand breaks (Simmons *et al.*, 2009). In contrast to previous studies that have measured spontaneous SOS induction, the live cell imaging approach of this work allowed not only the quantification of the frequency of spontaneous induction but also measuring the fraction of cells (18.6%, Fig. 4B and C) that recovered from this stress. In the course of SOS induction, affected cells have likely suffered from an increased mutation rate, e.g. due to the action of error-prone polymerases. Consequently, recovering cells might have considerable impact on the overall mutation rate of the population; thereby also affecting its adaptability and potential to evolve (Galhardo *et al.*, 2007). A recovery of about 18.6% is in a similar range as previous measurement of *E. coli* SOS+ cells, where about 35% of the cells isolated via FACS formed colonies on plates (Pennington and Rosenberg, 2007). The authors concluded from negative propidium iodide staining that the rest of the SOS+ cells remained in a senescence-like, dormant state. However, previous studies ignored another factor impacted by a spontaneous induction of the SOS response, namely the activation of lambdoid prophages, mobile elements or cryptic (degenerated) prophages. In contrast to SOS, their induction often leads to a 'dead end'.

In this work, we could show that under the chosen assay conditions more than 8% of the SOS+ cells also induced the cryptic prophage CGP3 (Fig. 5B). In all cases, the signal of the SOS reporter preceded the signal of the phage reporter supporting the causal connection

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

between the SOS induction and prophage induction (Video S1–3). Due to the limited observation time, this value, however, rather represents a lower bound for SOS-induced CGP3 induction.

Remarkably, this study also demonstrated a considerable high fraction of SOS-independent CGP3 induction (> 30%, Fig. 5C). This is in line with the finding that a Δ*recA* mutant still exhibits spontaneous CGP3 induction (Fig. S1, Video S4). Alternative (RecA-independent) pathways for the induction of lambdoid prophages have been described in a few studies and include for instance the induction of *Pseudomonas aeruginosa* prophages by the accumulation of acyl-homoserine lactones (AHL) as a density-dependent mechanism (Ghosh *et al.*, 2009) or the RcsA- or DsrA-mediated induction described for *E. coli* (Rozanov *et al.*, 1998). The identification of SOS-independent induction mechanisms of CGP3 will be the target of future studies.

DNA replication has been previously reported as an important source for the occurrence of sporadic DNA damage, e.g. replication fork collapses (Sassanfar and Roberts, 1990; Kuzminov, 2001). Recent single-cell studies also have addressed the question whether spontaneous SOS induction is triggered by a generation- or time-dependent mechanism (Pennington and Rosenberg, 2007; Shee *et al.*, 2013). Those studies suggested that spontaneous DNA breakage is precisely correlated with the number of cell divisions (Shee *et al.*, 2013). In contrast, the starvation experiment conducted in this study revealed continued SOS induction in the starvation phase (Fig. 6). Given that colony area and cell number stayed constant for about 10 h, these data suggest that also time-dependent (generation-independent) mechanisms feed into the spontaneous induction of the SOS response. In fact, it is possible that, for instance, oxidative stress or changes in iron availability may cause spontaneous DNA damage in resting cells (Binnenkade *et al.*, 2014). However, even the spontaneous assembly of replisomes in a small fraction of non-growing cells is possible and might represent a source of spontaneous SOS induction. In contrast, CGP3 induction occurred only in proliferating cells which is in agreement with the finding of Pearl *et al.*, who reported that persistent cells, existing in a dormant state, are protected from prophage induction (Pearl *et al.*, 2008). In fact, persistence might represent an important mechanism of bacterial populations and communities to avoid complete eradication upon prophage induction under certain stressful conditions.

In all cases, SOS-dependent and SOS-independent, induction of the cryptic prophage CGP3 led to a stop of growth and likely cell death. If CGP3 induction, or, more generally, the induction of several (cryptic) prophages in other species, leads to cell death, why does this degenerated viral DNA remain in bacterial genomes in the course of evolution? Which selection pressure enforces the maintenance of degenerated prophage elements? Interestingly, several recent studies highlighted the beneficial effects prophages may have on the fitness of their bacterial host (reviewed in Bondy-Denomy and Davidson, 2014; Nanda *et al.*, 2015). Cryptic *E. coli* phages were reported to contribute significantly to the overall fitness and stress resistance of populations (Wang *et al.*, 2009; 2010b). Sacrificing a small fraction of cells by SPI is exploited by several bacterial strains for the development of biofilms (Rice *et al.*, 2008; Wang *et al.*, 2009; Zegans *et al.*, 2009; Carrolo *et al.*, 2010; G deke *et al.*, 2011), the release of toxins (Livny and Friedman, 2004) or host microbe interactions (Mitchell *et al.*, 2007; Matos *et al.*, 2013).

In conclusion, SPI represents a common but often unnoted phenomenon of lysogenic bacterial populations. This study provided a first insight into SPI at spatiotemporal resolution. Whereas our data confirmed the spontaneous induction of the SOS response as a prominent trigger, evidence was provided for further RecA-independent factors contributing to SPI in bacterial populations. Future studies will reveal how host–phage interaction has shaped the diversity of trigger inputs to adjust SPI to an optimal level depending on the particular environmental and physiological conditions of the host.

Experimental procedures

Bacterial strains and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 2. *E. coli* cells of the strain DH5α were cultivated in LB (Lysogeny Broth) medium or on agar plates at 37°C. For growth studies and fluorescence based assays with *C. glutamicum* (e.g. preparation of cells for flow cytometry) cells were first cultivated in BHI (brain heart infusion, Difco™ BHI, BD, Heidelberg, Germany) media at 30°C and were then used to inoculate a main culture in CGXII (Keilhauer *et al.*, 1993) with 2% glucose. When necessary, 50 μg ml^{-1} (*E. coli*) or 25 μg ml^{-1} (*C. glutamicum*) kanamycin was added.

Cloning techniques

Standard methods for example PCR or DNA restriction were performed according to established protocols (Sambrook and Russell, 2001). Gibson assembly was used for plasmid construction (Gibson *et al.*, 2008). DNA sequencing and oligonucleotides synthesis were conducted by Eurofins MWG Operon (Ebersberg, Germany). Plasmids and oligonucleotides used in this study are listed in Table 2 and Table S2 respectively. The chromosomal integration of the SOS reporter (P_{recA}-*venus*) was performed using the two-step homologues recombination method (Niebisch and Bott, 2001). Correct integration into the intergenic region of cg1121 and cg1122 was verified by colony PCR with the oligonucleotides Cg1121-Int-fw and Cg1122-Int-rv. Construction of the *recA* mutant of was performed analogously to the genomic integration by homologues recom-

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

Table 2. Bacterial strains and plasmids used in this study.

Strains	Relevant characteristics	Reference
E. coli DH5α	*supE44 ΔlacU169* (φ80*lacZ*ΔM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Invitrogen
C. glutamicum ATCC 13032	Biotin-auxotrophic wild type	Kinoshita *et al.* (2004)
ATCC 13032::P_{recA}-*venus*	ATCC 13032 derivative, promoter of *recA* was fused to *venus* and integrated into the cg1121-cg1122 intergenic region	This study
ATCC 13032 Δ*recA*	In frame deletion of the ATPase domain of *recA* (cg2141)	This study
ATCC 13032::P_{tac}-*venus*	Derivative of ATCC 13032, *lacI* gene and *tac* promoter fused to *venus* were integrated into the intergenic region of cg1121-cg1122	This study
ATCC 13032::P_{lys}-*eyfp*	Derivative of ATCC 13032 containing the prophage reporter P_{lys}-*eyfp* integrated into the intergenic region of cg1121-cg1122	This study
ATCC 13032 Δ*recA*::P_{lys}-*eyfp*	Derivative of ATCC 13032 P_{lys}-*eyfp*, in frame deletion of ATPase domain of *recA* (cg2141)	This study
Plasmids		
pK18mobsacB-cg1121/1122	Kan^R; plasmid for integration of foreign DNA into the intergenic region between cg1121-cg1122 ($oriV_{E.c.}$, *sacB*, *lacZ*α)	Baumgart *et al.* (2013)
pK18mobsacB-cg1121/1122-P_{recA}-*venus*	Derivative of pK18mobsacB-cg1121/1122; plasmid for integration into the intergenic region of cg1121-cg1122, the *recA* promoter was fused to *venus*	This study
pK18mobsacB-cg1121/1122-P_{lys}-*eyfp*	Derivative of pK18mobsacB-cg1121/1122; plasmid for integration of the *lysin* promoter fused to *eyfp* into the intergenic region of cg1121-cg1122	This study
pK18mobsacB-cg1121/1122-*lacI* P_{tac}-*venus*	Derivative of pK18mobsacB-cg1121/1122; plasmid for integration of *lacI* and *tac* promoter fused to *venus* in the intergenic region of cg1121-cg1122	This study
pK19*mobsacB*	Kan^R.; plasmid for allelic exchange in *C. glutamicum*; (pK18 $oriV_{E.c.}$, *sacB*, *lacZ*α)	Sch fer *et al.* (1994)
pK19mobsacB-Δ*recA*	Derivative of pK19*mobsacB*; plasmid for in frame deletion of the ATPase domain (A7 – E238) in *recA*	This study
pJC1	Kan^R, Amp^R, *C. glutamicum* shuttle vector	Cremer *et al.* (1990)
pJC1-P_{lys}-*e2-crimson*	pJC1 derivative containing the promoter region of a putative *lysin* gene within the CGP3 prophage region (cg1974), included are 30 bp of the coding sequence, a stop codon and an additional ribosome binding site (pET16) in front of *e2-crimson*	Nanda *et al.* (2014)
pJC1-*lacI*-P_{tac}-*crimson*	pJC1 derivative containing the *lacI* gene and a fusion of the *tac* promoter to *e2-crimson*	This study

bination (Niebisch and Bott, 2001). Correct deletion of the ATPase domain of *recA* was verified by colony PCR using the oligonucleotides D_recA_5 and D_recA_6.

Flow cytometry

Flow cytometry analysis and cell sorting experiments were performed with a FACSAria II flow cytometer (BD, Heidelberg, Germany). A blue solid state laser with an excitation wavelength of 488 nm (to excite Venus) and a red gas laser for excitation at a wavelength of 633 nm (to excite E2-Crimson) were used. Cytometer set-up and performance tracking was conducted with tracking beads labeled with a mixture of fluorochromes (BD, Heidelberg, Germany). Forward-scatter characteristics (FSC) and side-scatter characteristics (SSC) were detected as small and large angle scatters of the 488 nm laser. Fluorescence emitted by Venus was detected using a 502 nm long-pass and a 530/30 nm band pass filter set. E2-Crimson fluorescence was detected using a 660/20 nm band pass filter set. Analyses of cells were performed at a threshold rate of 2,000 to 5,000 events s^{-1}. Data were analyzed using FlowJo V10 (Tree Star, Ashland, OR).

Microfluidic device cultivation

An in-house developed microfluidic platform was used for *C. glutamicum* single-cell analysis (Gr nberger *et al.*, 2012; Gruenberger *et al.*, 2013). The microfluidic device incorporates a few hundred cultivation chambers with dimensions to ensure monolayer growth of isogenic microcolonies, with up to a few hundred cells maximum each. Phase contrast and fluorescence time-lapse imaging was performed at 8 min intervals.

For continuous cultivations, a dead-end channel configuration (mother machine) was utilized with thousands of microchannels arranged perpendicular to the main medium flow (Wang *et al.*, 2010a). For each channel, the exponential growth of the mother cells was followed for 24 h. The flow rate was adjusted to 1000 nL min^{-1} and images were taken in 5 min (phase contrast) and 20 min (epifluorescence) intervals respectively.

In both devices, medium was supplied continuously to ensure stable and constant environmental conditions. Minimal medium (CGXII + 4% glucose) with addition of 25 μg ml^{-1} kanamycin was infused at a rate of 300 nL min^{-1} with a high-precision syringe pump (neMESYS, Cetoni GmbH, Korbussen, Germany). For initiation of the starvation phase the medium was switched to minimal medium lacking glucose and protocatechuate (–carbon source) after an initial growth phase of 8 h. Cells were exposed to carbon limitation for ~24 h. Then growth on standard CGXII minimal medium (+ carbon source) was resumed. A constant cultivation temperature of 30°C was ensured by an incubation chamber (PeCon GmbH, Erbach, Germany).

The microfluidic chip was mounted on a fully-motorized, inverted epifluorescence microscope (TI-Eclipse, Nikon

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

Staphylococcus aureus encoding multiple superantigens. *J Bacteriol* **183:** 63–70.

Friedberg, E., Walker, G., Siede, W., and Schultz, R. (2006) *DNA Repair and Mutagenesis*. Washington: ASM Press.

Friedman, N., Vardi, S., Ronen, M., Alon, U., and Stavans, J. (2005) Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. *PLoS Biol* **3:** e238.

Frunzke, J., Bramkamp, M., Schweitzer, J.E., and Bott, M. (2008a) Population heterogeneity in *Corynebacterium glutamicum* ATCC 13032 caused by prophage CGP3. *J Bacteriol* **190:** 5111–5119.

Frunzke, J., Engels, V., Hasenbein, S., Gatgens, C., and Bott, M. (2008b) Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2. *Mol Mirobiol* **67:** 305–322.

Galhardo, R.S., Hastings, P.J., and Rosenberg, S.M. (2007) Mutation as a stress response and the regulation of evolvability. *Crit Rev Biochem Mol Biol* **42:** 399–435.

Ghosh, D., Roy, K., Williamson, K.E., Srinivasiah, S., Wommack, K.E., and Radosevich, M. (2009) Acyl-homoserine lactones can induce virus production in lysogenic bacteria: an alternative paradigm for prophage induction. *Appl Environ Microbiol* **75:** 7142–7152.

Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma* genitalium genome. *Science* **319:** 1215–1220.

G deke, J., Paul, K., Lassak, J., and Thormann, K.M. (2011) Phage-induced lysis enhances biofilm formation in *Shewanella oneidensis* MR-1. *ISME* **5:** 613–626.

Gruenberger, A., Probst, C., Heyer, A., Wiechert, W., Frunzke, J., and Kohlheyer, D. (2013) Microfluidic picoliter bioreactor for microbial single-cell analysis: fabrication, system setup, and operation. *J Vix Exp* 50560. doi: 10.3791/50560.

Gr nberger, A., Paczia, N., Probst, C., Schendzielorz, G., Eggeling, L., Noack, S., *et al.* (2012) A disposable picolitre bioreactor for cultivation and investigation of industrially relevant bacteria on the single cell level. *Lab Chip* **12:** 2060–2068.

Gr nberger, A., Wiechert, W., and Kohlheyer, D. (2014) Single-cell microfluidics: opportunity for bioprocess development. *Curr Opin Biotechnol* **29:** 15–23.

Huisman, O., D'Ari, R., and Gottesman, S. (1984) Cell-division control in *Escherichia coli*: specific induction of the SOS function SfiA protein is sufficient to block septation. *Proc Natl Acad Sci USA* **81:** 4490–4494.

Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S.L., and Danuser, G. (2008) Robust single-particle tracking in live-cell time-lapse sequences. *Nat Methods* **5:** 695–702.

Kalinowski, J. (2005) The genomes of amino acid-producing *Corynebacteria*. In *Handbook of* Corynebacterium glutamicum. Eggeling, L., and Bott, M. (eds). Boca Raton, FL: Academic Press, pp. 35–56.

Kamensek, S., Podlesek, Z., Gillor, O., and Zgur-Bertok, D. (2010) Genes regulated by the *Escherichia coli* SOS repressor LexA exhibit heterogeneous expression. *BMC Microbiol* **10:** 283.

Keilhauer, C., Eggeling, L., and Sahm, H. (1993) Isoleucine synthesis in *Corynebacterium glutamicum:* molecular analysis of the *ilvB-ilvN-ilvC* operon. *J Bacteriol* **175:** 5595–5603.

Kinoshita, S., Udaka, S., and Shimono, M. (2004) Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms. *J Gen Appl Microbiol* **50:** 331–343.

Kuzminov, A. (2001) Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci USA* **98:** 8241–8246.

Livny, J., and Friedman, D.I. (2004) Characterizing spontaneous induction of Stx encoding phages using a selectable reporter system. *Mol Mirobiol* **51:** 1691–1704.

Locke, J.C.W., and Elowitz, M.B. (2009) Using movies to analyse gene circuit dynamics in single cells. *Nat Rev Microbiol* **7:** 383–392.

Lwoff, A. (1953) Lysogeny. *Bacteriol Rev* **17:** 269–332.

McCool, J.D., Long, E., Petrosino, J.F., Sandler, H.A., Rosenberg, S.M., and Sandler, S.J. (2004) Measurement of SOS expression in individual *Escherichia coli* K-12 cells using fluorescence microscopy. *Mol Microbiol* **53:** 1343–1357.

Martincorena, I., Seshasayee, A.S., and Luscombe, N.M. (2012) Evidence of non-random mutation rates suggests an evolutionary risk management strategy. *Nature* **485:** 95–98.

Matos, R.C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., *et al.* (2013) *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. *PLoS Genet* **9:** e1003539–e1003539.

Mitchell, J., Siboo, I.R., Takamatsu, D., Chambers, H.F., and Sullam, P.M. (2007) Mechanism of cell surface expression of the *Streptococcus mitis* platelet binding proteins PblA and PblB. *Mol Microbiol* **64:** 844–857.

Molin, S., and Tolker-Nielsen, T. (2003) Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Curr Opin Biotechnol* **14:** 255–261.

Nanda, A.M., Heyer, A., Kr mer, C., Gr nberger, A., Kohlheyer, D., and Frunzke, J. (2014) Analysis of SOS-induced spontaneous prophage induction in *Corynebacterium glutamicum* at the single-cell level. *J Bacteriol* **196:** 180–188.

Nanda, A.M., Thormann, K.M., and Frunzke, J. (2015) Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *J Bacteriol* **197:** 410–419.

Neely, M.N., and Friedman, D.I. (1998) Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin. *Gene* **223:** 105–113.

Niebisch, A., and Bott, M. (2001) Molecular analysis of the cytochrome bc1-aa3 branch of the *Corynebacterium glutamicum* respiratory chain containing an unusual diheme cytochrome c1. *Arch Microbiol* **175:** 282–294.

Ogino, H., Teramoto, H., Inui, M., and Yukawa, H. (2008) DivS, a novel SOS-inducible cell-division suppressor in *Corynebacterium glutamicum*. *Mol Mirobiol* **67:** 597–608.

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., and

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

Adhya, S. (2005) Switches in bacteriophage lambda development. *Annu Rev Genetics* **39:** 409–429.
Pearl, S., Gabay, C., Kishony, R., Oppenheim, A., and Balaban, N.Q. (2008) Nongenetic individuality in the host-phage interaction. *PLoS Biol* **6:** e120–e120.
Pennington, J.M., and Rosenberg, S.M. (2007) Spontaneous DNA breakage in single living *Escherichia coli* cells. *Nat Genetics* **39:** 797–802.
Rice, S.A., Tan, C.H., Mikkelsen, P.J., Kung, V., Woo, J., Way, M., *et al.* (2008) The biofilm life cycle and virulence of *Pseudomonas aeruginosa* are dependent on a filamentous prophage. *ISME* **3:** 271–282.
Rosenberg, S.M., Shee, C., Frisch, R.L., and Hastings, P.J. (2012) Stress-induced mutation via DNA breaks in *Escherichia coli*: a molecular mechanism with implications for evolution and medicine. *Bioessays* **34:** 885–892.
Rozanov, D.V., D'Ari, R., and Sineoky, S.P. (1998) RecA-independent pathways of lambdoid prophage induction in *Escherichia coli*. *J Bacteriol* **180:** 6306–6315.
Sambrook, J., and Russell, D. (2001) *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.
Sassanfar, M., and Roberts, J.W. (1990) Nature of the SOS-inducing signal in *Escherichia coli*. The involvement of DNA replication. *J Mol Biol* **212:** 79–96.
Schallmey, M., Frunzke, J., Eggeling, L., and Marienhagen, J. (2014) Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. *Curr Opin Biotechnol* **26:** 148–154.
Sch fer, A., Tauch, A., J ger, W., Kalinowski, J., Thierbach, G., and P hler, A. (1994) Small mobilizable multipurpose cloning vectors derived from the *Escherichia-Coli* plasmids Pk18 and Pk19 – selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145:** 69–73.
Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9:** 676–682.
Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9:** 671–675.
Shee, C., Cox, B.D., Gu, F., Luengas, E.M., Joshi, M.C., Chiu, L.Y., *et al.* (2013) Engineered proteins detect spontaneous DNA breakage in human and bacterial cells. *Elife* **2:** e01222.
Simmons, L.A., Goranov, A.I., Kobayashi, H., Davies, B.W., Yuan, D.S., Grossman, A.D., and Walker, G.C. (2009) Comparison of responses to double-strand breaks between *Escherichia coli* and *Bacillus subtilis* reveals different requirements for SOS induction. *J Bacteriol* **191:** 1152–1161.
Waldor, M.K., and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272:** 1910–1914.
Wang, P., Robert, L., Pelletier, J., Dang, W.L., Taddei, F., Wright, A., and Jun, S. (2010a) Robust growth of *Escherichia coli*. *Curr Biol* **20:** 1099–1103.
Wang, X., Kim, Y., and Wood, T.K. (2009) Control and benefits of CP4-57 prophage excision in *Escherichia coli* biofilms. *ISME* **3:** 1164–1179.
Wang, X., Kim, Y., Ma, Q., Hong, S.H., Pokusaeva, K., Sturino, J.M., and Wood, T.K. (2010b) Cryptic prophages help bacteria cope with adverse environments. *Nature Comms* **1:** 147–147.
Weaver, W.M., Tseng, P., Kunze, A., Masaeli, M., Chung, A.J., Dudani, J.S., *et al.* (2014) Advances in high-throughput single-cell microtechnologies. *Curr Opin Biotechnol* **25:** 114–123.
Zegans, M.E., Wagner, J.C., Cady, K.C., Murphy, D.M., Hammond, J.H., and O'Toole, G.A. (2009) Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J Bacteriol* **191:** 210–219.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

© 2015 John Wiley & Sons Ltd, *Molecular Microbiology*

3.2 Silencing of cryptic prophages in *Corynebacterium glutamicum*

Pfeifer E.[1], Hünnefeld M.[1], Popa O.[2], Polen T.[1], Kohlheyer D.[1], Baumgart M.[1], and Frunzke J.[1]*

[1]IBG-1: Biotechnology, Forschungszentrum Jülich, Jülich, Germany

[2]Quantitative and Theoretical Biology, Heinrich-Heine University, Düsseldorf, Germany

*Corresponding author

Name of Journal: *Nucleic Acids Research*

Impact factor: *9.202*

Author contributions

Own contribution to the work: 80%

Contribution to project design and manuscript

	Name	Contribution
Project planning	**Pfeifer, E.**	70%
	Frunzke, J.	20%
	Polen, T.	5%
	Kohlheyer, D.	5%
Writing	Pfeifer, E.	65%
	Frunzke, J.	25%
	Baumgart, M.	2.5%
	Polen, T.	2.5%
	Kohlheyer, D.	2.5%
	Popa, O.	2.5%

Contribution to experiments, their evaluation and illustration

	Name	Experimental work	Evaluation	Processing & Preparation of figure/ table
Figure 1	**Pfeifer, E.**	100%	70%	100%
	Frunzke, J.	-	30%	-
	Baumgart, M	-	10%	-
Figure 2	**Pfeifer, E.**	80%	100%	100%
	Hünnefeld, M.	20%	-	-
Figure 3	**Pfeifer, E.**	60%	100%	100%
	Hünnefeld, M.	40%	-	-
Figure 4	**Pfeifer, E.**	40%	100%	100%
	Hünnefeld, M.	60%	-	-
Figure 5	**Pfeifer, E.**	50%	100%	80%
	Hünnefeld, M.	50%		20%
Figure 6	**Pfeifer, E.**	50%	80%	70%
	Hünnefeld, M.	50%	20%	30%
Figure 7	**Pfeifer, E.**	-	20%	10%
	Popa, O.	-	60%	90%
	Frunkze, J.	-	20%	-

	Name	Experimental work	Evaluation	Processing & Preparation of figure/ table
Table S1	**Pfeifer, E**. Hünnefeld, M.	80% 20%	- -	100% -
Table S2	**Pfeifer, E**. Hünnefeld, M.	80% 20%	- -	100% -
Table S3	**Pfeifer, E**. Hünnefeld, M.	80% 20%	- -	100% -
Figure S1	**Pfeifer, E**.	-	100%	100%
Figure S2	**Pfeifer, E**.	100%	100%	100%
Figure S3	**Pfeifer, E.**	-	100%	100%
Figure S4	**Pfeifer, E.**	Based on Fig. 3	100%	100%
Figure S5	**Pfeifer, E.**	Based on Fig. 3	100%	100%
Figure S6	**Pfeifer, E.**	100%	100%	100%
Figure S7	Hünnefeld, M.	100%	100%	100%
Figure S8	Hünnefeld, M.	100%	100%	100%
Figure S9	**Pfeifer, E.**	100%	100%	100%
Figure S10	**Pfeifer, E.** Popa, O. Frunzke, J.	- -	10% 70% 20%	10% 90% -
Video S1 & S2	**Pfeifer, E.**	100%	100%	100%
Table S4	**Pfeifer, E.** Polen, T.	Based on Fig. 3 -	70% 30%	70% 30%
Table S5	**Pfeifer, E.** Hünnefeld, M.	Based on Fig. 5	70% 30%	100% -
Table S6	**Pfeifer, E.** Popa, O. Frunzke, J.	- -	10% 70% 20%	10% 90% -

Nucleic Acids Research Advance Access published August 4, 2016

Nucleic Acids Research, 2016 **1**
doi: 10.1093/nar/gkw692

Silencing of cryptic prophages in *Corynebacterium glutamicum*

Eugen Pfeifer[1], Max Hünnefeld[1], Ovidiu Popa[2], Tino Polen[1], Dietrich Kohlheyer[1], Meike Baumgart[1] and Julia Frunzke[1,*]

[1]Institute of Bio- und Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, 52425 Jülich, Germany and [2]Quantitative and Theoretical Biology, Heinrich-Heine-Universität Düsseldorf, 40225, Düsseldorf, Germany

Received May 11, 2016; Revised July 25, 2016; Accepted July 26, 2016

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

ABSTRACT

DNA of viral origin represents a ubiquitous element of bacterial genomes. Its integration into host regulatory circuits is a pivotal driver of microbial evolution but requires the stringent regulation of phage gene activity. In this study, we describe the nucleoid-associated protein CgpS, which represents an essential protein functioning as a xenogeneic silencer in the Gram-positive *Corynebacterium glutamicum*. CgpS is encoded by the cryptic prophage CGP3 of the *C. glutamicum* strain ATCC 13032 and was first identified by DNA affinity chromatography using an early phage promoter of CGP3. Genome-wide profiling of CgpS binding using chromatin affinity purification and sequencing (ChAP-Seq) revealed its association with AT-rich DNA elements, including the entire CGP3 prophage region (187 kbp), as well as several other elements acquired by horizontal gene transfer. Countersilencing of CgpS resulted in a significantly increased induction frequency of the CGP3 prophage. In contrast, a strain lacking the CGP3 prophage was not affected and displayed stable growth. In a bioinformatics approach, *cgpS* orthologs were identified primarily in actinobacterial genomes as well as several phage and prophage genomes. Sequence analysis of 618 orthologous proteins revealed a strong conservation of the secondary structure, supporting an ancient function of these xenogeneic silencers in phage-host interaction.

INTRODUCTION

Viral DNA, in the form of functional prophages or degenerated (cryptic) phage elements, is ubiquitously found in bacterial genomes and may constitute up to 20% of the host genome (1–3). The mosaic-like structure of bacterial genomes indicates that phage-mediated horizontal gene transfer is a pivotal driver of bacterial evolution (4). Recent studies demonstrated that these elements might contribute significantly to the fitness of their respective host by improving stress tolerance, antibiotic resistance, biofilm formation or virulence (5,6). Phage-mediated gene transfer may provide the cell with novel adaptive traits, improving the fitness of the receptor cell, but this does not occur without risks. The integration of selfish replicators, including transposable elements, integrative/conjugative elements (ICE) or phages, can lead to high transcriptional and translational costs or even cell death (7,8). Hence, bacteria possess a number of different systems that confer resistance to foreign genetic elements, e.g. CRISPR/Cas and restriction modification (R-M) systems (9,10).

However, to harness the adaptive potential of foreign DNA and enable its integration into the host regulatory circuitry, bacteria have evolved a rather mediative mechanism called xenogeneic silencing (XS) (11–13). This mechanism relies on the function of small nucleoid-associated proteins (NAPs) to target and inhibit the expression of foreign DNA, which is recognizable by its typically higher AT content in comparison to the host genome (1,14). The major role of XS proteins is the binding of foreign DNA elements and the inhibition of transcription by a complex formation of AT-rich DNA stretches causing either the occlusion or trapping of the RNA polymerase (15,16). Currently known XS proteins belong to one of four classes, consisting of H-NS-type proteins found in several proteobacteria (12,17), Lsr2-like proteins of the actinomycetes (18), MvaT of *Pseudomonas* species (16) and Rok of *Bacillus subtilis* (19).

To date, most studies have focused on host-encoded XS proteins acting as silencers of foreign DNA. However, it may also be of benefit for the foreign element to bring its own silencer protein to improve tolerance within the host cell. Here, we describe a novel prophage-encoded XS protein of the Lsr2-type in *Corynebacterium glutamicum* ATCC 13032. The genome of this important industrial amino acid producer contains three cryptic prophages (20,21). Whereas CGP1 and CGP2 are highly degenerated, CGP3 comprises almost 6% of the entire genome (187 kb) and is inducible in an SOS-dependent manner (22,23). Even under

*To whom correspondence should be addressed. Tel: +49 2461 615430; Email: j.frunzke@fz-juelich.de

© The Author(s) 2016. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

non-inducing conditions, spontaneous prophage induction (SPI) was observed, preceded by a spontaneous activation of the SOS response in >60% of cases (20,22,23). However, the precise regulatory control of CGP3 induction has not been studied thus far.

In this study, we demonstrate the essential role of a prophage-encoded NAP, which is a homolog to the mycobacterial Lsr2 protein and functions as a silencer of cryptic phage elements in *C. glutamicum* (CgpS, *C. glutamicum* prophage silencer). Genome-wide profiling of the CgpS–DNA interaction revealed its association with AT-rich DNA regions located primarily within prophage regions. Countersilencing of CgpS activity *via* the expression of its truncated oligomerization domain resulted in the induction of CGP3, causing cell death. A bioinformatics analysis revealed homologous proteins mainly in actinomycetes, but, interestingly, also in several phage and prophage genomes. These data demonstrate the importance of XS proteins for the tolerance of viral DNA and indicate that this mechanism is exploited by both the host and the virus.

MATERIALS AND METHODS

Bacterial strains and growth conditions

The bacterial strains and plasmids used in this study are listed in Supplementary Table S1. *Corynebacterium glutamicum* ATCC 13032 was used as wild-type strain (24). *E. coli* DH5α was used as host for cloning procedures and cultivated in Lysogeny Broth (LB) medium or on agar plates at 37°C (25). For growth studies and fluorescence assays (e.g. preparation of cells for fluorescence microscopy), *C.glutamicum* cells were pre-cultivated in BHI (brain heart infusion, DifcoTM BHI, BD, Heidelberg, Germany) medium at 30°C for 6 h. This first preculture was used to inoculate an overnight culture in CGXII minimal medium (26) containing 2% (w/v) glucose and 30 $mg \cdot l^{-1}$ protocatechuat acid. The CGXII culture was finally used to inoculate the main culture in the same medium (CGXII with 2% (w/v) glucose) to a start OD_{600} of 1, unless specified otherwise. If necessary, 50 $\mu g \cdot ml^{-1}$ (*E. coli*) or 25 $\mu g \cdot ml^{-1}$ (*C. glutamicum*) kanamycin and/or 34 $\mu g \cdot ml^{-1}$ (*E. coli*) or 10 $\mu g \cdot ml^{-1}$ (*C. glutamicum*) chloramphenicol were added.

Recombinant DNA work

Plasmids and oligonucleotides used in this study are listed in Supplementary Table S2, respectively. Standard methods including PCR, DNA restriction and ligation, were performed according to established protocols (25). In some cases, Gibson assembly (27) was used for the constructions of plasmids. DNA sequencing and oligonucleotides synthesis were conducted by Eurofins MWG Operon (Ebersberg, Germany). The chromosomal integration of the Strep tagged *cgpS* gene variant was performed using the two-step homologous recombination method (28). The 500 bp up and downstream regions of *cgpS* were amplified using the oligonucleotides LF_cgpS_pK19_fw and LF_cgpS_rv and, accordingly, RF_cgpS_fw and RF_cgpS_pK19_rv. Amplification of the Strep-tagged *cgpS* gene was done by using the plasmid pAN6-*cgpS-strep* as template for the oligonucleotide pair cgpS_strep_fw and cgpS_strep_rv. The three resulting PCR products and the digested pK19*mobsacB* plasmid (with *Bam*HI, *Eco*RI) were assembled using Gibson assembly (27). Correct integration into the *cgpS* locus was confirmed by sequencing of the colony PCR product with the oligonucleotides Cgps_indel-fw and CgpS_indel_rv.

Cultivation in the BioLector System

Growth experiments were performed predominantly in the BioLector® microcultivation system of m2p-labs (Aachen, Germany) as described by (29). Cultivation was performed in 48-well FlowerPlates (m2p labs, Germany) at 30°C and a shaking frequency of 1200 rpm. The cells were cultivated in 750 µl of CGXII minimal media with 2% (w/v) glucose containing different additives (e.g. Isopropyl β-D-1-thiogalactopyranoside (IPTG), MMC, kanamycin), as indicated. Measurements were taken at 15-min intervals.

DNA affinity chromatography with the promoter region of *alpAC*

The promoter region of *alpAC* was amplified by PCR with the oligonucleotides PalpAC-Biotin-Tag-fw and PalpAC rv (product size 516 bp). To flag the amplified product further PCRs were performed but with the Biotin-Primer (MWG Eurofins, Ebersberg, Germany) and the PalpAC rv. At least 220 pmol of the biotinylated products were purified by size exclusion chromatography with the usage of an 8 ml sepharose s400-HR column from GE Healthcare (Freiburg, Germany). A total of 5 mg of the M-280 Streptavidin Dynabeads® (Invitrogen, Carlsbad, CA, USA) were washed twice with the binding and wash (BW) buffer (10 mM Tris-HCl pH 7.5, 2 M NaCl), subsequently suspended in BW buffer containing biotinylated products and incubated for 1 h at room temperature. To eliminate unbound DNA fragments the beads were washed three times with the BW buffer and finally suspended in the binding and storage (BS) buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 10% (v/v) glycerin, 0.01% (v/v) Triton-X-100, 100 mM NaCl, 1 mM DTT). A total fo 500 ml of cells were grown in CGXII minimal media with glucose as carbon source (as described in bacterial strains and growth conditions) to an OD_{600} of ∼5. After the cells were harvested by centrifugation (20 min, 5300*g* and washed once with phosphate buffered saline (PBS) buffer (137 mM NaCl, 2.7 mM KCl, 20 mM Na_2HPO_4, 1.8 mM KH_2PO_4), cell pellets were suspended in BS buffer supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF). Cell disruption was performed by five passages at 172 MPa through a French pressure cell (Heinemann, Schwaebisch Gmuend, Germany). The DNA binding reactions were set up with complete prepared crude extracts, the DNA-coupled beads and 500 µg of chromosomal DNA for 45 min at room temperature. After the binding reaction, beads were washed once with BS buffer, twice with BS buffer and 400 µg chromosomal DNA and, as a final washing step, again with BS buffer. The elution was fulfilled in two subsequent steps with BS buffer containing 2 M sodium chloride. After TCA precipitations (30) of the pooled elution fractions the samples were analyzed via

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (31). Identification of proteins was conducted by MALDI-ToF analysis as described in the section below.

Preparation of ChAP-Seq samples

Cells of the wild-type strain ATCC 13032 and the variant containing the Strep-tagged CgpS protein (WT::*cgpS-strep*) were first grown in BHI for 6 h and then 1 ml was used to inoculate minimal media cultures (CGXII with 2% (w/v) glucose). After cultivation overnight, these precultures were used to inoculate 500 ml of the same minimal medium, were grown to an OD_{600} 5 to 6, and finally harvested by centrifugation (10 min, 11 325*g* at 4°C). After washing the cells with CGXII medium without (w/o) MOPS, the cells were resuspended in 10 ml MOPS-free CGXII containing 1% (v/v) formaldehyde. The fixation was conducted by incubation at room temperature for 20 min. Subsequently, glycine was added to a final concentration of 125 mM and the cells were incubated for further 5 min at room temperature. Then, the cells were washed twice with buffer A (100 mM Tris-HCl, pH 8.0, 1 mM EDTA) and resuspended in 10 ml buffer A supplemented with *cOmplete* Protease Inhibitor (Roche, Basel, Switzerland) and 5 mg RNase A. Cell disruption was performed as described in the DNA affinity chromatography section (five passages through a French Press cell). The chromosomal DNA of the lysates were sheared by sonication 3 × 30 s with a Branson sonifier 250 (Heinemann, Schwaebisch Gmuend, Germany) using a pulse length of 40% and an intensity of one to give an average fragment size of 200–1500 bp as confirmed by agarose gel electrophoresis. Cell debris was first removed by centrifugation at 5300*g* for 20 min and then centrifuged for 1 h at 150 000*g* both steps at 4°C. The supernatant was used for protein–DNA purification according to the standard Strep-tag® purification protocol (see below, protein purification). The pooled elution fractions were incubated overnight at 65°C, followed by a treatment with proteinase K (final concentration 400 $mg \cdot ml^{-1}$) for 3 h at 55°C. Finally, the DNA of the samples was purified by phenol–chloroform extraction (32), precipitated with ethanol, washed with 70% (v/v) ethanol, dried and resuspended in 50–100 µl ddH_2O.

ChAP-Seq

The obtained DNA fragments of each sample (2 µg) were used for library preparation and indexing using the TruSeq DNA PCR-free sample preparation kit according to the manufacturer's instruction, yet omitting the DNA size selection steps (Illumina, Chesterford, UK). The resulting libraries were quantified using the KAPA library quant kit (Peqlab, Bonn, Germany) and normalized for pooling. Sequencing of pooled libraries was performed on a MiSeq (Illumina, San Diego, US) using paired-end sequencing with a read-length of 2 × 150 bases. Data analysis and base calling were accomplished with the Illumina instrument software and stored as fastq output files. The obtained sequencing data of each sample were imported into CLC Genomics Workbench (Version 7.5.1, Qiagen Aarhus A/S) for trimming and base quality filtering. The output was mapped to accession BX927147 as *C. glutamicum* reference genome (21). For peak detection the resulting mapping coverage of each sample was exported and imported into the in-house software Genome Data Viewer (unpublished). A peak was automatically annotated if the coverage of a region is above the 3-fold average of the averaged genome coverage. All peaks were inspected and confirmed manually.

qPCR

The relative amount of circular phage DNA was determined *via* quantitative PCR (qPCR). Therefore, *C. glutamicum* wild type cells containing empty pAN6 plasmid (control), pAN6-*cgpS* gene or pAN6-*N-cgpS* were grown in 48-well FlowerPlates containing CGXII minimal medium at 30°C and 900 rpm in a microtron (Infors-HT, Bottmingen, Switzerland). The overexpression of *cgpS* and the N-terminal part were induced with 150 µM IPTG (for control samples no IPTG was added). After 24 h, 750 µl of the cells were harvested and the DNA was extracted using the NucleoSpin microbial DNA Kit (Macherey Nagel, Dueren, Germany) and DNA concentration was quantified using a nanophotometer (Implen, München, Germany). Each sample contained 1 µg total DNA as a template. For the reaction an innuMIX qPCR MasterMix SyGreen (Analytic Jena, Jena, Germany) and a qTOWER 2.2 (Analytic Jena) was used. The reaction protocol was divided into two parts (i) polymerase chain reaction (PCR) ((a) 3 min preincubation at 95°C, (b) 5 s denaturation at 95°C, (c) 25 s elongation at 62°C, 40x repetition of step (b) to (c)) and a (ii) melting curve analysis ($\Delta T = 1°C/6$ s). The PCR product size using oligonucleotides belonging to the circular phage product is 150 bp (listed in Supplementary Table S2). As reference gene *ddh* was used with the oligonucleotides listed in Supplementary Table S2 resulting in a 150 bp product. For data analysis the qPCR software qPCR 3.1 (Analytik Jena) and the Livak method were used (33) to determine the $2^{-\Delta\Delta Ct}$ based on the measured C_T-values.

DNA microarrays

For a comparative transcriptome analysis of *C. glutamicum* ATCC 13032/pAN6 with cells carrying the pAN6-*N-cgpS*- (used for countersilencing) were cultivated in CGXII with 2% (w/v) glucose and 100 µM IPTG as described in bacterial strains and growth conditions. The preparation of labeled cDNA and DNA microarray analysis was performed as described previously (34). Array data were deposited in the GEO database (ncbi.nlm.nih.gov/geo) under accession number GSE80674.

Cultivation and perfusion in microfluidic device

For single-cell analysis an in-house developed microfluidic platform was used (22,35–37). Phase-contrast and fluorescence time-lapse imaging was performed at 6 min intervals. Medium was supplied continuously to ensure stable and constant environmental conditions. CGXII minimal medium with 2% (w/v) glucose and 25 $\mu g \cdot ml^{-1}$ kanamycin was infused at a rate of 300 $nl \cdot min^{-1}$ using a high-precision syringe pump (neMESYS, Cetoni GmbH, Korbussen, Germany). For the expression of the N-terminal part of CgpS

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

150 μM IPTG were added to the medium. A constant cultivation temperature of 30°C was ensured (PeCon GmbH, Erbach, Germany). The cells were cultivated for 16 h.

Fluorescence microscopy

The cultivations were done as described in bacterial strains and growth conditions. After 6 h of cultivation, 1–3 μl were pipetted on a microscope slide coated with a thin 1% (w/v) agarose layer that was based on tris-acetate buffer. To stain the DNA with the Hoechst Dye, 33 342 1 ml cells were harvested (5300*g*, 5 min), subsequently resuspended in PBS buffer containing 100 ng·ml^{-1} Hoechst 33342 and incubated at room temperature for 20 min. Images were taken on an AxioImager M2 (Zeiss, Oberkochen, Germany) equipped with a Zeiss AxioCam MRm camera. Fluorescence was monitored with the filter set 46 HE YFP for eYFP, 63 HE filter was used for mCherry fluorescence and Hoechst fluorescence was examined with the filter set 49. An EC Plan-Neofluar 100x/1.3 Oil Ph3 objective was used. Images were acquired and analyzed with the AxioVision 4.8 software (Carl Zeiss).

Protein purification

CgpS tagged C-terminal with a Strep-tag® was heterologously produced in *E. coli* BL21 (DE3). Cells were grown to an OD_{600} of 0.4 at 37°C. Upon induction with 50 μM IPTG the cultivation was continued at 16°C overnight. Cells were harvested by centrifugation at 5300*g* and 4°C for 10 min and resuspended in buffer B (250 mM NaCl, 50 mM Tris-HCl, pH 7.5). Cell disruption was performed by two passages through a French pressure cell at 172 MPa. Cell debris was removed by centrifugation at 20 min, 5300*g* and 4°C, followed by an ultracentrifugation (60 min, 229 000*g*, 4°C). The supernatant was applied to an equilibrated 1 ml Strep-Tactin®-Sepharose® (IBA, Göttingen, Germany) column. It was subsequently washed with 10 ml buffer B and the protein was eluted with 10 ml buffer B containing 1 mM d-desthiobiotin (Sigma Aldrich).

Electrophoretic mobility shift assays (EMSA)

EMSA studies of CgpS and selected DNA regions identified by ChAP-Seq were performed with selected regions (500 bp fragments, for oligo sequences see Supplementary Table S3). The corresponding regions were amplified by PCR and purified by using the PCR clean-up Kit of Macherey Nagel (Dueren, Germany). The promoter region of *gntK* was used as control fragment (560 bp). A total of 90 ng DNA per lane were incubated with different concentrations (1 μM and 2 μM) of purified CgpS protein for 20 min in EMSA buffer (250 mM Tris-HCl pH 7.5, 25 mM $MgCl_2$, 200 mM KCl, 25% (v/v) glycerol). Subsequently, samples were loaded onto a native 10% polyacrylamide gel (TBE-based, TBE (89 mM Tris base, 89 mM boric acid, 2 mM Na_2EDTA, loading dye: 0.01% (w/v) xylene cyanol dye, 0.01% (w/v) bromophenol blue dye, 20% (v/v) glycerol, 1x TBE). The DNA was stained with SYBR Green I (Sigma Aldrich, St. Louis, MO, USA).

Protein pull down and MALDI-TOF analysis

C. glutamicum cells containing the plasmids pAN6, pAN6-*cgpS*-strep or pAN6-N-*cgpS*-strep were cultivated as described in bacterial strains and growth conditions. The cultures were grown in 500 ml CGXII with 2% (w/v) glucose to an OD_{600} of 5 and subsequently induced with 150 μM IPTG for further 4 h. The cells were harvested (5300*g*, 20 min, 4°C), washed in buffer B (see protein purification) and disrupted as descripted in the DNA affinity chromatography section. Purification was performed as described in the section above. The eluted fractions were analyzed by SDS-PAGE (31) using a 4–20% Mini-PROTEAN® gradient gel (Bio Rad, Munich, Germany). The gels were stained with a Coomassie dye based RAPIDstain solution (G-Biosciences, St. Louis, MO, USA). MALDI-TOF-MS measurements were performed with an Ultraflex III TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) for the identification of the proteins as described (38).

Homology search

BLAST 'nr' database (ver. February 2015) was downloaded from NCBI (http://www.ncbi.nlm.nih.gov/). CgpS amino acid sequence was extracted from the GenBank file *Corynebacterium glutamicum* ATCC 13032, accession: NC_006958.1 and locus_tag: cg1966. A PSI-BLAST ((39)) search with CgpS sequence as the query was executed against the ncbi nr database. The e-value threshold was set to 0.005, the number of iteration was not limited and the search iteration was performed until it converged. A total of 5230 (1920 unique) homologous hits were achieved from which 618 could be allocated to a particular bacterial species or a phage. Sequence global identity was calculated by pairwise comparison between the CgpS sequence with all 618 PSI-BLAST hits using the Needleman–Wunsch algorithm (40) implemented in the EMBOSS package (41) needle.

Secondary structure prediction

The amino acid sequence of the CgpS protein and the sequences of the 618 homologous hits were used to predict the secondary structure by psipred (42). The visualization of the psipred output was done in R (43).

Statistics and visualization

All statistical analysis and data visualization from the bioinformatic section was performed in R (43).

RESULTS

A small nucleoid-associated protein encoded by a cryptic prophage element

To decipher the control of prophage induction and activation of cryptic elements in *C. glutamicum* ATCC 13032, we performed DNA affinity chromatography with the promoter of the early phage operon *alpAC* using the crude extract of log-phase cells grown in glucose minimal medium ((34), Figure 1A). SDS-Page analysis of the proteins bound

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

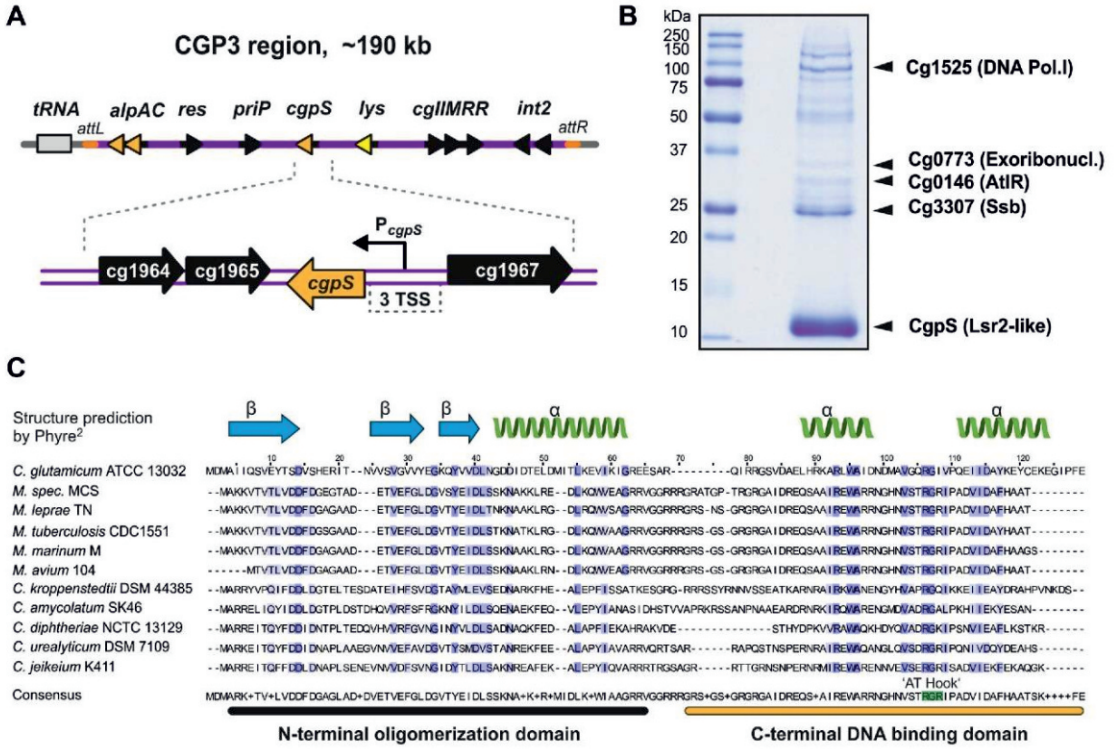


Figure 1. CgpS is a prophage-encoded nucleoid-associated protein in *C. glutamicum* ATCC 13032. (**A**) Genomic organization of the CGP3 prophage region containing the *cgpS* gene in *C. glutamicum* strain ATCC 13032. (**B**) DNA affinity chromatography was used to identify putative regulatory proteins bound to the early phage promoter of *alpAC* (34). Protein samples were separated by SDS-PAGE, and enriched proteins were identified using MALD-ToF analysis. The nucleoid-associated protein CgpS (Cg1966) was detected as a prominent band with a molecular mass of 13.4 kDa. (**C**) Amino acid sequences of CgpS, five related proteins of *Corynebacteria* and five Lsr2 homologs of *Mycobacteria* were used for a multiple sequence alignment conducted using the Clustal Omega platform (73). The predicted domain organization is depicted above the sequence alignment. The Blosum62 identity score is highlighted in blue and matches between 21–26% across the species. The secondary structure of CgpS was predicted with 99.7% confidence of 54 residues (46% of CgpS sequence) by Phyre^2 (74). The motif 'RGI,' which is similar to the AT hook motif 'RGR' of Lsr2 and H-NS (44,59), was identified between two predicted alpha helices.

to the *alpAC* promoter revealed a prominent band corresponding to the 13.4 kDa protein Cg1966 encoded within the CGP3 prophage region (Figure 1B). In particular, the C-terminal domain of Cg1966 shares significant sequence similarity with the nucleoid-associated protein Lsr2 of *Mycobacterium tuberculosis* (Supplementary Figure S1). This domain corresponds to the DNA binding domain of Lsr2 (IPR024412), which was previously found to bind AT-rich DNA *via* an AT-hook motif and functions as a silencer of xenogeneic DNA (44,45). Based on the data described in the following sections, we renamed Cg1966 as CgpS (*Corynebacterium glutamicum* prophage silencer). Secondary structure predictions of CgpS as well as of CgpS homologs suggest a significant structural similarity with Lsr2 and reveal the presence of an AT-hook-like motif 'RGI' between the two predicted C-terminal alpha helices (Figure 1C) (18,45).

CgpS functions as a silencer of CGP3 activity

To study the impact of *cgpS* expression on the activity of the CGP3 prophage, we overexpressed *cgpS* in a strain carrying a reporter construct (WT-P$_{lys}$-*eyfp*) indicative for the activation of CGP3 by the production of the yellow fluorescent protein eYFP under the control of a phage promoter (22). Upon induction with mitomycin C, the control strain carrying the empty plasmid displayed increased reporter activity. Consistent with our assumption, overexpression of *cgpS* reduced the reporter output to nearly the background level (Figure 2A).

To study the intracellular localization of CgpS in *C. glutamicum* cells, we C-terminally fused this protein to mCherry and analyzed its distribution *via* fluorescence microscopy. As shown by Hoechst staining, this NAP appeared associated with the nucleoid but formed distinct foci in the cell (Figure 2B and C). Remarkably, CgpS-mCherry foci co-localized with foci of an AlpA-eYFP fusion that was previously described as a CGP3 DNA adaptor protein binding to the *alpAC* promoter region (34) (Figure 2C). The

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

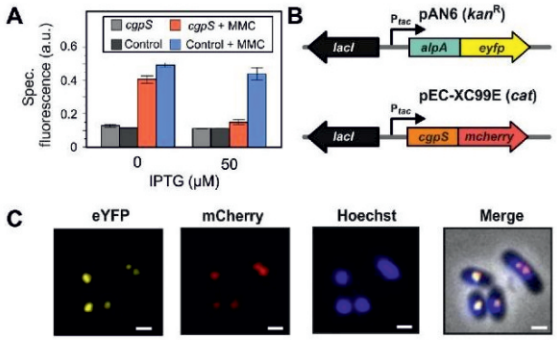


Figure 2. CgpS functions as a silencer of CGP3 prophage activity. (**A**) Silencing of CGP3 induction. The prophage reporter strain *C. glutamicum* ATCC 13032::P_{lys}-*eyfp* containing the *cgpS* overexpression plasmid pAN6-*cgpS-Strep* was cultivated in CGXII minimal medium in the presence or absence of IPTG (50 μM) and MMC (600 μM). The prophage reporter strain carrying the empty plasmid pAN6 served as a control. EYFP fluorescence was measured after 20 h of microplate cultivation. (**B** and **C**) CgpS is located in the nucleoid and displays colocalization with the phage adaptor protein AlpA (34). For co-localization studies, a C-terminal CgpS-mCherry fusion and a C-terminal fusion of the prophage adaptor protein AlpA to eYFP (pEC-XC99E) were analyzed. Both gene fusions were under control of the inducible *tac* promoter. DNA was stained with Hoechst Dye 33342. Fluorescence images were taken after 2.5 h of cultivation in CGXII with 150 μM IPTG. Scale bar, 2 μm.

functionality of this CgpS-mCherry fusion was confirmed by the counteraction of CGP3 activation upon addition of MMC (Supplementary Figure S2).

Genome-wide binding profile of CgpS

The data of the co-localization experiments suggest binding of CgpS to the CGP3 prophage region. In the following, the genome-wide binding profile was analyzed by combining affinity chromatography purification of crosslinked CgpS–DNA complexes followed by sequencing of associated DNA (ChAP-Seq). For this purpose, we replaced the native *cgpS* gene in the genome of ATCC 13032 with *cgpS-Strep* encoding a C-terminal Strep-tagged CgpS variant. This analysis revealed that CgpS associates with 1.5% of the ATCC 13032 genome and with ~20.5% of the cryptic CGP3 prophage region (Supplementary Figure S3). In total, 90 peaks were detected, 58 of which were within and 32 were located outside the CGP3 prophage (Figure 3A, Supplementary Table S4). The majority of the peak maxima were located within promoter regions (60%), but CgpS binding was also observed within genes (31%) or intergenic regions (9%) (Supplementary Figure S4B and C). To deduce a binding motif of CgpS, sequences of the 90 peaks (Supplementary Table S4) were extracted and analyzed using the MEME-ChIP software platform (46). A 21-bp long AT-rich motif was predicted, which was present in 87 of 90 sequences (Figure 3B). The occurrences of the found DNA binding sites were validated using a FIMO search (Find Individual Motif Occurrences, (47)) in the ATCC 13032 genome, which revealed significant matches (>75%) of the predicted and experimentally identified CgpS binding sites (Supplementary Figure S5). Remarkably, the %GC content of the 90 peak sequences is considerably lower than the average GC content of the ATCC 13032 strain, indicating the preferred binding of CgpS to AT-rich DNA (Figure 3C). Moreover, the GC contents of the CgpS bound regions within the prophage revealed no significant differences from that of the regions bound outside the prophage (Figure 3C).

Most of the identified CgpS targets were located within the CGP3 prophage and code for hypothetical proteins. The two strongest signals were found within transposase-encoding genes (cg1950-cg1951) and in the promoter region of *cgpS* itself, indicating a negative autoregulation similar to that of H-NS (48). Other potential target genes encode the actin-like protein and the corresponding adaptor protein (*alpAC*, cg1890 and cg1891 (34)), a resolvase (cg1929), a prophage primase (cg1959), a putative phage lysin (cg1974) and a phage integrase (cg2071), which are spread across the cryptic prophage element. In addition to regions within CGP3, CgpS target sites are located in the low GC island 1 (LCG1), in the cryptic phage element CGP1, or proximal to transposases encoding genes. Furthermore, promoter regions of genes coding for R-M systems (Pcg1028 and P*cglIM*, (Pcg1996)) are also bound by CgpS, which in several studies were shown to be transferred horizontally (49–52). A considerably high peak was observed for the promoter region of cg0150 that encodes a putative regulatory protein or toxin possessing a predicted fido domain (IPR003812).

The binding profile obtained by the ChAP-Seq analysis was validated by EMSAs (Supplementary Figure S6). For this purpose, CgpS was purified as a C-terminal Strep-tag fusion and incubated with DNA fragments covering selected putative CgpS binding sites as identified by ChAP-Seq (Figure 3D and Supplementary Figure S6). This *in vitro* approach confirmed the binding of CgpS for all selected target regions (including the promoters of cg0150, *alpAC* and *cgpS* itself) in comparison to the control fragment (*gntK* promoter) (Figure 3D). Overall, these data are consistent with CgpS acting as a xenogeneic silencer by targeting AT-rich DNA regions, several of which have likely been acquired by HGT.

Countersilencing of CgpS activity

Several independent efforts to inactivate the *cgpS* gene failed (data not shown), suggesting that *cgpS* represents an essential gene for *C. glutamicum* ATCC 13032. However, previous studies revealed that deletions of all three cryptic phage elements, including the *cgpS* gene, are possible and do not lead to a significant growth defect of the particular strain (53). In fact, trials to construct an in-frame deletion of *cgpS* resulted in the isolation of strains lacking large parts of the CGP3 prophage, indicating that the essentiality of *cgpS* is a consequence of the de-repression of toxic phage genes in the absence of CgpS.

For the conditional inactivation of CgpS, we adapted a countersilencing approach similar to the H-NST system described by Williamson and Free (54). This protein was reported as a truncated H-NS derivative that antagonizes H-NS function by interfering with the multimerization of H-NS. Co-purification assays with the N-terminal domain of CgpS confirmed the interaction of this truncated variant

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

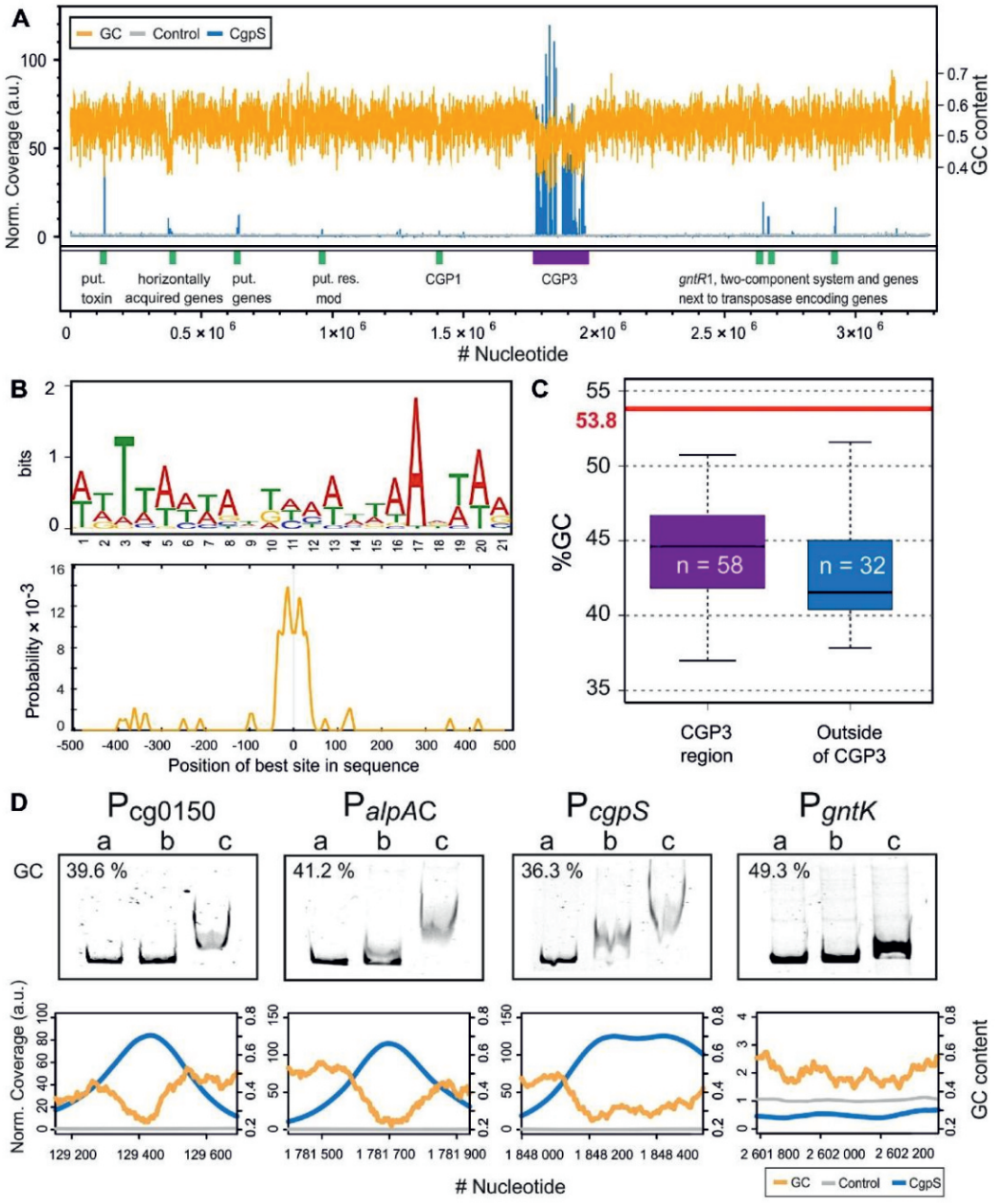


Figure 3. Genome-wide profiling of CgpS-binding using ChAP-Seq. (**A**) Genome wide binding profile of CgpS obtained by ChAP-Seq experiments. Enriched DNA regions purified in complex with CgpS (blue) or with the control sample (gray, empty vector control) were normalized to their mean and plotted against the ATCC 13032 genome. For calculation of the GC content (orange) and the coverages, a step size of 50 bp and a window size of 500 bp were used. Regions with high coverage are indicated by green (outside of CGP3 region) and purple boxes (CGP3 region). A total of 90 peaks were detected by applying a threshold of 3-fold of the mean coverage (SD of the control sample = ± 0.28 (a.u.)). (**B**) Sequences of the 90 peaks were used to derive a DNA binding motif using MEME-ChIP (75). A 21-bp long AT-rich motif was identified (E-value = 1.1×10^{-58}), and the highest probability was centered approximately ±38 bp around the maximum peak position. (**C**) A total of 58 detected signals were within and 32 were outside of the CGP3 region. GC contents of the peak sequences were calculated and compared to the average GC content of ATCC 13032 (21). (**D**) Electrophoretic mobility shift assays (EMSAs) were performed with promoter regions of the putative target genes. DNA fragments (around 500 bp) covering the promoter regions of cg0150, *alpAC*, *cgpS* and *gntK* (negative control) were incubated without (lane a) or with (lane b: 1 μM; c: 2 μM) purified CgpS protein. The corresponding ChAP-seq results of the particular regions are shown below the EMSA pictures (step size: 5 and windows size: 50).

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

with the full-length protein (Figure 4A). Based on previous data and the H-NST mechanism (Figure 4B), we constructed the pAN6-*N-cgpS* plasmid to overproduce a truncated variant of CgpS (amino acids 1–65) under the control of P_{tac}. Homology studies indicated that amino acids 1–65 cover the domain of CgpS required for the oligomerization of this NAP. Remarkably, production of the truncated CgpS-N domain in the wild-type strain resulted in a significant growth defect, whereas no impact on growth was observed in a strain lacking the CGP3 prophage (Figure 5A). This finding was supported by single-cell analysis of a strain containing a prophage reporter construct (P_{lys}-*eyfp*) (22) and the countersilencing construct pAN6-*N-cgpS*. Production of the N-terminal domain of CgpS led to a strong increase in fluorescence accompanied by growth arrest and a branched cell morphology (Figure 5C, Video S1 and S2). Quantitative real-time PCR revealed a 3-fold increase in the level of circular CGP3 DNA in comparison to uninduced cells, which is consistent with the induction of this cryptic prophage (Figure 5B) (20).

To monitor the impact of countersilencing CgpS activity on gene expression, we performed a comparative transcriptome analysis (Figure 5D, Supplementary Table S5). More than 194 genes were affected, 12 of which exhibited a reduced mRNA level (mRNA ratio ≤ 0.5, P-value < 0.05), and 182 genes were upregulated (mRNA ratio ≥ 2, P-value < 0.05). The majority of upregulated genes (148) were genes of the prophage CGP3. Additional genes that displayed an increased mRNA level were the ferritin gene (*ftn,* cg2782) and cg1517 of the CGP1 prophage (Supplementary Table S5), both of which were also identified as putative CgpS targets by ChAP-Seq. Together, these data demonstrate that CgpS is an essential NAP due to its function as a silencer of cryptic phage elements in *C. glutamicum*.

CgpS homologs are found in actinomycetes and their phages

Our data support a function for CgpS as a xenogeneic silencer that binds to AT-rich DNA similar to the Lsr2 of *M. tuberculosis* as well as the H-NS of *E. coli.* This is underlined by the fact that both proteins, Lsr2 and CgpS, are able to complement the phenotype of an *hns* mutant strain ((18), Supplementary Figure S7). These findings highlight the conserved mechanism of a highly diverse set of proteins.

In the following, we overexpressed the N-terminal oligomerization domains of CgpS orthologs from *Corynebacterium amycolatum* DSM 44737 (CORAM0001_2081) and *Corynebacterium diphtheria* DSM 44123 (CDC7B_2240) and the Lsr2 from *M. tuberculosis* H37R (Rv3597c; Lsr2) (Figure 6A and B). Whereas the production of the oligomerization domain strongly affected cellular growth in all cases (Figure 6A), only the N-terminal domain of the ortholog of *C. amycolatum* (DSM 44737) led to a significant induction of CGP3 (Figure 6B). No significant reporter output was observed with production of the truncated orthologs of *C. diphtheria* or *M. tuberculosis*, suggesting a high level of plasticity within this family of xenogeneic silencers (Figure 6B).

Furthermore, we used a bioinformatics approach to obtain a more general overview of the distribution of CgpS orthologous proteins. For this purpose, a PSI-BLAST (Position-Specific Iterated BLAST) search was performed on CgpS and resulted in 5230 hits, of which 1920 protein sequences were unique (threshold e-value ≤ 0.005). Of these, 98.3% were found in the domain of bacteria and 1.7% in phages, mostly belonging to the *Siphoviridae* (Figure 7A, Supplementary Table S6). Of 302 bacterial genomes containing prophage regions predicted by PhiSpy (55), 22 contain *cgpS* orthologs (Supplementary Table S6). The remaining 280 hits were found outside of any predicted prophage region. Moreover, secondary structure predictions were performed for 618 unique sequences, which were clearly assigned to bacterial or phage species, exhibiting high resemblances. The structural similarity suggests a common function, although the identity of the amino acid sequences is low (~23%) (Figure 7B, C and Supplementary Figure S10).

XS exclusion hypothesis

A recent bioinformatics study on the distribution of XS genes revealed that members of the same family can appear within a particular species but that members of different families are never found together (56). To test the proposed exclusion mechanism, we expressed the *hns* gene from *E. coli* MG1655 in a *C. glutamicum* ATCC 13032 strain containing the prophage reporter (::P_{lys}-*eyfp*). As expected, the overexpression of *hns* caused a severe growth defect, coinciding with a highly increased output of the prophage reporter (Figure 6C and D). The effect of *hns* overexpression was comparable to the countersilencing of CgpS activity with the production of a truncated CgpS variant (Figure 6E). When *hns* was expressed in a ΔCGP3 background the effect on growth was only moderate (Supplementary Figure S8). However, *hns* expression still negatively affected the growth of the CGP3 mutant strain which can likely be explained by unspecific binding and interference of H-NS at other genomic regions. These findings are in agreement with the hypothesis that different XS proteins interfere at AT-rich DNA regions, leading to a disruption of silencing complexes and thereby to an activation of foreign DNA elements. Nevertheless, in some cases the scenario is clearly more complex, as illustrated by the finding that the expression of *cgpS* in the *E. coli* wild-type strain was not able to counteract H-NS expression at the *bgl* operon (Supplementary Figure S9).

DISCUSSION

CgpS functions as a silencer of cryptic phage elements

In this study, we identified the prophage-encoded XS protein CgpS that inherits an essential role as a silencer of cryptic prophages in *C. glutamicum*. Genome-wide profiling of CgpS binding sites reveals an association of this protein to AT-rich DNA stretches primarily located within horizontally acquired genomic islands and shows a remarkable accumulation of binding sites within the large and cryptic CGP3 prophage. Countersilencing of CgpS activity by overproduction of its N-terminal oligomerization domain resulted in a strong increase in CGP3 activity leading to cell death. Furthermore, several CgpS binding sites were identified outside the CGP3 region, and the essentiality of the *cgpS* gene was attributed to the presence of the CGP3

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

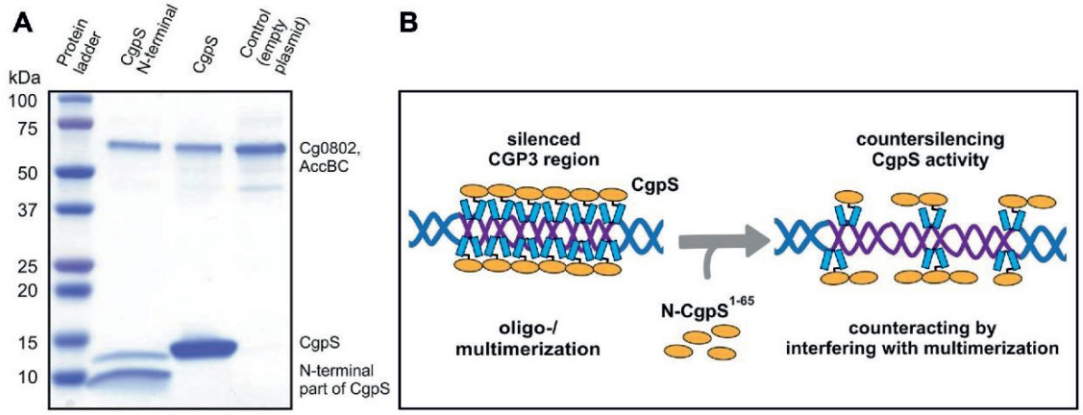


Figure 4. Principle of CgpS countersilencing. (**A**) Protein pulldown assays were conducted with *C. glutamicum* cells containing the plasmids pAN6-*cgpS-Strep* and pAN6-*N-cgpS-Strep*, which were used for the overexpression of Strep-tagged *cgpS* and its truncated variant. The pulldown of full-length CgpS by the truncated variant (aa 1–65) confirmed the N-terminal part of CgpS as oligo-/multimerization domain of CgpS proteins. (**B**) Model of CgpS silencing and countersilencing. The expression of genes depends on the accessibility of their particular promoter regions to the RNA polymerase. AT-rich regions such as CGP3 are bound by CgpS, likely resulting in an oligomerization of the CgpS protein (18,76,77), thereby interfering with the binding of RNA polymerase. The production of an N-terminal truncated CgpS variant interferes with the silencing ability of the native protein by binding to its N-terminal oligomerization domain while lacking the site for DNA binding.

prophage. This is consistent with the finding that the *cgpS* gene is located on the CGP3 island, suggesting that evolution favored a physical association between this XS and its main target.

Sequence analysis of CgpS revealed a low sequence identity (27%, Supplementary Figure S1) with the mycobacterial Lsr2 protein that was described in previous studies as an H-NS-like protein targeting AT-rich sequences in *M. tuberculosis* (18). Both XS proteins, Lsr2 and CgpS, complemented the *bgl*-based phenotype (57) of an *Escherichia coli* Δ*hns* strain, supporting the overall analogous functions of these XS proteins (Supplementary Figure S7) (18). Whereas both *lsr2* and *cgpS* are essential for viability in their native hosts, *E. coli hns* mutant strains are viable although exhibiting severe growth defects (58). *Salmonella Typhimurium* null mutants of *hns* are not viable unless mutations in *rpoS* (general stress response) or *phoP* (virulence gene regulator) counteract this deletion (12). Because the presence and diversity of phage elements contributes to major strain-specific differences within a bacterial species, our study illustrates that the essentiality of XS genes is highly dependent on the particular strain background. The *C. glutamicum* strain MB001, cured of all prophage regions as well as the *cgpS* gene located on prophage CGP3, displays wild-type-like growth behavior (53).

CgpS binds AT-rich xenogeneic DNA regions

Secondary structure predictions of CgpS-related proteins evince two α-helices flanking an 'RGI' motif (Figures 1C and 7C). This motif resembles the prokaryotic AT-hook motif 'Q/RGR' found in H-NS and Lsr2 and may also be responsible for the binding of AT-rich DNA as a general rule for XS functioning (44,59). A certain plasticity of the AT-hook motif is supported by experiments with AT-hook muteins of H-NS and Lsr2, showing that the exchange of a single arginine residue to an alanine reduces DNA binding but does not completely abolish it (59). Moreover, another member of the H-NS family, the Ler protein, has a hydrophobic amino acid ('VGR' motif) instead of an arginine at this position (60).

However, significant differences were observed for the number of target genes affected by the binding of the particular XS proteins. ChIP-on-Chip analysis revealed a direct influence of *S. Typhimurium* H-NS on the expression of more than 740 ORFs (12,61), and the binding of Lsr2 affected more than 800 regions within the *M. tuberculosis* genome and >900 in *Mycobacterium smegmatis* (45). ChAP-Seq profiling of CgpS binding, however, yielded only 90 potential target regions. Typical for XS function, an AT-rich DNA motif was derived from the ChAP-Seq results, which clusters at a high density within the CGP3 prophage region (Supplementary Figure S5). In general, promoter regions are more often bound by CgpS than genes or intergenic regions (Supplementary Figure S4), which is not surprising because promoter regions usually possess a higher AT content (62,63). CgpS targets outside the CGP3 region show a similar or lower GC content (Figure 3C) but less altered expression levels, and this may suggest the importance of motif density for XS function. Here, a variation of the AT-hook motif likely represents a mechanism to adjust the binding behavior of the XS protein to meet the needs of a particular host species.

In addition to CGP3 as a main CgpS target, further targets were identified which were also likely acquired by horizontal gene transfer, such as the LCG1 island, the cryptic prophage CGP1 (21), R-M systems, transposases and also regulatory proteins such as putative transcriptional regulators (Cg0725, Cg1340, Cg2426), the gluconate-responsive repressor GntR1 (Cg2783) (64) and an operon encoding the two-component system CgtSR6 (Cg3060) (Supplementary Table S4). Several previous studies reported similar tar-

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

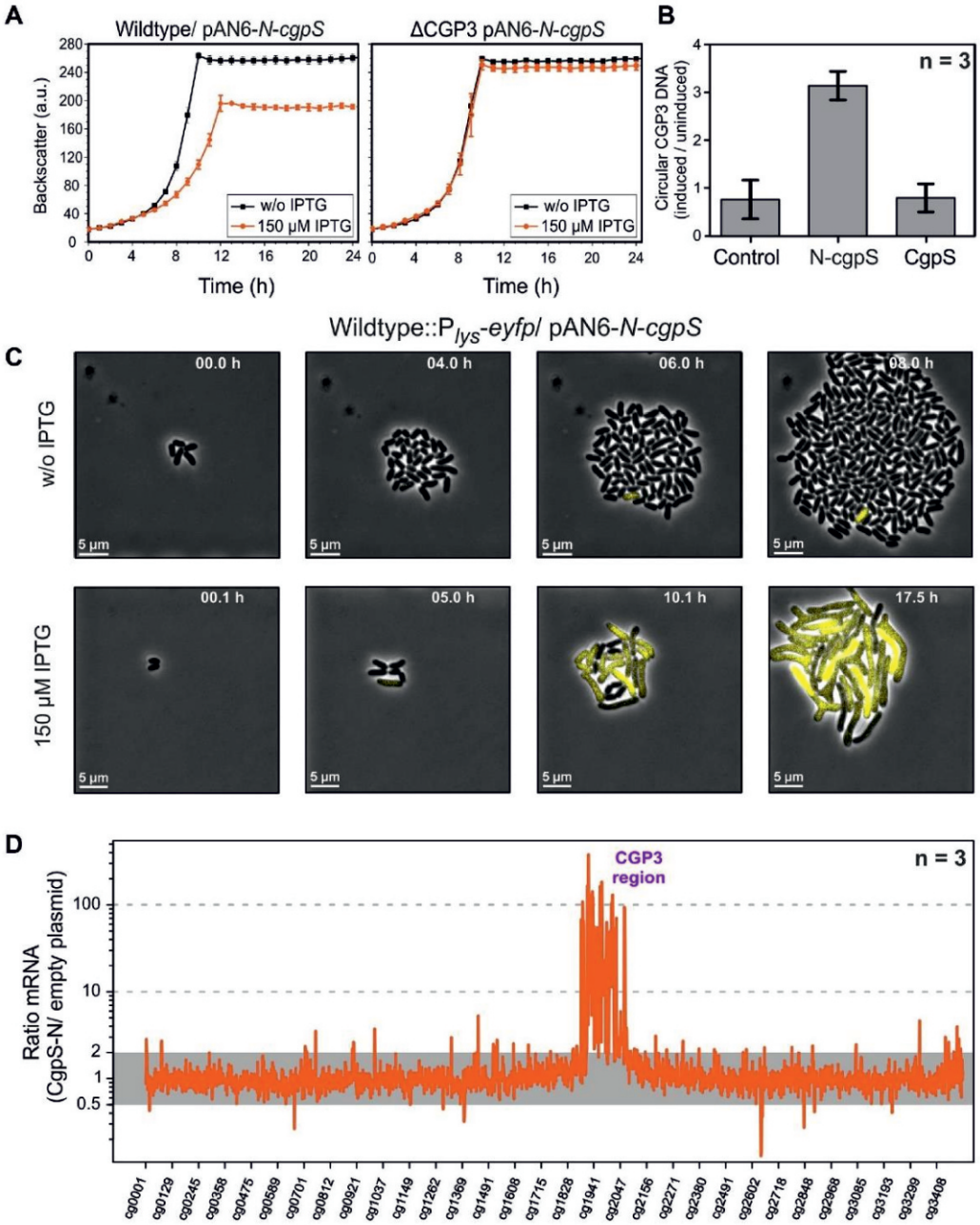


Figure 5. Countersilencing by overexpression of the N-terminal oligomerization domain of CgpS. (**A**) Growth studies of the wild-type ATCC 13032 and a strain lacking the CGP3 prophage (ΔCGP3), both carrying the pAN6-N-*cgpS* overexpression plasmid. Strains were grown in CGXII minimal medium with and without IPTG (150 µM). Data represent average values and standard deviations of three biological replicates. (**B**) Relative quantification of CGP3 excision using qPCR (20). The N-terminal domain of CgpS and the full-length protein were overproduced as described in (A). Samples for qPCR analysis were taken after 24 h. The relative amounts of circular phage DNA of induced and uninduced samples were compared. As a control wild-type cells with the empty plasmid were used. Data represent average values and standard deviations of three biological replicates. (**C**) Time-lapse fluorescence microscopy of the *C. glutamicum* prophage reporter strain (ATCC 13032::P*lys-eyfp*) carrying the pAN6-*N-cgpS*. Cells were grown in PDMS-based microfluidic chip devices under continuous supply of CGXII with 25 $\mu g \cdot ml^{-1}$ kanamycin and with or without 150 µM IPTG to induce the expression of the truncated CpgS variant (36) (300 $nl \cdot min^{-1}$) (Video S1 and S2). (**D**) Comparative transcriptome analysis of *C. glutamicum* ATCC 13032 containing the overexpression plasmid pAN6-*N-cgpS* and a strain containing the empty vector control was performed as described in the Materials and Methods section.

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

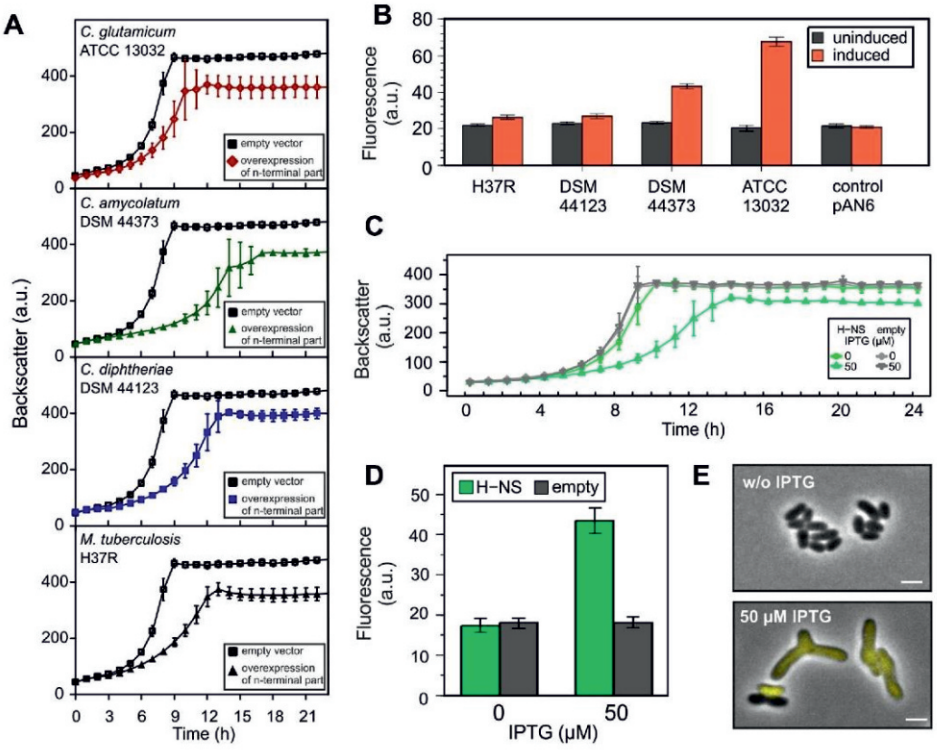


Figure 6. Impact of H-NS and Lsr2/Lsr2-like truncated variants on CGP3 prophage induction. (**A** and **B**) Impact of truncated CgpS orthologs on the growth and CGP3 prophage activity using the reporter strain ATCC 13032::P*lys*-*eyfp*. Shown is the cultivation in microtiter plates in CGXII minimal media with 25 $\mu g \cdot ml^{-1}$ kanamycin and 150 µM IPTG. As a control the reporter strain containing the empty plasmid was used. In (**B**), the fluorescence output after 20 h is shown in comparison to the uninduced samples. (**C** and **D**) Growth experiments were performed with WT::P*lys*-*eyfp* cells carrying *hns* on the overexpression plasmid pAN6 under the conditions described in (**A**). Expression of *hns* was induced with 50 µM IPTG. Fluorescence output of the prophage reporter after 20 h is shown in (**D**). (**E**) After 24 h, fluorescence images were taken of cells (of C and D) placed on agar pads. Scale bar is 2 µm. Data represent average values and standard deviations of three biological replicates.

get genes or regions for H-NS, Lsr2 and MvaT, demonstrating the convergent evolution of XS in bacterial species (12,47,61,65).

Overall, more that 80% of CgpS-bound regions also exhibited a more than 2-fold altered expression level under countersilencing conditions (Figure 5D) confirming the postulated silencing effect of CgpS. Several potential targets outside of the CGP3 region, however, showed only a moderate impact on the expression level suggesting a more complex regulatory scheme at the corresponding promoter regions. Therefore, the role of CgpS for the control of these potential targets, including, e.g. the *gntR1* gene or the *cgtSR6* operon, remains to be elucidated in further studies.

How to overcome CgpS silencing?

Several different mechanisms were described to counteract H-NS-mediated silencing, including structural interference with H-NS-bound nucleoids by transcription factors, temperature or osmolarity effects, and the binding of alternative sigma factors or other NAPs preventing multimerization of the XS protein (11,66,67). To interfere with CgpS XS activity, we produced a truncated part of the native protein covering the N-terminal oligomerization domain of CgpS (Figure 5). This overcomes the problem of *cgpS* being essential in the presence of CGP3 and was inspired by the study of Williamson and Free, who described the antagonistic function of a truncated H-NS variant found in an enteropathogenic *E. coli* strain (54). As expected, production of the N-terminal CgpS domain resulted in strong activation of CGP3, leading to cell death.

In recent studies we described the spontaneous induction of the CGP3 prophage occuring in the absence of an external trigger (20,22,23). Single-cell analysis demonstrated that a considerable fraction of this SPI is preceded by an activation of the SOS response, which is likely the result of spontaneous DNA damage during replication (68,69). However, these studies also highlighted a certain (>30%) fraction of SOS-independent SPI, suggesting that other factors influence this common phenomenon of bacterial populations (5). The present study shows the sensitive reaction of *C. glutamicum* cells to the downregulation of CgpS activity (Video S2). It is therefore interesting to determine whether cells can adjust the level of XS proteins to manipulate the frequency of SPI according to their particular requirements.

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

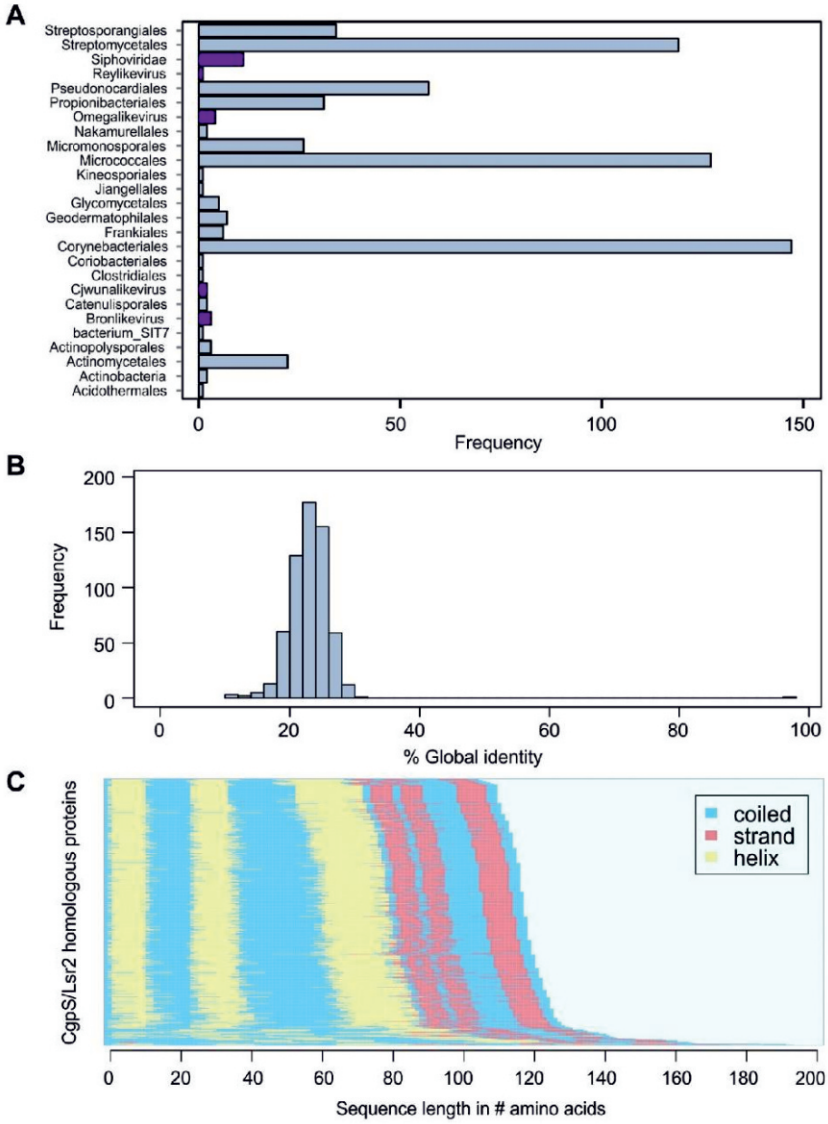


Figure 7. Orthologous sequences of the CgpS protein in actinomycetes and their phages. (**A**) Bar chart depicting the frequency of orthologous hits (n = 618) as predicted by PSI-BLAST (39) (e-value $\leq$ 0.005) across several orders of the phylum Actinobacteria and phages as annotated in the NCBI database (http://www.ncbi.nlm.nih.gov/). Orthologous found in phages are highlighted in purple. (**B**) Histogram of pairwise global identities between the CgpS amino acid sequence and its orthologous counterparts. The distribution reveals an overall low similarity (mean x = 23.07 and standard deviation σ = 4.05) to the orthologous sequences. Global identity was calculated using the Needleman–Wunsch (40) algorithm from the EMBOSS package (41). (**C**) Secondary structure prediction calculated by psipred (42) shows conserved protein structure for CgpS and the orthologous amino acid sequences. The corresponding secondary structure of each sequence was ordered in the direction from C-Terminus to N-terminus. Predicted coiled structures are shown in blue, strand regions in red and helices are colored in yellow.

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

CgpS homologs in phage genomes

Sequence analysis revealed the presence of CgpS/Lsr2 homologs in phage and prophage genomes displaying a low sequence identity but highly conserved secondary structure prediction (Figure 7). This finding is not surprising because bacterial evolution has been shaped by a tight interaction with bacteriophages. For the integration of viral DNA into the host genome, both the bacterium and phage benefit from tolerance and a smooth integration into the host genetic circuitry. Because the activation of silent prophages or mobile elements often causes serious detrimental effects to host cells (11,70,71), the stringent control of xenogeneic elements is required.

Several examples of XS proteins involved in the control of mobile elements or phages have been described in the recent literature, including H-NS of *S. Typhimurium* (12), Rok from *B. subtilis* (19) and MvaT from *P. aeruginosa* (72). Their corresponding genes, however, are all located on the host chromosome and are characterized as a type of immunity system protecting hosts against foreign DNA (11,66). A PSI-BLAST search of CgpS-related proteins revealed that the majority (>98% of all hits, >92% of prophage containing strains) are found in bacterial genomes (Supplementary Table S6). However, several examples located in phages or prophage regions were identified. The functions of these phage-encoded XS-like proteins remain to be studied, but their presence suggests the following: (i) like CgpS, they may be required to secure tolerance of their carrier DNA within the respective host; (ii) they may, however, also function as antagonistic proteins, interfering with the host XS protein similar to the situation described for H-NST (54); or (iii) they may interfere with the function of another class of XS proteins. This hypothesis is based on the exclusion theory suggested by Perez-Rueda and Ibarra, who postulated that XS from different families do not appear in the same bacterial organism (56). Consistent with this bioinformatics study, our data show that the expression of *E. coli hns* results in strong activation of the cryptic prophage CGP3 and consequently cell death. The finding that expression of the *C. glutamicum cgpS* gene in *E. coli* MG1655 does not counteract H-NS-mediated silencing at the *bgl* operon shows, however, that the scenario is more complex and strongly depends on the particular strain and its regulatory equipment. However, our data on prophage activation in *C. glutamicum* provide evidence for an interference of analogous XS proteins at AT-rich DNA regions. Here, likely the incompatibility of the oligomerization domains inhibits the formation of XS multimeric structures required for silencing. Considering the presence of XS encoding genes in phage and prophage genomes, this principle is likely to be harnessed by any phage predator by encoding an interfering XS.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

The authors thank Karin Schnetz (University of Cologne) for helpful advice and for providing us with the *E. coli* Δ*hns* mutant strain.

FUNDING

Deutsche Forschungsgemeinschaft priority program SPP1617 [FR 2759/2-2 and KO 4537/1-2]; Helmholtz Association [VH-NG-716]. Funding for open access charge: Helmholtz Association [VH-NG-716].
Conflict of interest statement. None declared.

REFERENCES

1. Juhas,M., van der Meer,J.R., Gaillard,M., Harding,R.M., Hood,D.W. and Crook,D.W. (2009) Genomic islands: Tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.*, **33**, 376–393.
2. Canchaya,C., Proux,C., Fournous,G., Bruttin,A. and Brussow,H. (2003) Prophage genomics. *Microbiol. Mol. Biol. Rev.*, **67**, 238–276.
3. Casjens,S. (2003) Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.*, **49**, 277–300.
4. Soucy,S.M., Huang,J. and Gogarten,J.P. (2015) Horizontal gene transfer: Building the web of life. *Nat. Rev. Genet.*, **16**, 472–482.
5. Nanda,A.M., Thormann,K. and Frunzke,J. (2015) Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *J. Bacteriol.*, **197**, 410–419.
6. Wang,X., Kim,Y., Ma,Q., Hong,S.H., Pokusaeva,K., Sturino,J.M. and Wood,T.K. (2010) Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.*, **1**, 147.
7. d'Hérelle,F. (1917) Sur un microbe invisible antagoniste des bacilles dysentérique. *Acad. Sci. Paris*, **165**, 373–375.
8. Diaz Ricci,J.C. and Hernandez,M.E. (2000) Plasmid effects on *Escherichia coli* metabolism. *Crit. Rev. Biotechnol.*, **20**, 79–108.
9. Labrie,S.J., Samson,J.E. and Moineau,S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, **8**, 317–327.
10. Samson,J.E., Magadan,A.H., Sabri,M. and Moineau,S. (2013) Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.*, **11**, 675–687.
11. Navarre,W.W., McClelland,M., Libby,S.J. and Fang,F.C. (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev.*, **21**, 1456–1471.
12. Navarre,W.W., Porwollik,S., Wang,Y., McClelland,M., Rosen,H., Libby,S.J. and Fang,F.C. (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science*, **313**, 236–238.
13. Will,W.R., Navarre,W.W. and Fang,F.C. (2015) Integrated circuits: how transcriptional silencing and counter-silencing facilitate bacterial evolution. *Curr. Opin. Microbiol.*, **23**, 8–13.
14. Langille,M.G., Hsiao,W.W. and Brinkman,F.S. (2010) Detecting genomic islands using bioinformatics approaches. *Nat. Rev. Microbiol.*, **8**, 373–382.
15. Dame,R.T., Luijsterburg,M.S., Krin,E., Bertin,P.N., Wagner,R. and Wuite,G.J. (2005) DNA bridging: A property shared among H-NS-like proteins. *J. Bacteriol.*, **187**, 1845–1848.
16. Ding,P., McFarland,K.A., Jin,S., Tong,G., Duan,B., Yang,A., Hughes,T.R., Liu,J., Dove,S.L., Navarre,W.W. *et al.* (2015) A novel AT-rich DNA recognition mechanism for bacterial xenogeneic silencer MvaT. *PLoS Pathog.*, **11**, e1004967.
17. Oshima,T., Ishikawa,S., Kurokawa,K., Aiba,H. and Ogasawara,N. (2006) *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res.*, **13**, 141–153.
18. Gordon,B.R., Imperial,R., Wang,L., Navarre,W.W. and Liu,J. (2008) Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J. Bacteriol.*, **190**, 7052–7059.
19. Smits,W.K. and Grossman,A.D. (2010) The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in *Bacillus subtilis*. *PLoS Genet.*, **6**, e1001207.
20. Frunzke,J., Bramkamp,M., Schweitzer,J.E. and Bott,M. (2008) Population Heterogeneity in *Corynebacterium glutamicum* ATCC 13032 caused by prophage CGP3. *J. Bacteriol.*, **190**, 5111–5119.
21. Kalinowski,J., Bathe,B., Bartels,D., Bischoff,N., Bott,M., Burkovski,A., Dusch,N., Eggeling,L., Eikmanns,B.J., Gaigalat,L. *et al.* (2003) The complete *Corynebacterium glutamicum* ATCC 13032

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.*, **104**, 5–25.
22. Helfrich,S., Pfeifer,E., Krämer,C., Sachs,C.C., Wiechert,W., Kohlheyer,D., Nöh,K. and Frunzke,J. (2015) Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations. *Mol. Microbiol.*, **98**, 636–650.
23. Nanda,A.M., Heyer,A., Krämer,C., Grünberger,A., Kohlheyer,D. and Frunzke,J. (2014) Analysis of SOS-induced spontaneous prophage induction in *Corynebacterium glutamicum* at the single-cell level. *J. Bacteriol.*, **196**, 180–188.
24. Kinoshita,S., Udaka,S. and Shimono,M. (1957) Studies on the amino acid fermentation - Part I. Production of L-glutamic acid by various microorganisms. *J. Gen. Appl. Microbiol.*, **50**, 331–343.
25. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
26. Keilhauer,C., Eggeling,L. and Sahm,H. (1993) Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J. Bacteriol.*, **175**, 5595–5603.
27. Gibson,D.G., Young,L., Chuang,R.Y., Venter,J.C., Hutchison,C.A. and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.
28. Niebisch,A. and Bott,M. (2001) Molecular analysis of the cytochrome bc1-aa3 branch of the *Corynebacterium glutamicum* respiratory chain containing an unusual diheme cytochrome c1. *Arch. Microbiol.*, **175**, 282–294.
29. Heyer,A., Gätgens,C., Hentschel,E., Kalinowski,J., Bott,M. and Frunzke,J. (2012) The two-component system ChrSA is crucial for haem tolerance and interferes with HrrSA in haem-dependent gene regulation in *Corynebacterium glutamicum*. *Microbiology*, **158**, 3020–3031.
30. Kim,S.C., Chen,Y., Mirza,S., Xu,Y., Lee,J., Liu,P. and Zhao,Y. (2006) A clean, more efficient method for in-solution digestion of protein mixtures without detergent or urea. *J. Proteome Res.*, **5**, 3446–3452.
31. Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
32. Evans,G.A. Molecular cloning: A laboratory manual. Second edition. Volumes 1, 2, and 3. Current protocols in molecular biology. Volumes 1 and 2. *Cell*, **61**, 17–18.
33. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.
34. Donovan,C., Heyer,A., Pfeifer,E., Polen,T., Wittmann,A., Krämer,R., Frunzke,J. and Bramkamp,M. (2015) A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria. *Nucleic Acids Res.*, **43**, 5002–5016.
35. Grünberger,A., Paczia,N., Probst,C., Schendzielorz,G., Eggeling,L., Noack,S., Wiechert,W. and Kohlheyer,D. (2012) A disposable picolitre bioreactor for cultivation and investigation of industrially relevant bacteria on the single cell level. *Lab Chip*, **12**, 2060–2068.
36. Grünberger,A., Probst,C., Helfrich,S., Nanda,A., Stute,B., Wiechert,W., von Lieres,E., Nöh,K., Frunzke,J. and Kohlheyer,D. (2015) Spatiotemporal microbial single-cell analysis using a high-throughput microfluidics cultivation platform. *Cytometry*, **87A**, 1101–1115.
37. Grünberger,A., van Ooyen,J., Paczia,N., Rohe,P., Schiendzielorz,G., Eggeling,L., Wiechert,W., Kohlheyer,D. and Noack,S. (2013) Beyond growth rate 0.6: *Corynebacterium glutamicum* cultivated in highly diluted environments. *Biotechnol. Bioeng.*, **110**, 220–228.
38. Bussmann,M., Baumgart,M. and Bott,M. (2010) RosR (Cg1324), a hydrogen peroxide-sensitive MarR-type transcriptional regulator of *Corynebacterium glutamicum*. *J. Biol. Chem.*, **285**, 29305–29318.
39. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J.H., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
40. Needleman,S.B. and Wunsch,C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, **48**, 443–453.
41. Rice,P., Longden,I. and Bleasby,A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.*, **16**, 276–277.
42. Jones,D.T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.*, **292**, 195–202.
43. R Development Core Team (2016) R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, http://www.R-project.org.
44. Ali,S.S., Xia,B., Liu,J. and Navarre,W.W. (2012) Silencing of foreign DNA in bacteria. *Curr. Opin. Microbiol.*, **15**, 175–181.
45. Gordon,B.R., Li,Y., Wang,L., Sintsova,A., van Bakel,H., Tian,S., Navarre,W.W., Xia,B. and Liu,J. (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 5154–5159.
46. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Ren,J., Li,W.W. and Noble,W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.
47. Grant,C.E., Bailey,T.L. and Noble,W.S. (2011) FIMO: Scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017–1018.
48. Dersch,P., Schmidt,K. and Bremer,E. (1993) Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol. Microbiol.*, **8**, 875–889.
49. Khan,F., Furuta,Y., Kawai,M., Kaminska,K.H., Ishikawa,K., Bujnicki,J.M. and Kobayashi,I. (2010) A putative mobile genetic element carrying a novel type IIF restriction-modification system (PluTI). *Nucleic Acids Res.*, **38**, 3019–3030.
50. Nobusato,A., Uchiyama,I. and Kobayashi,I. (2000) Diversity of restriction-modification gene homologues in *Helicobacter pylori*. *Gene*, **259**, 89–98.
51. Jeltsch,A. and Pingoud,A. (1996) Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. *J. Mol. Evol.*, **42**, 91–96.
52. Furuta,Y. and Kobayashi,I. (2014) *Bacterial Integrative Mobile Genetic Elements*. Austin.
53. Baumgart,M., Unthan,S., Rückert,C., Sivalingam,J., Grünberger,A., Kalinowski,J., Bott,M., Noack,S. and Frunzke,J. (2013) Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Appl. Environ. Microbiol.*, **79**, 6006–6015.
54. Williamson,H.S. and Free,A. (2005) A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS antagonist. *Mol. Microbiol.*, **55**, 808–827.
55. Akhter,S., Aziz,R.K. and Edwards,R.A. (2012) PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res.*, **40**, e126.
56. Perez-Rueda,E. and Ibarra,J.A. (2015) Distribution of putative xenogeneic silencers in prokaryote genomes. *Comput. Biol. Chem.*, **58**, 167–172.
57. Dole,S., Kühn,S. and Schnetz,K. (2002) Post-transcriptional enhancement of *Escherichia coli bgl* operon silencing by limitation of BglG-mediated antitermination at low transcription rates. *Mol. Microbiol.*, **43**, 217–226.
58. Yamada,H., Yoshida,T., Tanaka,K., Sasakawa,C. and Mizuno,T. (1991) Molecular analysis of the *Escherichia coli hns* gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. *Mol. Gen. Genet.*, **230**, 332–336.
59. Gordon,B.R., Li,Y., Cote,A., Weirauch,M.T., Ding,P., Hughes,T.R., Navarre,W.W., Xia,B. and Liu,J. (2011) Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10690–10695.
60. Cordeiro,T.N., Schmidt,H., Madrid,C., Juarez,A., Bernado,P., Griesinger,C., Garcia,J. and Pons,M. (2011) Indirect DNA readout by an H-NS related protein: structure of the DNA complex of the C-terminal domain of Ler. *PLoS Pathog.*, **7**, e1002380.
61. Lucchini,S., Rowley,G., Goldberg,M.D., Hurd,D., Harrison,M. and Hinton,J.C. (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.*, **2**, e81.
62. Pedersen,A.G., Jensen,L.J., Brunak,S., Staerfeldt,H.H. and Ussery,D.W. (2000) A DNA structural atlas for *Escherichia coli*. *J. Mol. Biol.*, **299**, 907–930.
63. Ussery,D.W., Tindbaek,N. and Hallin,P.F. (2004) Genome update: promoter profiles. *Microbiology*, **150**, 2791–2793.
64. Frunzke,J., Engels,V., Hasenbein,S., Gätgens,C. and Bott,M. (2008) Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

transcriptional regulators, GntR1 and GntR2. *Mol. Microbiol.*, **67**, 305–322.
65. Castang,S., McManus,H.R., Turner,K.H. and Dove,S.L. (2008) H-NS family members function coordinately in an opportunistic pathogen. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 18947–18952.
66. Navarre,W.W. (2009) In: Dame,RT and Dorman,CJ (eds). *Bacterial Chromatin*. Springer, Netherlands.
67. Stoebel,D.M., Free,A. and Dorman,C.J. (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology*, **154**, 2533–2545.
68. Cox,M.M., Goodman,M.F., Kreuzer,K.N., Sherratt,D.J., Sandler,S.J. and Marians,K.J. (2000) The importance of repairing stalled replication forks. *Nature*, **404**, 37–41.
69. Pennington,J.M. and Rosenberg,S.M. (2007) Spontaneous DNA breakage in single living *Escherichia coli* cells. *Nat. Genet.*, **39**, 797–802.
70. Buckling,A. and Rainey,P.B. (2002) Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. Biol. Sci.*, **269**, 931–936.
71. Lee,S.W. and Edlin,G. (1985) Expression of tetracycline resistance in pBR322 derivatives reduces the reproductive fitness of plasmid-containing *Escherichia coli*. *Gene*, **39**, 173–180.
72. Li,C., Wally,H., Miller,S.J. and Lu,C.D. (2009) The multifaceted proteins MvaT and MvaU, members of the H-NS family, control arginine metabolism, pyocyanin synthesis, and prophage activation in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.*, **191**, 6211–6218.
73. Sievers,F., Wilm,A., Dineen,D., Gibson,T.J., Karplus,K., Li,W., Lopez,R., McWilliam,H., Remmert,M., Soding,J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**, 539.
74. Kelley,L.A., Mezulis,S., Yates,C.M., Wass,M.N. and Sternberg,M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
75. Ma,W., Noble,W.S. and Bailey,T.L. (2014) Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat. Protoc.*, **9**, 1428–1450.
76. Cam,E.L., Culard,F., Larquet,E., Delain,E. and Cognet,J.A. (1999) DNA bending induced by the archaebacterial histone-like protein MC1. *J. Mol. Biol.*, **285**, 1011–1021.
77. Chen,J.M., Ren,H., Shaw,J.E., Wang,Y.J., Li,M., Leung,A.S., Tran,V., Berbenetz,N.M., Kocincova,D., Yip,C.M. *et al.* (2008) Lsr2 of *Mycobacterium tuberculosis* is a DNA-bridging protein. *Nucleic Acids Res.*, **36**, 2123–2135.

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on August 4, 2016

3.3 Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium.

Pfeifer E., Gätgens C., Polen T.*, and Frunzke J.*

IBG-1: Biotechnology, Forschungszentrum Jülich, Jülich, Germany

*Corresponding author

Current state: *Under Review*

Name of Journal: *Metabolic Engineering*

Impact factor: *8.201*

Author contributions

Own contribution to the work: 90 %

Contribution to project design and manuscript

	Name	Contribution
Project planning	**Pfeifer, E.**	85%
	Frunzke, J.	15%
Writing	**Pfeifer, E.**	85%
	Frunzke, J.	15%

Contribution to experiments, their evaluation and illustration

	Name	Experimental work	Evaluation	Processing & Preparation of figure/ table
Figure 1	**Pfeifer, E.**	100%	100%	100%
Figure 2	**Pfeifer, E.**	95%	100%	100%
	Gätgens, C	5%		
Figure 3	**Pfeifer, E.**	100%	100%	100%
Figure 4	**Pfeifer, E.**	100%	85%	100%
	Frunzke, J.	-	10%	-
	Polen, T.	-	5%	-
Figure 5	**Pfeifer, E.**	70%	100%	100%
	Gätgens, C	30%	-	-
Figure 6	**Pfeifer, E.**	80%	100%	100%
	Gätgens, C	20%	-	-
Table 1	**Pfeifer, E.**	-	70%	100%
	Frunzke, J.	-	30%	-
Table 2	**Pfeifer, E.**	-	100%	100%
Table 3	**Pfeifer, E.**	40%	100%	100%
	Gätgens, C	60%	-	-

	Name	Experimental work	Evaluation	Processing & Preparation of figure/ table
Table S1	**Pfeifer, E.**	-	100%	100%
Table S2	**Pfeifer, E.**	-	100%	100%
Table S3	**Pfeifer, E.**	-	100%	100%
Table S4	**Pfeifer, E.**	-	95%	100%
	Polen, T.	-	5%	-
Table S5	**Pfeifer, E.**	-	100%	100%
Table S6	**Pfeifer, E.**	-	70%	100%
	Polen, T.	-	30%	-
Table S7	**Pfeifer, E.**	-	80%	100%
	Polen, T.	-	20%	-
Figure S1	**Pfeifer, E.**	100%	100%	100%
Figure S2	**Pfeifer, E.**	100%	100%	100%
Figure S3	**Pfeifer, E.**	50%	100%	100%
	Gätgens, C.	50%	-	-
Figure S4	**Pfeifer, E.**	-	100%	100%
Figure S5	**Pfeifer, E.**	-	100%	100%
Figure S6	**Pfeifer, E.**	10%	100%	100%
	Gätgens, C.	90%	-	-
Figure S7	**Pfeifer, E.**	15%	100%	100%
	Gätgens, C.	85%	-	-

Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium.

Eugen Pfeifer, Cornelia Gätgens, Tino Polen*, and Julia Frunzke*

*Corresponding author

Institute of Bio- und Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Adaptive laboratory evolution (ALE) experiments persuaded to be promising key techniques in applied and fundamental sciences enabling the improvement of cellular properties and the identification of non-intuitive beneficial mutations. In this work, we performed a comparative ALE experiment of the important biotechnological platform strain *Corynebacterium glutamicum* ATCC 13032 and its prophage-free variant MB001 towards improved growth rates on glucose minimal medium. Both strains (each six independent cell lines) displayed a comparable adaptation behavior and no significant differences in genomic rearrangements and mutation frequencies. Remarkable, a significant fitness leap by about 20% was observed for both strains already after 100 generations. In a second ALE experiment the reproducibility of this prominent fitness leap was confirmed. Isolated top clones (UBw and UBm) showed an about 26% increased growth rate on glucose minimal medium. Genome sequencing of evolved populations as well as isolated clones resulted in the identification of key mutations in *pyk* (pyruvate kinase), *fruK* (1-phosphofructokinase) and *corA* encoding a Mg^{2+} importer. The reintegration of selected *pyk* and *fruK* mutations resulted in an increased glucose uptake and *ptsG* expression causative for the accelerated growth on glucose minimal medium. In contrast, the introduction of single *corA* mutations did not improve growth and thus likely represent an example for epistatic interactions. Overall, this study resulted in the identification of causative key mutations improving the growth of *C. glutamicum* on glucose. These identified mutational hot spots as well as the two evolved top strains, UBw and UBm, represent promising targets for future metabolic engineering approaches.

In recent years, adaptive laboratory evolution experiments (ALE) in combination with next-generation sequencing (NGS) became a key approach to study microbial adaptation in fundamental as well as in applied research (1-4). Especially in the field of metabolic engineering, researchers took advantage of the fast adaptation of microbes towards changing environments (4, 5). Here, ALE approaches represent a powerful complementary strategy to rational strain engineering to improve growth, product tolerance or stress resistance (2, 4-9). Now, even synthetic ALE scenarios are emerging where synthetic regulatory circuits are implemented to impose an artificial selection pressure on a particular phenotypic trait (e.g. small molecule production) (10-12). In the course of ALE experiments, spontaneous mutations and genomic rearrangements bearing a fitness advantage under the particular selection pressure establish within the population. However, the overall setup of the ALE experiment has a remarkable impact on the outcome. Important parameters include for example the mode of propagation (batch *versus* continuous culture), the passage size and the growth phase under which cells are transferred to the fresh medium (13-15). A number of recent reviews nicely summarize current efforts in this scientific field (see, for example, (4, 11, 15)).

Mutation and selection are the key drivers in ALE experiments. However, the activity of specialized mobile elements (16), in particular, transposable phages, genomic islands and cryptic prophages may have a considerable influence on the stability of bacterial genomes by causing genomic rearrangements e.g. integrations, deletions, disruptions or inversions (16-18). Consequently, these elements represent prime candidates for removal in several recent genome reduction projects aiming at the construction of stable and predictable platform strains (19, 20). However, the impact of suchlike genomic modifications on the long-term genomic stability, mutation frequency and evolvability of the particular strain is usually not characterized.

In this study, we focused on *Corynebacterium glutamicum* representing one of the most important industrial platform organisms used for the production of L-glutamate and L-lysine (about 3.1 million and 2.2 million tons per year, respectively) (21, 22) and various further value-added products (23-26). Although laboratory evolution experiments with *C. glutamicum* proved to achieve promising results (6, 10, 27) investigations on long-term scales have not yet been performed. The genome of this Gram-positive soil bacterium contains three cryptic prophage elements (CGP1-3) of which the largest element, CGP3, is still inducible leading to death of the affected cell (28, 29). In a recent study, all three cryptic elements have been removed from the genome of *C. glutamicum* ATCC13032 resulting in strain MB001 with a genome reduced by 6% (19). This prophage-free variant prevailed to be a stable strain for metabolic engineering as reflected by several recent studies (19, 30-33).

In this work, we compared *C. glutamicum* ATCC 13032 wild type strain and its prophage-free variant MB001 in a long-term evolution experiment on glucose minimal medium. For both strains, a fast adaptation resulting in about ~20% increased growth rates was observed within the first ~100 generations. Genome sequencing of population samples as well as selected isolates revealed frequently occurring key mutations in *pyk* (pyruvate kinase) and *fruK* genes (*pfkB*, 1-phosphofructokinase) leading to increased glucose uptake and growth rates when reintroduced into the wild type background. Overall, our results revealed no significant differences between the two strains in terms of mutation frequency and genomic stability but even emphasized a positive trend of MB001 to evolve to higher growth rates under the chosen conditions.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Bacterial strains used in this study are listed in Table S1. *C. glutamicum* ATCC 13032 served as wild type strain (34). Strain MB001 represents a prophage-free variant of ATCC 13032 (19). *E. coli* DH5α was used for cloning and was cultivated at 37 °C in lysogeny broth (LB) (if not indicated otherwise). *C. glutamicum* cells were grown in complex medium consisting of Brain-

Heart-Infusion (BHI, 37 g/L) (DifcoTM BHI, Becton, Dickinson and Company (BD)) or in minimal medium CGXII (35) with 30 $mg \cdot L^{-1}$ biotin. Unless stated otherwise, 2% (w/v) glucose was added as carbon source. For growth studies, *C. glutamicum* cells from a fresh BHI agar plate were cultivated for 6 h in BHI at 30 °C. Subsequently, this preculture was used to inoculate a second preculture in CGXII minimal medium with 2% (w/v) glucose which was incubated at 30°C overnight. On the next day, the main culture was inoculated with cells of the 2nd preculture in defined CGXII media, as indicated. If necessary, kanamycin was added in a final concentration of 50 $\mu g \cdot mL^{-1}$ for *E. coli* and 25 $\mu g \cdot mL^{-1}$ for *C. glutamicum*.

Recombinant DNA work. Plasmids and oligonucleotides used in this work are specified in Table S2. Standard cloning methods including PCR, restriction and ligation of DNA were conducted according to established protocols (36). Cloning of plasmids was conducted using Gibson assembly (37). Synthesis of oligonucleotides and DNA sequencing were performed by Eurofins MWG Operon. The chromosomal integration of the barcode sequences, the mutations *pyk*_T12A, *pyk*_A20V, *pyk*_A271T, *fruK*_T6I, *fruK*_R71L and *corA*_Q307* and the deletion of 20 bp within the *corA* gene were done using the two-step homologous recombination method (38). Approx. 500 bp of the up- and downstream region of the replacement site were amplified by PCR using corresponding oligonucleotides listed in Table S3. Integrations of the 20 bp barcode sequences were checked by colony-PCR with the oligonucleotides WT_BC_fw and WT_BC_rv in the wild type strain and with MB_BC_fw and MB_BC_rv in the prophage-free MB001 strain. Insertions of the mutations and of the 12 bp deletion in *corA* were verified by sequencing applying the oligonucleotides pyk_T12A-seq_fw, pyk_T12A-seq_rv, pyk_A271T-seq_fw, pyk_A271T-seq_rv, *corA*_del_seq_fw, *corA*_del_seq_rv, fruK_T6I_seq_fw and fruK_T6I_seq_rv.

Long-term adaptive evolution experiment (LTAEE). The LTAEE was conducted with genomically barcoded ATCC 13032 and prophage-free MB001 strains. The cultivations were performed in 48-well FlowerPlates® at 30°C and a shaking frequency of 900 rpm in a microtron (Infors HT). The first and second precultivation were done as described in the bacterial strains and growth conditions part. The long-term experiment started with six biological replicates of each strain at an initial cell density of OD_{600} = 1. CGXII with 2% (w/v) glucose was used as standard defined minimal media. To keep differences in medium composition to a minimum, CGXII was prepared in sufficient amounts at the beginning of the experiment, aliquoted and stored at -80 °C. Every 48 h to 72 h 10 µl of stationary cells were used to inoculate 790 µl of fresh medium (dilution 1:40) in the microtiter plate. In total, 90 serial transfers were performed during this LTAEE. At defined time points (~approx. every 15 transfer) samples were taken and stored as glycerin cultures. For the first twenty cultivations, standard CGXII medium was alternated with iron depleted (1 µM Fe) CGXII medium in order to investigate potential effects of spontaneous prophage induction stimulated by alternations in iron availability (data not shown). Since, this did not result in any significant differences in growth or stress resistance between the wild type and MB001, this step was avoided in further transfers (data not shown) (final 70 transfers to standard CGXII medium). To reproduce and resolve the enhanced growth phenotype observed after ~100 generations, a repetition of the initial phase of the LTAEE was performed. Here, 16 serial transfers in standard CGXII medium containing 2% (w/v) glucose were conducted every 24 h in three biological replicates. Every second cultivation samples for glycerin cultures were taken. Studies to characterize the growth were conducted using the BioLector® microcultivation system of m2p-labs as described in (39). 48-well FlowerPlates® were applied for the cultivation at 30°C and 1200 rpm shaking frequency. The volume of the main culture was 750 µl of defined CGXII media with 2% (w/v) glucose (or alternative carbon source) and/ or further additives (e.g. Isopropyl-β -D-1-thiogalactopyranoside (IPTG), MMC, kanamycin). Typically, cultures were inoculated to a start OD_{600} of 1, unless specified otherwise. Measurements were taken every 15 minutes. Number of generations were calculated throughout the experiment by estimating a start OD of about ~0.5 and a final OD_{600} = 35-40 resulting in approx. 6-7 generations for each cultivation.

Genome sequencing. Genomic DNA of *C. glutamicum* cells were purified using the NucleoSpin® Microbial DNA Kit (Macherey Nagel). About ~100 mg (of cell pellet) that corresponds to 1 ml BHI-overnight culture and yielded in approx. 10 µg DNA. 4 µg of genomic DNA were used for library preparation and indexing with the TruSeq DNA PCR-free sample preparation kit (Illumina). Quantifications of the resulting libraries were conducted using KAPA library quant kits (Peqlab) and were normalized for pooling. A MiSeq sequencing device (Illumina) was used for paired-end sequencing with a read-length of 2 x 150 bases. Data analysis and base calling were accomplished with the Illumina instrument software and stored as fastq output files. Obtained sequencing data were imported into CLC Genomics Workbench (Qiagen Aarhus A/S) for trimming and base quality filtering. The output was mapped to accession BX927147 as the *C. glutamicum* ATCC 13032 reference genome (40) or CP005959 as the reference genome for MB001 (19). The resulting mappings were used for the quality-based SNP/ variant detection with CLC Genomics Workbench. The detected SNPs were manually inspected for relevance.

Dilution drop stress tests. *C. glutamicum* cells of an BHI overnight culture were diluted to an OD_{600} = 1 in phosphate buffered saline (PBS) solution (137 mM NaCl, 2.7 mM KCl, 20 mM Na_2HPO_4, 1.8 mM KH_2PO_4). Eight serial dilutions (each 1:10) were prepared and 3 µl were spotted on CGXII agar plates containing 2% (w/v) glucose. To test for differences in stress resistance, cells were treated by different procedures before spotting on agar plates (heat stress, 1 h at 42°C; UV stress, UV light exposure for 1 or 5 min (254 nm, 6 W); osmotic stress, agar plate containing 1 M NaCl). If not indicated otherwise, plates were incubated for 48 h at 30 °C.

Flow cytometry. Flow cytometry was used to determine the ratio between eYFP and E2-Crimson positive cells during the competitive growth experiment of ATCC 13032 and MB001 (Figure 1). Measurements and sorting were performed with a FACS Aria II (BD). A blue solid state laser with an excitation wavelength of 488 nm (to excite eYFP) and a red gas laser for excitation at a wavelength of 633 nm (to excite E2-Crimson) were used. Cytometer set-up, measurements and sorting were conducted as described in Helfrich et al. (29).

Determination of glucose uptake rate. Glucose concentrations of culture supernatants were determined by using the D-Glucose Kit (Roche), which is an enzyme-based UV test, according to manufactory's protocol. Samples of 500 µl were taken at defined time points (Figure S3) and cells were separated from the media by centrifugation (16,000 *g*, 3 min). The obtained supernatant was directly used for the assay or stored at -20 °C. Uptake rates were calculated as described in Frunzke et al. (41) by the following equation:

$$\frac{S}{M} \times \mu \left[\frac{mmol \cdot L^{-1} \cdot OD^{-1}}{gDW \cdot L^{-1} \cdot OD^{-1}} \cdot h^{-1}\right] = \left[\frac{mmol}{gDW \cdot h}\right]$$

S is the slope of a regression line which was calculated by plotting the glucose concentrations against the optical density at 600 nm OD_{600}. M describes the correlation between dry weight and OD by assuming an OD_{600} of 1 corresponds to 0.25 g dry weight per Liter. The growth rate µ was calculated as described in following section.

Determination of growth rate. Based on the growth curves obtained by cultivations in the BioLector® system or by shaking flask experiments growth rates were calculated using R (42) by applying following exponential fit to the exponential growth phase.

$$N(t) = N_0 * e^{-\mu \cdot t}$$

In this equation μ describes the growth rate, t is the time, N(t) population size at time point t and N_0 represents initial population size. Prior to the exponential fit, measured backscatter data were corrected by subtracting background values, which were defined as backscatter values measured after the first time point (t = 15 min). In calculations of growth rates that are based on shaking flask experiments correction to background values was not needed.

Pyruvate Kinase Assay. The activity of the pyruvate kinase was measured in cell crude extracts using the Pyruvate Kinase Activity Assay Kit (Sigma). This enzymatic assay is based on the coupling with a pyruvate oxidase. The product is proportional to formed pyruvate and can be measured at 570 nm. Measurements were conducted using an Infinite 200 PRO reader (Tecan). Crude extracts were prepared of cells that were harvested in mid-exponential phase at an OD_{600} ~ 5.

RESULTS

Competitive growth of the prophage-free *C. glutamicum* strain MB001. In previous studies, we reported on the construction of a prophage-free variant of *C. glutamicum* strain ATCC 13032 named MB001 displaying several positive features for metabolic engineering (19). Interestingly, different studies also revealed a slightly increased growth rate in comparison to the wild type strain. To test for a competitive growth advantage of MB001, fluorescent reporter genes were integrated into the genome of the wild type strain ATCC 13032 and MB001 (19). With the resulting strains (mixed 1:1), a competitive growth experiment was conducted in glucose minimal medium. The composition of the population was analyzed by flow cytometry for twelve serial transfers. Remarkably, already after three serial transfers a significant competitive advantage for strain MB001 (for both reporter strains, *vice versa*) was observed (Figure 1).

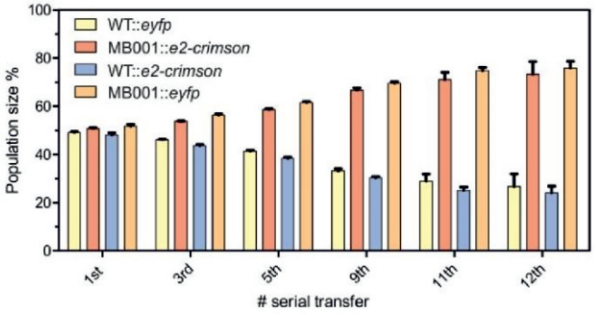


FIG 1 Competitive growth experiment of *C. glutamicum* ATCC 13032 and MB001. *C. glutamicum* ATCC 13032 and MB001 strains were both labelled *vice versa* with a genomically encoded yellow or far red fluorescent protein (eYFP and E2-crimson) (19). The resulting strains were compared *vice versa* in a competitive growth study on CGXII minimal medium containing 2% (w/v) glucose. Cells were serially transferred into fresh medium after 24 h. Production of the fluorescent proteins was induced by adding IPTG in a final concentration of 0.5 mM six hours before samples for flow cytometry analysis were taken. Strains from the co-cultivation experiment were discriminated by flow cytometry determining the fraction of eYFP and E2-Crimson positive cells.

This finding supports a slightly increased growth rate of the prophage-free MB001 strain compared to the wild type. However, we wondered whether the lack of all prophage regions (~6% of the genome) would somehow affect the genomic stability or evolvability of MB001 and addressed this question by a long-term experiment presented in the following.

Long-term adaptive evolution experiment. In a long-term evolution experiment, the robustness and adaptive potential of *C. glutamicum* ATCC 13032 and the prophage-free strain MB001 were compared. Prior to the ALE experiment, both strains were genomically barcoded by a short unique DNA sequence to trace the particular strains throughout the experiment and detect contaminations. Cells were grown in repetitive batch cultures in glucose minimal media (CGXII with 2% (w/v) glucose) and serially transferred (90x) from the stationary phase into fresh medium (Figure 2). For each strain, six independent cells lines were adaptively evolved in parallel for about ~630 generations. Frozen glycerin stocks were prepared throughout the experiment to allow for further analysis and/or re-inoculations (for details see material and methods). The overall fitness in terms of growth rate was analyzed at indicated time points at the population level as well as for 24 isolates for each strain (Figure 3). Remarkable, a significant increase in fitness was observed for all six cell lines after approx. 100 generations (Figure 3). The box plot shown in Figure 3A illustrates the high diversity of the populations narrowing again after about 140 generations coinciding with an overall increase of competitive fitness. In this initial phase of the experiment ATCC 13032 exhibited an increase in growth rate from 0.51 ± 0.01 h^{-1} to 0.62 ± 0.04 h^{-1}, whereas MB001 strains adaptively evolved from 0.53 ± 0.01h^{-1} to 0.65 ± 0.03 h^{-1}.This adaptation step was reproduced in a second ALE experiment, where a significant fraction of isolated clones exhibited a leap in their fitness after about 80 generations. Differences in the final OD_{600} values of stationary cells were not observed (data not shown) and both evolved strains behaved comparable regarding their sensitivity towards stress conditions, e.g. heat shock, UV and osmotic stress (Figure S2). Finally, we present the two isolated top strains of this ALE experiment named UBw (originating from ATCC 13032, μ = 0.64 ± 0.01 h^{-1}) and UBm (MB001, μ = 0.67 ± 0.01 h^{-1}). These strains featured the highest growth rates (UB = Usain Bolt) observed for the isolated clones and were therefore analyzed in more detail in the following experiments. Improved growth behavior was also demonstrated in standard shaking flask experiments, where both strains revealed an average increase of more than 20% in growth rate in comparison to their parental strain (Figure S3).

Adaptation on the genomic level. A central goal of ALE experiments is the identification of key mutations causing an improvement in fitness under the respective selective conditions. For this purpose, genome sequencing of two cell lines (out of six) of each strain was performed at different time points (after 100, 140, 540 and 630 generations). Mapping of the read data to the published genome sequences (19, 40) revealed several transposon insertions and nearly 400 SNPs (Table S6 & S7). By excluding mutations that were already present in ancestor strains 70 potentially causative SNPs were identified (Figure S4). Additionally, five transposon insertions were identified for the wild type as well as for the prophage-free strain (Table S7).

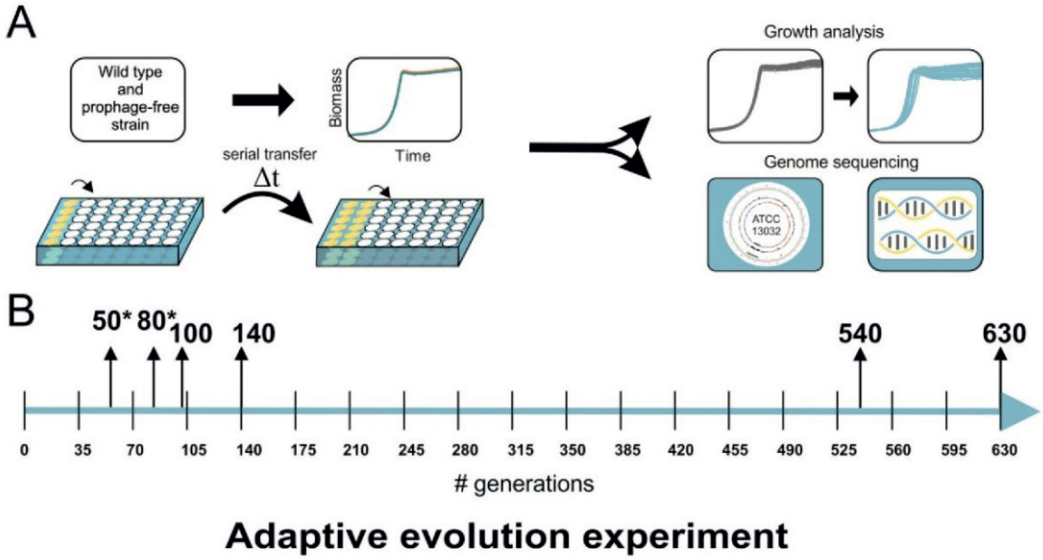


FIG 2 Long-term adaptive evolution experiment. **A.** *C. glutamicum* ATCC 13032 and its prophage-free variant MB001 were adaptively evolved towards increased growth rates on glucose minimal medium (CGXII medium with 2% (w/v) glucose). Overall, six independent cell lines for each strain were evolved by repetitive batch cultivation using 48-well plates incubated at 30°C (experimental details are given in the material and methods section). **B**. Time line of the ALE experiment. Highlighted are the samples further analyzed by genome sequencing. Samples of the second reproduction of the experiment are marked with an asterisk.

As expected the number of SNPs increased over time for all investigated cell lines (Figure S5), but no significant differences, neither in the mutation frequency nor in transposon activity, was observed between ATCC 13032 and MB001 (Figure S5, Table S7). To identify causative key mutations, the SNPs were clustered with respect to their time of appearance in the ALE experiment and their frequency within a particular gene or genomic region (Table 1). Considering that all six initial cell lines displayed improved fitness after ~100 generations on glucose minimal media (Figure 3), we were particularly interested in mutations appearing in this early stage of the ALE experiment. Remarkably, the *pyk* gene, encoding pyruvate kinase, was one of the most frequently affected genes with five different mutations (Table 1, Figure 4). Three of these mutations (D175G, A20V and T12A) were observed independently in different strains at different time points. In particular, these *pyk* mutations already appeared in an early stage of the experiment (after 100 & 140 generations) in low frequencies and displayed a competitive dynamic progression within the population (Figure 4, Table S4). However, significant differences in the persistence of *pyk* mutations were observed: Whereas, for example, mutation A271T increased in frequency from 0 to almost 100%, mutation A20V (orange circles) and T12A (black circles) were almost equal after 540 generations but a shift towards T12A was observable after 630 generations (Figure 4, Table S4). Further targets frequently hit by mutations were *corA*, encoding a CorA-like Mg^{2+}/ Co^{2+} transporter, *hmuV*, encoding a component of a hemin transport system, *soxA* (sarcoside oxidase and *fruK* (*pfkB*) encoding a phosphofructokinase. Especially mutations in *hmuV* were observed many times but displayed no competitive progression in the population. The frequencies of mutations found in *hmuV* and *soxA* were low (<30%) in comparison to other mutations. However, analysis of the sequencing data revealed the phosphofructokinase gene *fruK* as another important key target. Within this experiment, *fruK* was hit by two mutations and by an insertion of a transposon (Figure 4, Table S5). It is noteworthy that one of the two *fruK* SNPs (T6I) occurred independently at different time points in both prophage-free cell lines whereas mutation R71L was found only in strain UBm showing the highest growth rate of all isolated clones. In the repetition of the initial phase of the ALE experiment *fruK* was hit by a transposon insertion at a time point where a significant jump in fitness was occurring (Table S5). Besides *pyk* and *fruK*, also different SNPs as well as transposon insertions were found in the *corA* gene. These were, however, only observed at later stages in the ALE experiment suggesting epistatic interactions of *corA* mutations supporting the evolution towards higher growth rates on glucose (Figure 3, Table S4, Table S7).

Impact of *pyk*, *fruK* and *corA* mutations. In the following, we studied the impact of selected *pyk*, *fruK* and *corA* mutations on *C. glutamicum* fitness by reintroduction into the ATCC 13032 (WT) strain background (Table 3). Due to low frequencies of *soxA* and *hmuV* mutations (<30 %, Figure 4, Table S4) we assumed them not to be of central importance for the improved growth on glucose. The *pyk* mutations A271T, A20V and T12A were selected, since the latter two differed in their competitive behavior (A20V and T12, Figure 4) and A271T was identified as the only *pyk* mutation in UBm. Besides the two *fruK* mutations T6I and R71L, the non-sense mutation Q307* in *corA* and the impact of a 12-bp deletion (near the C-terminus) were also examined in the parental strain background. Remarkably, comparative growth

analysis revealed a significant positive effect of all tested *pyk* and *fruK* mutations, whereas single mutations of the *corA* gene showed no significant effect (Q307*) or resulted in a reduced growth rate (12 bp deletion) (Figure 5).

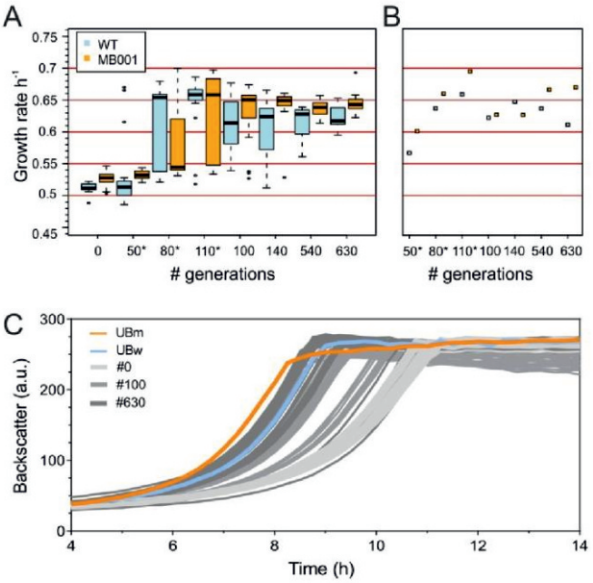


FIG 3 Fitness improvement of evolved strains. Growth rates of 24 clones isolated from each strain at the indicated time point **(A)** as well as for the population mixture **(B)** are shown (for the results of the second cell line, see Figure S1). To highlight outliers, whiskers of the boxplot represent values within 1.5 IQR. Time points marked with an asterisk are derived from the repetition of the ALE experiment. **(C)** The growth of the top isolated strains UBm (MB001 derivative) and UBw (based on ATCC 13032) is compared to clones isolated at the beginning of the experiment (#0), after 100 and after 630 generations.

Combination of *pyk* A271T and *fruK* R71L (identified in strain UBm) led to an even higher growth rate on glucose demonstrating the synergistic effect of these key targets (Figure 5). However, none of the tested mutants reached the fitness level of the isolated UBm strain (Table 3). Based on these data and under consideration of the high frequency of *corA* mutations in this ALE approach, it can be assumed that *corA* mutations are also of synergistic nature but require other adaptations beforehand. Further analysis of *pyk* mutations *via in vitro* enzyme assays revealed a significantly decreased pyruvate kinase activity for the single mutants T12A and A20V and the strain UBw (containing A20V). In contrast, no difference in pyruvate kinase activity was measured for A271T (identified in UBm) under the tested conditions. Remarkably, this mutant showed the strongest impact on growth (Table 3). Analysis of glucose consumption rates of evolved and parental strains revealed an increase of up to 40% reached by the *C. glutamicum* MB001 derived strain UBm (Table 3). Also the *pyk* A271T and *fruK* R71L mutants displayed increased uptake rates, whereas *pyk* T12A mutation did not significantly impact glucose import (Table 3). This effect on glucose uptake was also supported by the analysis of a promoter fusion of *ptsG* to *eyfp,* which revealed an increased *ptsG* expression in UBw and UBm as well as strains carrying *pyk* or *fruK* mutations causing a fitness increase (Table 3, Figure S6). The *ptsG* gene encodes the enzyme II of the phosphotransferase system representing the major system involved in glucose uptake in *C. glutamicum*.

Evolutionary trade-offs - Utilization of different carbon sources. In the described ALE experiment, we selected for strains harboring mutations that improve the fitness of *C. glutamicum* on glucose minimal medium. However, the missing selective pressure may tolerate the fixation of mutations detrimental for the catabolism of other carbon sources. Here, we investigated the growth of UBw and UBm on CGXII minimal medium containing fructose, gluconate, ribose, lactate or acetate as sole carbon sources. In addition, we also tested growth on BHI complex medium (Figure 6). Evolved strains already displayed a slightly decelerated growth rate on BHI medium reflecting the systemic adaption to minimal media.

TABLE 1 Key mutations identified in the ALE experiment. Mutations were clustered according to their frequencies in the respective genes. For a complete list of all mutations, see able S6.

Locus	Gene	#Mutations	Annotation
cg0080	*corA*	7	putative CorA-like Mg^{2+}/ Co^{2+} transporter protein, MIT-family
cg2291	*pyk*	5	pyruvate kinase (EC:2.7.1.40)
cg0469	*hmuV*	5	hemin transport system, ATP-binding protein
cg1781_cg1783*	*soxA*	4	sarcosine oxidase, cg1781 encodes the C-terminal fragment, cg1783 encodes the N-terminal fragment
cg2807, cg2600	*tnp*	4	transposase, putative pseudogene
cg0418	-	3	putative aminotransferase, involved in cell wall biosynthesis
cg3213	-	3	putative secreted protein
cgtRNA_3558	Leu tRNA	3	cgtRNA_3558, Leu tRNA
cg2119	*fruK (pfkB)*	3	1-phosphofructokinase (EC:2.7.1.56)
cg2136	*gluA*	2	glutamate uptake system, ABC-type, ATP-binding protein
cg1419-cg1420	*gatB* (cg1420)	2	putative Na^{+}-dependent transporter

*cg1782 encodes a transposase

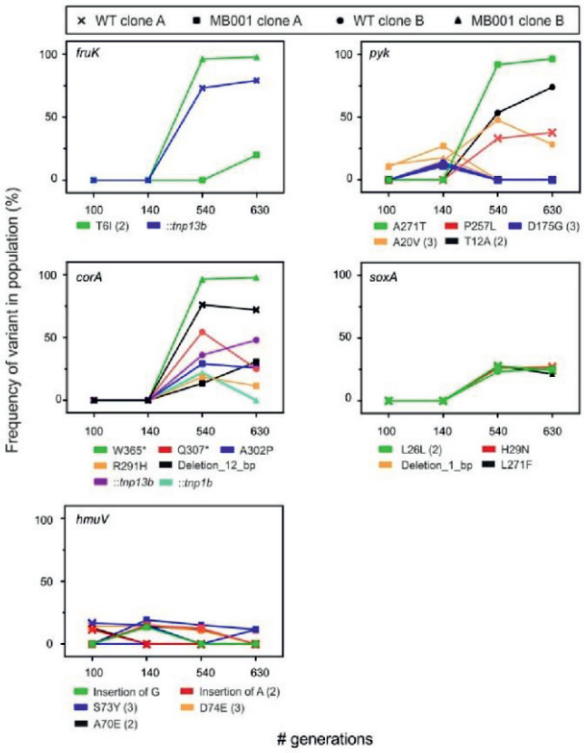


FIG 4 Frequencies of mutations in selected genes at different time points. Mutations affecting the same gene are grouped in one graph. The number in brackets behind the mutation indicates how often a particular mutation was observed in independent cell lines. Data of two wild type (cross, circle) and prophage-free cell lines (square, triangle) are shown for each investigated time point (approx. 100, 140, 540 and 630 generations) of the main ALE experiment. Frequency of mutations was defined as the count of the mutation divided by the coverage measured at this position.

On gluconate, ribose, and acetate UBw showed a strongly impaired growth in comparison to the parental strains. Growth on lactate was even totally abolished due to a transposon insertion into the lactate permease gene (Table S6). In contrast, UBm cells revealed competitive or even improved growth on all tested carbon sources (Figure 6). As a next step, we investigated the effects of *pyk* and *fruK* mutations on the utilization of fructose and gluconate. Whereas the *fruK* T6I mutation led to a strong growth defect on fructose, R71L improved the fitness. On gluconate minimal medium, the growth of *pyk* mutants A20V and T12A was strongly impeded, but the A271T mutant was not affected. In contrast, all *fruK* mutations caused a significant improvement when gluconate served as sole carbon source (Table 3, Figure S6). These results provide only a brief glimpse on the systemic changes of the evolved strains clearly affecting a variety of different metabolic routes and cellular processes.

DISCUSSION

In this study, we compared the potential of the biotechnological platform organism *C. glutamicum* ATCC 13032 and its prophage-free variant MB001 to evolve towards higher growth rates on glucose minimal medium. Sequencing of two independent cell lines for each clone as well as the genomes of isolated top strains revealed key mutations causing accelerated growth under the chosen conditions. Both strains showed a comparable adaptive behavior and already in the first 100 generations an improvement in fitness by about 20%. In Lenski's famous long-term experiment an increase of about 20% was achieved after ~1000 generations (43). Although hard to compare, our experiment reveals a high potential of *C. glutamicum* to evolve towards higher growth rates in comparison to *E. coli* wild type, which already starts at a higher fitness level on glucose minimal medium.

TABLE 2 Mutations identified in the evolved strains UBw and UBm

Mutation	Annotation	UBw Freq.%	UBm Freq.%
S270S	cg3197, *psp5*, putative secreted protein	100	0
Deletion 12 bp, CGTCGACGATGG, position 593 to 604	cg0080, *corA*, putative CorA-like Mg^{2+}/Co^{2+} transporter protein, MIT-family	0	98.55
Deletion 1 bp, T at position 1456	cg0934, hypothetical protein, conserved	0	94.38
Deletion 1 bp, C, position 270	cg2380, putative membrane protein	97.58	0
Exchange A20V	cg2291, *pyk*, pyruvate kinase (EC:2.7.1.40)	100	0
Exchange A271T	cg2291, *pyk*, pyruvate kinase (EC:2.7.1.40)	0	100
Exchange R71L	cg2119, *fruK* (*pfkB*), 1-phosphofructokinase (EC:2.7.1.56)	0	100
Exchange S73Y	cg0469, *hmuV*, hemin transport system, ATP-binding protein	10.87	0
Exchange V123L	cg2069, *psp1*, putative secreted protein, CGP3 region	100	0
Stop Q93*	cg2380, putative membrane protein	0	100
G to A mutation, 181 bp upstream	cg0850, *whmD* homolog	0	100
Insertion of *tnp13b* (cg1782)	cg0080, *corA*, putative CorA-like Mg^{2+}/ Co^{2+} transporter protein, MIT-family	78	0
Insertion of *tpn5a* (cg0842)	cg3226, lactate permease	91	0
Insertion of *tnp13b* (cg1782) in IGR	cg1696, permease of the major facilitator superfamily and cg1697, *aspA*, aspartate ammonia-lyase (aspartase) (EC 4.3.1.1)	0	60

In another study, where *E. coli* cells were kept constantly in the exponential phase, growth rates improved even faster and after ~2000 generations an overall increase over 50% was observed (2).

Key mutations in *pyk* and *fruK*, which occurred independently in *C. glutamicum* in different cell lines and both strains, were verified as causative mutations for higher growth rates on glucose medium. The second independent ALE experiment confirmed the reproducibility of the adaptive progress of the strains and revealed a strong overlap in the genetic regions being affected. This ALE experiment resulted in the isolation of two top strains UBw (ATCC 13032 derivative) and UBm (from MB001), which were characterized in further detail. Here, the positive features of strain UBm deserve special emphasis. This strain derived from the prophage-free strain MB001 showed an increase in fitness of about ~28% and displayed – in contrast to UBw - almost no significant trade-off or even better growth on the other carbon sources tested (Figure 6).

A previous study on *E. coli* K12 revealed a high importance of its cryptic prophage elements on growth and stress resistance (18). However, we observed no significant differences in the adaptive behavior, mutation frequency or transposon activity of *C. glutamicum* ATCC 13032 wild type and its prophage-free derivative MB001 in our comparative ALE experiment (Figure S5, Table S7). Furthermore, all cryptic prophage elements were stably maintained throughout the ALE of the wild type strain. Although a small fraction of cells is continuously killed by the spontaneous activation of the cryptic prophage (CGP3) (28, 29) this burden has apparently no strong impact on the fitness under the chosen conditions. A reason for this is likely the presence of the xenogeneic silencer CgpS encoded in the prophage CGP3 and acting as a silencer of phage gene expression (44). Therefore, CgpS clearly inherits a central role in the maintenance of foreign DNA elements in the host genome.

The *pyk* gene, encoding pyruvate kinase (PK), was found to represent a key genetic region for an improvement of growth on glucose in *C. glutamicum*. This is in line with evolutionary studies of *E. coli*, where also *pyk* mutations were identified in independent studies (2, 45). Barrick et al. showed that re-introduction of *pyk* mutations into the wild type background improved growth rates and suggested that these mutations cause a decrease in PK activity (1, 46).

TABLE 3 Comparative analysis of evolved strains and strains carrying selected key mutations

Strain	Growth rate (h^{-1}) (glucose)	Growth rate fold change (%) (glucose)	Glucose uptake rate (nmol·min^{-1}·mg^{-1})	*ptsG* expression* (a.u.) (glucose)	PK activity (mU·(mg protein)$^{-1}$)	Growth on fructose**	Growth on gluconate**
ATCC 13032	0.52 ± 0.01	100.0 ± 1	89.3 ± 1.1	1 ± 0.012	299.4 ± 42.4	+/-	+/-
MB001	0.52 ± 0.01	-0.4 ± 1.2	92.8 ± 1.6	1.01 ± 0.01	259.0 ± 23.6	+/-	+/-
UBw	0.64 ± 0.011	+23.9 ± 1.7	99.4 ± 9.4	1.51 ± 0.02	85.0 ± 50.7	+/-	--
UBm	0.67 ± 0.01	+28.1 ± 1.2	125.6 ± 2.47	1.52 ± 0.01	266.7 ± 44.6	++	+
WT_*pyk*_T12A	0.62 ± 0.01	+18.9 ± 2.3	90.76 ± 3.8	1.35 ± 0.01	18.2 ± 15.0	+/-	--
WT_*pyk* A20V	0.64 ± 0.01	+22.4 ± 1.1	N.D.	1.29 ± 0.01	69.3 ± 17.4	+/-	--
WT_*pyk* A271T	0.57± 0.02	+10.4 ± 4.0	101.4 ± 8.7	1.24 ± 0.01	337.1 ± 53.1	+	+/-
WT_*fruK* R71L	0.56 ± 0.003	+8.3 ± 0.5	106.5 ± 9.8	1.16 ± 0.004	N.D.	++	+
WT_*fruK*_T6I	0.60 ± 0.01	+5.2 ± 1.7	N.D.	1.26 ± 0.01	N.D.	-	+
WT_*corA*_ Q307*	0.53 ± 0.01	+1.5 ± 1.3	N.D.	0.99 ± 0.01	N.D.	N.D.	N.D.
WT_*corA*_Δ12bp	0.44 ± 0.01	-14.5 ± 1.8	N.D.	1.01 ± 0.02	N.D.	N.D.	N.D.
WT_*pyk*_A271T _*fruK*_R71L	N.D.***	N.D.***	N.D.	1.30 ± 0.01	N.D.	++	+

N.D.: not determined
*normalized to wild type level (Figure S6B)
** The evaluation is based on the growth analysis show in Figure S7, the wild type strain served as reference (Figure S7)
++ very good, + good, +/- no difference, - bad, -- very bad
*** Growth experiments (see Fig. 5) were conducted with strains harboring a plasmid (reporter for *ptsG* expression) and are therefore not comparable with prior experiments

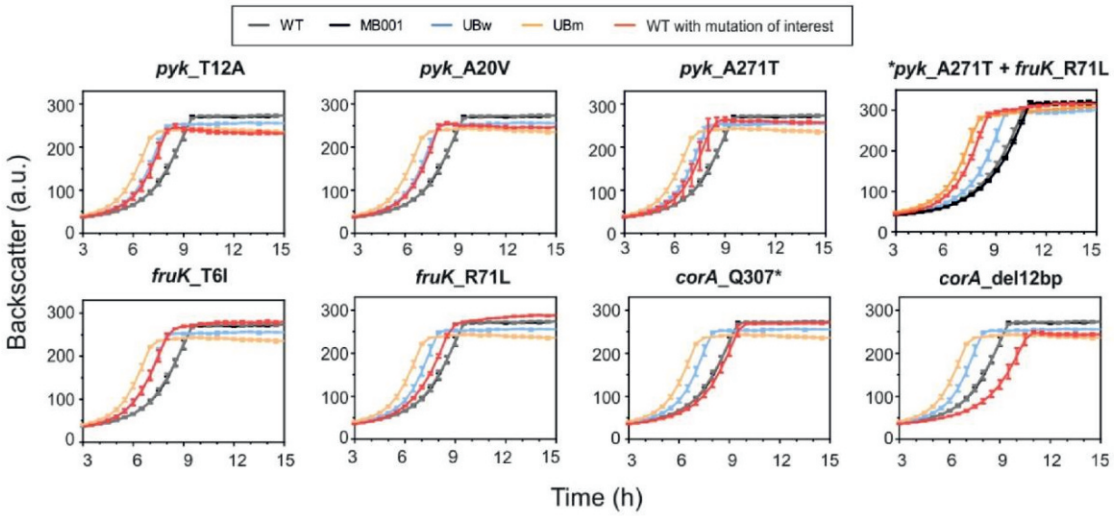


FIG 5 Impact of key mutations on the fitness in glucose minimal medium. Key mutations identified in *pyk*, *fruK* and *corA* (see Table 2) were introduced into the parental wild type strain *C. glutamicum* ATCC 13032 and investigated regarding their impact on growth in CGXII minimal medium containing 2% (w/v) glucose. Each graph is showing the average data of three biological replicates including standard deviation. *Growth data of *pyk*_A271T, *fruK*_R71L are based on plasmid-containing strains.

Lower enzymatic activity of PK would consequently lead to a higher PEP pool driving the phosphotransferase system (PTS) (45, 46). These data are in agreement with our study, as we confirm the positive effect of *pyk* mutations by reintroducing selected SNPs into the wild type background. Furthermore, we showed that two of the mutations (T12A, A20V), in fact, led to a decrease in PK activity, whereas the third mutation A271T did not significantly alter the PK activity (Table 3). Contrary to the suggested hypothesis, we have not observed an increased glucose uptake rate for strains harboring a less active PK, but for the A271T mutant (Table 3). A similar observation was also done in a recent study, in which no significant differences in glucose consumption were observed between the wild type and the Δ*pyk* strain (47). Comparing PK amino acid sequence of *C. glutamicum* to the published structure of *E. coli* (48) (sequence identity of 41%) revealed that the two residues T12 and A20 are located in close vicinity to the catalytic center. This finding is in line with the observed decrease in enzymatic activity. Strikingly, residue A271 was subject of another study where mutagenesis of this respective amino acid resulted in decreased PK activity upon increased levels of the allosteric activator fructose 1,6-bisphosphate (FBP) (49). Hence, we conclude PK is indeed an important target for fitness improvement. This is supported by observed mutations which either affect the catalytic PK activity or the allosteric regulation.

A further player of the central metabolism which was affected in several independent cell lines was the *fruK* gene, encoding the 1-phosphofructokinase (Pfk1). Pfk1 catalyzes the transfer of a phosphoryl group from ATP to fructose 1-phosphate (F1P) yielding FBP (Figure 7A). This reaction is supposed to be mainly relevant for fructose catabolism. However, since fructose was absent in this experiment a certain promiscuity of Pfk1 or at least a high potential to evolve other catalytic functions can be speculated. It is noteworthy that although *fruK* mutations were not identified in the *E. coli* ALE study, this gene was highly upregulated in their evolved strains (2). Introduction of the two *fruK* mutations, T6I and R71L, into our parental wild type strain resulted in a significant fitness increase (Table 3, Figure 5). Growth experiments on fructose minimal medium revealed a negative impact of the T6I mutation (Figure S6) comparable to a *fruK* deletion strain (50, 51). In contrast, R71L mutants acquired benefits by surpassing the growth of the parental strain (Table 3, Figure S6). Considering the crystal structures available for *E. coli* Pfk (52) suggests that both SNPs are not close to the active site but may result in conformational changes. In recent studies, a link between *fruK* deletion and an increase in glucose uptake was indeed described for *C. glutamicum* (50, 51). The authors suggested that *fruK* deletion results in an increased pool of F1P or in other hexose phosphates relieving SugR repression of *ptsG* (50). SugR is a pleiotropic transcriptional repressor of PTS genes and sugar phosphates including F1P, FBP and glucose-6-phosphate (G6P) (53) were described as potential effector molecules leading to the de-repression of PTS targets. Our results confirmed an increased level of *ptsG* and a higher glucose uptake rate, in particular for *fruK* mutants (Table 3). As fructose was not added to the growth medium, one can assume that this effect is unlikely a result of F1P accumulation, but may be caused by increased levels of other hexose phosphates (Figure 7).

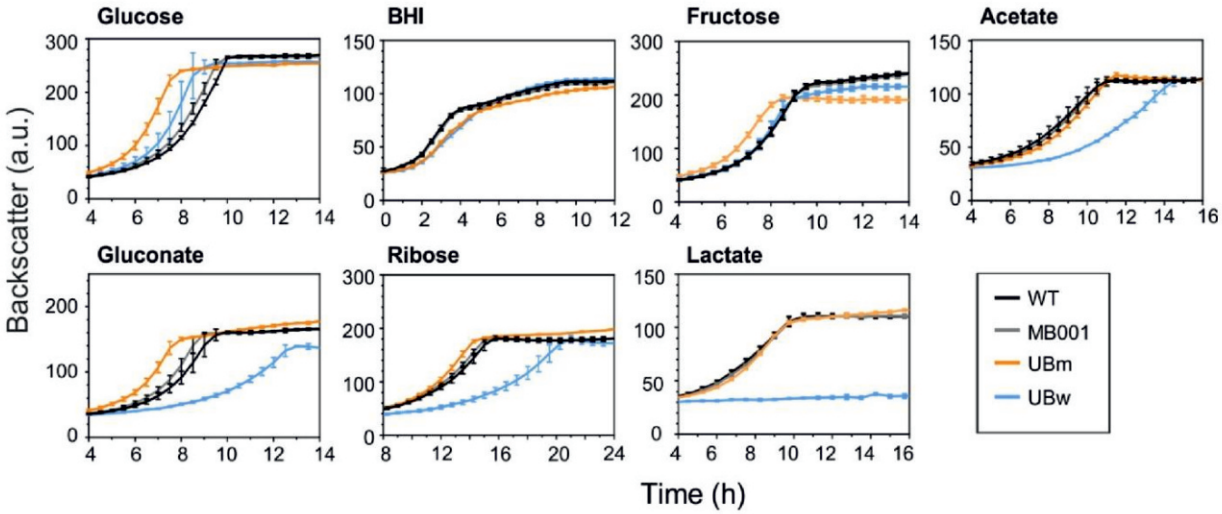


FIG 6 Analysis of trade-off effects of evolved strains growing on different carbon sources. The growth of *C. glutamicum* UBm, UBw and parental strains was analyzed on CGXII minimal medium supplemented with different carbon sources: glucose 2% (w/v), BHI (37g/L), fructose 2% (w/v), gluconate 2% (w/v), ribose 2% (w/v), lactate 1.5% (w/v) and acetate 1.5% (w/v). For each condition three biological replicates were analyzed.

However, in a *fruK ptsF* double knockout strain F1P is still detectable indicating the formation of F1P by unknown side reactions e.g. due to activities of phosphohexomutases (50).

Remarkably, the highest number of different mutations was identified in the *corA* gene, encoding a putative CorA-like Mg^{2+}/Co^{2+} transporter protein (Table 1). CorA transporters are beside MgtE the primary transporters of Mg^{2+} in bacteria (54). Both classes of transporters are identified in the genome of *C. glutamicum* (CorA: cg0080, MgtE1 and MgtE2: cg0275 and cg1276) and appear to be similar in mode of action as the electrochemical gradient is used to transport Mg^{2+} (54). Nevertheless, mutations were only found in the *corA* gene (Table S6). Two transpositions, two nonsense mutations (W365*, Q307*) and a deletion of 12 bp were identified - strongly indicating a loss of CorA function. As introduction of the *corA*Δ12bp mutant resulted in an impaired growth on glucose, we suggest that *corA* mutations are depending on epistatic interactions with other mutations occurring earlier in the experiment. Also in the *E. coli* ALE experiment of LaCroix et al. *corA* mutations were identified, but the impact on glucose uptake was not elucidated (2). Recently, it was shown that extracellular magnesium concentrations are directly linked to glucose consumption (55). Furthermore, in a magnesium limitation experiment *E. coli* cells consumed faster glucose by harboring higher metabolic rates (56). Altogether, these findings support the hypothesis that reduced CorA activity has a synergistic impact on fitness when glucose uptake is increased by other key events beforehand.

CONCLUSION

In summary, our study reports on the first long-term evolution experiment of *C. glutamicum* and highlights its strong metabolic adaptability. A significant fitness increase by about 22% was already observed after 100 generations. Furthermore, we describe key mutations in *pyk* and *fruK* which cause a significant increase of the growth rate on glucose minimal medium. Evidence for epistatic interactions is provided by the finding of several independent mutations in *corA* encoding a Mg^{2+} importer. The experiment resulted in the isolation of the top strains UBw (based on ATCC 13032) and UBm (MB001 derivative), where especially the MB001 derived strain displays several positive features and no obvious trade-off under the conditions tested in this study. UBm represents to date the *C. glutamicum* strain with the fastest growth rate on glucose minimal medium and, thus, provides a strong basis for future metabolic engineering approaches. The fact that similar genetic regions were affected in independent studies with the model strains *E. coli* and *C. glutamicum* also highlights the broad significance of these data for the understanding of microbial metabolic networks and for the identification of major bottlenecks for metabolic flux. The identification and further analysis of causative key mutations is not only relevant for future metabolic engineering but also enhances our understanding of biological systems.

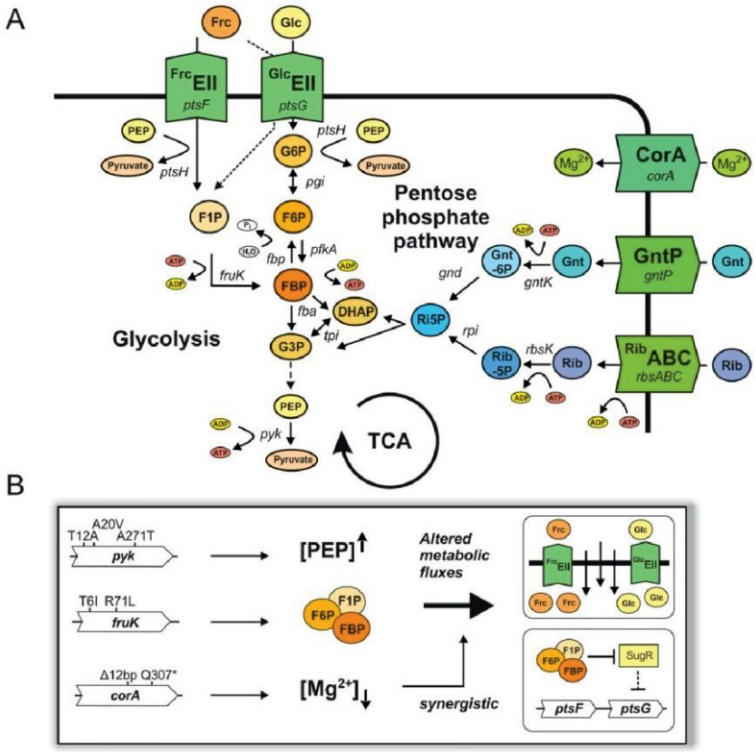


FIG 7: Impact of key mutations on central carbon metabolism. **A.** Overview of glycolysis, pentose phosphate pathway (PPP) and tricarboxylic acid cycle (TCA) in *C. glutamicum*. Genes are given in italics. Glc: glucose, Frc: fructose, Gnt: gluconate, Rib: ribose, G6P: glucose 6-phosphate, F6P: fructose 6-phospahte, FBP: fructose 1,6-bisphosphate, G3P: glyceraldehyde 3-phosphate, DHAP: dihydroxyacetone phosphate, PEP, phosphoenolpyruvate, Gnt-6P: gluconate 6-phosphate, Rib5P: ribose 5-phosphate, Ri5P: ribulose 5-phopshate, glycerin EII: sugar specific transport system, CorA: Mg^{2+} transporter, GntP: gluconate permease, ABC: Rib-ATP: ATP dependent ABC-transporter. **B.** Key mutations in *pyk*, *fruK* and *corA*. We hypothesize an impact on the PEP pool (*pyk* mutations), hexose phosphates (*fruK*) and Mg^{2+} homeostasis (*corA*), which also may act synergistically to increase glucose uptake.

ACKNOWLEDGMENTS

We thank Doris Rittman (Institute for Bio- and Geosciences 1, IBG-1) for assistance in sample preparation for genome sequencing and we wish her all the best for a happy and fulfilling retirement.

FUNDING

This study was supported by the priority program SPP1617 of the Deutsche Forschungsgemeinschaft (FR 2759/2-2) and by the Helmholtz Association (grant NH-VG-716).

REFERENCES

1. **Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF.** 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature **461:**1243-1247.
2. **LaCroix RA, Sandberg TE, O'Brien EJ, Utrilla J, Ebrahim A, Guzman GI, Szubin R, Palsson BO, Feist AM.** 2015. Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of *Escherichia coli* K-12 MG1655 on glucose minimal medium. Appl Environ Microbiol **81:**17-30.
3. **Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, Wu GC, Wielgoss S, Cruveiller S, Medigue C, Schneider D, Lenski RE.** 2016. Tempo and mode of genome evolution in a 50,000-generation experiment. Nature **536:**165-170.
4. **Abatemarco J, Hill A, Alper HS.** 2013. Expanding the metabolic engineering toolbox with directed evolution. Biotechnol J **8:**1397-1410.
5. **Portnoy VA, Bezdan D, Zengler K.** 2011. Adaptive laboratory evolution - harnessing the power of biology for metabolic engineering. Curr Opin Biotech **22:**590-594.
6. **Lee JY, Seo J, Kim ES, Lee HS, Kim P.** 2013. Adaptive evolution of *Corynebacterium glutamicum* resistant to oxidative stress and its global gene expression profiling. Biotechnol Lett **35:**709-717.
7. **Oide S, Gunji W, Moteki Y, Yamamoto S, Suda M, Jojima T, Yukawa H, Inui M.** 2015. Thermal and solvent stress cross-tolerance conferred to *Corynebacterium glutamicum* by adaptive laboratory evolution. Appl Environ Microbiol **81:**2284-2298.
8. **Ramanan S, Grampp G.** 2014. Drift, evolution, and divergence in biologics and biosimilars manufacturing. BioDrugs **28:**363-372.
9. **Reyes LH, Gomez JM, Kao KC.** 2014. Improving carotenoids production in yeast via adaptive laboratory evolution. Metab Eng **21:**26-33.
10. **Mahr R, Gätgens C, Gätgens J, Polen T, Kalinowski J, Frunzke J.** 2015. Biosensor-driven adaptive laboratory evolution of l-valine production in *Corynebacterium glutamicum*. Metab Eng **32:**184-194.
11. **Williams TC, Pretorius IS, Paulsen IT.** 2016. Synthetic Evolution of Metabolic Productivity Using Biosensors. Trends Biotechnol **34:**371-381.
12. **Chou HH, Keasling JD.** 2013. Programming adaptive control to evolve increased metabolite production. Nat Commun **4:**2595.

13. **Barrick JE, Lenski RE.** 2013. Genome dynamics during experimental evolution. Nat Rev Genet **14:**827-839.
14. **Conrad TM, Lewis NE, Palsson BO.** 2011. Microbial laboratory evolution in the era of genome-scale science. Mol Syst Biol **7:**509.
15. **Dragosits M, Mattanovich D.** 2013. Adaptive laboratory evolution -- principles and applications for biotechnology. Microb Cell Fact **12:**64.
16. **Darmon E, Leach DR.** 2014. Bacterial genome instability. Microbiol Mol Biol Rev **78:**1-39.
17. **Taylor AL.** 1963. Bacteriophage-Induced Mutation in Escherichia Coli. Proc Natl Acad Sci U S A **50:**1043-1051.
18. **Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, Wood TK.** 2010. Cryptic prophages help bacteria cope with adverse environments. Nat Commun **1:**147.
19. **Baumgart M, Unthan S, Rückert C, Sivalingam J, Grünberger A, Kalinowski J, Bott M, Noack S, Frunzke J.** 2013. Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. Appl Environ Microbiol **79:**6006-6015.
20. **Martinez-Garcia E, Jatsenko T, Kivisaar M, de Lorenzo V.** 2015. Freeing *Pseudomonas putida* KT2440 of its proviral load strengthens endurance to environmental stresses. Environ Microbiol **17:**76-90.
21. **Ajinomoto Co. I.** 2017. Ajinomoto Co., Inc. Consolidated Results Interim period ended September 30, 2016. Accessed 2017-01-26.
22. **Ajinomoto Co. I.** 2017. Feed-Use Amino Acids Market Price. Accessed 2017-01-26.
23. **Heider SA, Wendisch VF.** 2015. Engineering microbial cell factories: Metabolic engineering of *Corynebacterium glutamicum* with a focus on non-natural products. Biotechnology Journal **10:**1170-1184.
24. **Ikeda M.** 2016. Lysine Fermentation: History and Genome Breeding. Adv Biochem Eng Biotechnol.
25. **Wendisch VF, Jorge JM, Perez-Garcia F, Sgobba E.** 2016. Updates on industrial production of amino acids using *Corynebacterium glutamicum*. World J Microbiol Biotechnol **32:**105.
26. **Becker J, Giesselmann G, Hoffmann SL, Wittmann C.** 2016. *Corynebacterium glutamicum* for Sustainable Bioproduction: From Metabolic Physiology to Systems Metabolic Engineering. Adv Biochem Eng Biotechnol.
27. **Lee J, Saddler JN, Um Y, Woo HM.** 2016. Adaptive evolution and metabolic engineering of a cellobiose- and xylose- negative *Corynebacterium glutamicum* that co-utilizes cellobiose and xylose. Microb Cell Fact **15:**20.
28. **Frunzke J, Bramkamp M, Schweitzer JE, Bott M.** 2008. Population Heterogeneity in *Corynebacterium glutamicum* ATCC 13032 caused by prophage CGP3. J Bacteriol **190:**5111-5119.
29. **Helfrich S, Pfeifer E, Kramer C, Sachs CC, Wiechert W, Kohlheyer D, Noh K, Frunzke J.** 2015. Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations. Mol Microbiol **98:**636-650.
30. **Heider SA, Wolf N, Hofemeier A, Peters-Wendisch P, Wendisch VF.** 2014. Optimization of the IPP Precursor Supply for the Production of Lycopene, Decaprenoxanthin and Astaxanthin by *Corynebacterium glutamicum*. Front Bioeng Biotechnol **2:**28.
31. **Kortmann M, Kuhl V, Klaffl S, Bott M.** 2015. A chromosomally encoded T7 RNA polymerase-dependent gene expression system for *Corynebacterium glutamicum*: construction and comparative evaluation at the single-cell level. Microb Biotechnol **8:**253-265.
32. **Unthan S, Baumgart M, Radek A, Herbst M, Siebert D, Brühl N, Bartsch A, Bott M, Wiechert W, Marin K, Hans S, Krämer R, Seibold G, Frunzke J, Kalinowski J, Rückert C, Wendisch VF, Noack S.** 2015. Chassis organism from *Corynebacterium glutamicum*--a top-down approach to identify and delete irrelevant gene clusters. Biotechnol J **10:**290-301.
33. **Lubitz D, Wendisch VF.** 2016. Ciprofloxacin triggered glutamate production by *Corynebacterium glutamicum*. BMC Microbiol **16:**235.
34. **Kinoshita S, Udaka S, Shimono M.** 1957. Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms. J Gen Appl Microbiol **50:**331-343.
35. **Keilhauer C, Eggeling L, Sahm H.** 1993. Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the ilvB-ilvN-ilvC operon. J Bacteriol **175:**5595-5603.
36. **Sambrook J, Russell DW.** 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.
37. **Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO.** 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods **6:**343-U341.
38. **Niebisch A, Bott M.** 2001. Molecular analysis of the cytochrome bc1-aa3 branch of the *Corynebacterium glutamicum* respiratory chain containing an unusual diheme cytochrome c1. Arch Microbiol **175:**282-294.
39. **Heyer A, Gätgens C, Hentschel E, Kalinowski J, Bott M, Frunzke J.** 2012. The two-component system ChrSA is crucial for haem tolerance and interferes with HrrSA in haem-dependent gene regulation in *Corynebacterium glutamicum*. Microbiology **158:**3020-3031.
40. **Kalinowski J, Bathe B, Bartels D, Bischoff N, Bott M, Burkovski A, Dusch N, Eggeling L, Eikmanns BJ, Gaigalat L, Goesmann A, Hartmann M, Huthmacher K, Krämer R, Linke B, McHardy AC, Meyer F, Mockel B, Pfefferle W, Pühler A, Rey DA, Ruckert C, Rupp O, Sahm H, Wendisch VF, Wiegräbe I, Tauch A.** 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. J Biotechnol **104:**5-25.
41. **Frunzke J, Engels V, Hasenbein S, Gätgens C, Bott M.** 2008. Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2. Mol Microbiol **67:**305-322.
42. **R Development Core Team.** 2016. R: A language and environment for statistical computing, R Foundation for Statistical Computing, http://www.R-project.org.
43. **Wiser MJ, Ribeck N, Lenski RE.** 2013. Long-term dynamics of adaptation in asexual populations. Science **342:**1364-1367.
44. **Pfeifer E, Hünnefeld M, Popa O, Polen T, Kohlheyer D, Baumgart M, Frunzke J.** 2016. Silencing of cryptic prophages in *Corynebacterium glutamicum*. Nucleic Acids Res **44:**10117-10131.
45. **Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE.** 2006. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. Proc Natl Acad Sci U S A **103:**9107-9112.
46. **Schneider D, Duperchy E, Coursange E, Lenski RE, Blot M.** 2000. Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. Genetics **156:**477-488.
47. **Chai X, Shang X, Zhang Y, Liu S, Liang Y, Zhang Y, Wen T.** 2016. A novel pyruvate kinase and its application in lactic acid production under oxygen deprivation in *Corynebacterium glutamicum*. BMC Biotechnol **16:**79.
48. **Mattevi A, Valentini G, Rizzi M, Speranza ML, Bolognesi M, Coda A.** 1995. Crystal structure of *Escherichia coli* pyruvate kinase type I: molecular basis of the allosteric transition. Structure **3:**729-741.
49. **Valentini G, Chiarelli L, Fortin R, Speranza ML, Galizzi A, Mattevi A.** 2000. The allosteric regulation of pyruvate kinase. J Biol Chem **275:**18145-18152.
50. **Wang Z, Chan SH, Sudarsan S, Blank LM, Jensen PR, Solem C.** 2016. Elucidation of the regulatory role of the fructose operon reveals a novel target for enhancing the NADPH supply in *Corynebacterium glutamicum*. Metab Eng **38:**344-357.
51. **Hasegawa S, Tanaka Y, Suda M, Jojima T, Inui M.** 2017. Enhanced Glucose Consumption and Organic Acid Production by Engineered *Corynebacterium glutamicum* Based on Analysis of a *pfkB1* Deletion Mutant. Appl Environ Microbiol **83**.
52. **Baez M, Merino F, Astorga G, Babul J.** 2008. Uncoupling the MgATP-induced inhibition and aggregation of *Escherichia coli* phosphofructokinase-2 by C-terminal mutations. FEBS Lett **582:**1907-1912.
53. **Gaigalat L, Schluter JP, Hartmann M, Mormann S, Tauch A, Pühler A, Kalinowski J.** 2007. The DeoR-type transcriptional regulator SugR acts as a repressor for genes encoding the phosphoenolpyruvate:sugar phosphotransferase system (PTS) in *Corynebacterium glutamicum*. BMC Mol Biol **8:**104.
54. **Groisman EA, Hollands K, Kriner MA, Lee EJ, Park SY, Pontes MH.** 2013. Bacterial Mg2+ homeostasis, transport, and virulence. Annu Rev Genet **47:**625-646.

55. **Christensen DG, Orr JS, Rao CV, Wolfe AJ.** 2017. Increasing growth yield and decreasing acetylation in *Escherichia coli* by optimizing the carbon-to-magnesium ratio in peptide-based media. Appl Environ Microbiol.
56. **Chubukov V, Sauer U.** 2014. Environmental dependence of stationary-phase metabolism in *Bacillus subtilis* and *Escherichia coli*. Appl Environ Microbiol **80:**2901-2909.

4. Discussion

4.1 Prophages in bacterial genomes

Viral elements are inherent parts of bacterial chromosomes that crucially influence bacterial physiology by providing beneficial genes and regulatory switches (Feiner *et al.*, 2015, Wang *et al.*, 2010). Spontaneous activation of prophage-like elements (SPI), which is the in focus of this work, is a common phenomenon affecting the dynamics of bacterial populations. Usually, the SOS response is considered to be the key trigger of SPI and is induced by sporadically occurring DNA damage (Johnson & Grossman, 2015, Nanda *et al.*, 2015, Pennington & Rosenberg, 2007). However, a few cases are described to activate prophage-like elements by alternative pathways (Rozanov *et al.*, 1998, Cardinale *et al.*, 2008, Ghosh *et al.*, 2009, Hong *et al.*, 2010) reflecting fundamental insights into how the regulatory circuits of viruses and their hosts can be connected. Nucleoid-associated proteins (NAP), which fulfill functions as xenogeneic silencers (XS), are generally used to keep the expression of exogenous genetic elements under control (Navarre *et al.*, 2007). It was shown that the release of XS-mediated repression leads to increased expression of foreign genes and may even cause the activation of defective prophages (Hong *et al.*, 2010, Navarre *et al.*, 2007). Furthermore, it was suggested that XS proteins facilitate the incorporation of foreign, accessory genes into the hosts' regulatory circuits (Navarre *et al.*, 2007, Singh *et al.*, 2016).

The focus of the present thesis was a quantitative description of the dynamics of spontaneous SOS response and SPI of the cryptic prophage CGP3, as well as its impact on bacterial physiology in a long-term evolution experiment. Moreover, a further aim was to identify the central regulatory mechanisms controlling CGP3 activity and their impact on SPI.

Previous studies demonstrated that the defective CGP3 prophage is activated spontaneously under non-induced conditions (Frunzke *et al.*, 2008) and first single-cell experiments revealed a positive correlation with the SOS response (Nanda *et al.*, 2014). In this work, promoter-reporter fusions were constructed to follow the response of the two pathways at the single-cell level using fluorescence microscopy. Live-cell imaging in microfluidic chip devices enabled quantitative investigations into the fate of the SOS- and prophage-induced cells and to trace their history within lineage trees of the respective microcolonies. Two-thirds of phage positive cells were activated according to the classic, SOS-linked model, but a remarkable observation of this study was that one third of the cells were induced SOS-independently.

Fishing for transcriptional regulators controlling the CGP3 activity resulted in the identification of the small, prophage-encoded protein CgpS, which has similar characteristics to

host-encoded XS proteins. Furthermore, it could be shown that this protein is used by CGP3 as a key regulatory player to maintain the silent, lysogenic state and, at the same time, represent a link to SOS-independent activation.

In all cases (SOS-dependent or independent), the induction of CGP3 led to the stop of cellular growth rather than to lysis. A competitive approach between the wild type and the prophage-free strain, MB001, revealed that the presence of CGP3, in fact, imposes low but significant fitness burden, which is sufficient to result in a slight growth advantage of the prophage-free variant (under standard laboratory conditions). To further investigate the impact of CGP3 and the stability of this defective prophage, a comparative long-term study (>600 generations) was conducted, in which a wild type strain and its prophage-free variant were juxtaposed for the first time. Astonishingly, no selective impacts regarding growth, mutation frequencies or genomic rearrangements were detected, emphasizing CGP3 to be very well-maintained in his host organism.

4.2 SOS-dependent and independent SPI

Since the introduction of the well-studied λ phage (Lederberg & Lederberg, 1953) investigations on activation mechanisms enabling silent prophages to switch into their virulent state are of high interest. Several studies suggested that the host's SOS response was a key trigger of prophage induction (Canchaya *et al.*, 2003, Casjens, 2003, Sekulovic & Fortier, 2016), but investigations at the single-cell level have not been done before, which, for example, would allow enhanced examination of alternative, responsible pathways. For CGP3 induction, the SOS-dependency was clearly proven by MMC experiments since its addition to *C. glutamicum* cells caused a high induction rate (Donovan *et al.*, 2015). Furthermore, activation of this cryptic prophage was found to occur in the absence of any external triggers in a small fraction of the population (Frunzke *et al.*, 2008, Nanda *et al.*, 2014). Hence, the question arose as to whether sporadic DNA damage is the only reason for SPI and, for this reason, the aim was set to perform a quantitative description of SOS and CGP3 induction at the single-cell level using live cell imaging of the selected reporter strains (P_{recA}-*venus* for SOS and P_{lys}-*crimson* for CGP3 induction). This imaging approach facilitated a spatiotemporal resolution (Grünberger *et al.*, 2015) and overcame limitations of conventional snapshot analysis, such as given by flow cytometry or fluorescence microscopy with fixated cells (McCool *et al.*, 2004, Nanda *et al.*, 2014, Pennington & Rosenberg, 2007, Simmons *et al.*, 2009). By tracking >4000 cells, in 5.7% an activation of the SOS reporter and in 0.7% SPI of CGP3 were detected (Helfrich *et al.*, 2015).

Previous single-cell experiments revealed that spontaneous activation of the SOS response is caused by incidental DNA lesions, due to sporadically occurring DNA double-strand breaks (Pennington & Rosenberg, 2007). Especially stressful conditions may lead to spontaneous DNA breaks, which result in an upregulation of error-prone DNA polymerases (Foster, 2007). The SOS-coupled processes ultimately increase the mutation rate and finally may accelerate bacterial evolution (Rosenberg *et al.*, 2012). Comparing the SOS rate with that of the current literature, a relatively high fraction of SOS-positive cells was obtained (published range 0.09 – 3.1% (Simmons *et al.*, 2009, Kamensek *et al.*, 2010, Pennington & Rosenberg, 2007, Nanda *et al.*, 2014)). However, the high amount can be explained by the application of a rather low threshold, which was used to assess fluctuations in the SOS reporter output, whereas previous flow cytometry experiments used a stringent cut off to distinguish between SOS-"ON" and -"OFF" cells (Nanda *et al.*, 2014). Nonetheless, using the low threshold, a fraction of transient SOS positive cells (~18%) was determined, which were presumably able to recover from sporadic DNA damage and resume growth. Although recovering cells suffered from DNA lesions, it is likely, due to an increased number of mutations, that they may significantly impact the overall evolvability of the population (Galhardo *et al.*, 2007). Remarkably, none of the transient SOS-positive cells displayed prophage induction, suggesting a certain threshold value needs to be reached to finally trigger the CGP3 induction (Helfrich *et al.*, 2015). A similar threshold behavior was shown in UV prophage induction experiments, in which the number of induced lysogens increased with increasing UV dose (Kneser, 1966), implying that viruses avoid premature induction processes and therefore require strict regulation.

The obtained SPI rate is in line with previous studies (0.1 - 1% (Frunzke *et al.*, 2008, Nanda *et al.*, 2014)). It is striking that spontaneous activations of viral-like elements were already observed in early the 1950s by the famous microbiologist André Michel Lwoff. In the absence of any external trigger, he detected free phage particles in lysogenic bacterial populations (Lwoff, 1953). Remarkably, the so far highest spontaneously produced phage titer was ~10^9 particles per mL and was reported for BTP1, a prophage of the pathogenic strain *Salmonella* Typhimurium ST313 (Owen *et al.*, 2017) (for comparison only: the lowest spontaneously induced phage titer was reported for prophage LES400 of *P. aeruginosa* with ~10^2 particles (Fothergill *et al.*, 2011), whereas a moderate titer (~10^4 phages) was found for prophage Mu of *E. coli* (Howe & Bade, 1975)). The physiological relevance of BTP1 and the high spontaneous titers are currently under study, but the authors speculate that phage-mediated lysis may be tolerated due to beneficial traits that are associated with the lysogenic state (e.g. phage immunity) (Owen *et al.*, 2017). Nevertheless, spontaneous activations of viral-like elements are reported to significantly promote the fitness of bacterial communities (Nanda *et al.*, 2015). Especially for pathogenic strains, sacrificing a small fraction to increase the good of the whole population is a

crude, but efficient, method to release toxins, as, for example, was shown for *Aggregatibacter actinomycetemcomitans* (Stevens *et al.*, 2013) and a Shiga toxin-producing *E.coli* strain (Shimizu *et al.*, 2009).

Spontaneous induction of CGP3 in *C. glutamicum* cells did not result in cell lysis, but recovery was also not observed. Moreover, if CGP3 was activated, in all cases the cellular growth stopped indicating death or the entrance into a dormant/ senescent state (Helfrich *et al.*, 2015). Dormancy enables isogenic, persisting cells to survive drug treatment by shutting down their metabolism and the antibiotic targets (Van den Bergh *et al.*, 2017). However, since no example of a prophage-induced dormancy has so far been reported and CGP3-induced cells were not able to re-grow (or form progeny under the chosen conditions), it was concluded that CGP3 activation causes a 'dead end' for *C. glutamicum* cells. In contrast, various studies proved that the SOS response is involved in forming persisters, especially upon treatment with DNA-damaging quinolones (Van den Bergh *et al.*, 2017). A carbon-limitation experiment revealed that CGP3 induction was only triggered in proliferating cells, whereas spontaneous activity of the SOS reporter was still detectable (Helfrich *et al.*, 2015). In agreement with the study of Pearl et al., it can be deduced from these experiments that persistent cells are protected from prophage induction (Pearl *et al.*, 2008) and it is speculated that *C. glutamicum* cells benefit from SOS-induced dormancy by avoiding CGP3 induction in cells that may recover. However, since no experimental data are available to prove this interesting hypothesis, the lower SPI rate observed under carbon starvation conditions, might also underline the tendency of temperate phages, to favor lysogeny over the lytic cycle under nutrient limitation (Williamson *et al.*, 2002, Wilson & Mann, 1997).

In 63.7% of CGP3-induced cells, the SOS response preceded prophage induction as expected. This analysis, however, evinced that in a significantly high fraction of phage positive cells (36.7%) phage induction was triggered in an SOS-independent manner (Helfrich *et al.*, 2015). The dependency of the SOS pathway is further underlined by the fact that prophage activity is strongly reduced in an SOS deficient strain (Δ*recA*) under MMC-inducing conditions (Helfrich *et al.*, 2015). This finding is in line with the 'classic' mechanism behind prophage induction, in which RecA is required to trigger an autocatalytic cleavage of a phage repressor protein (Casjens & Hendrix, 2015). Moreover, the presence of LexA binding sites in upstream regions of CGP3 genes and the upregulation of 47 CGP3-encoded genes in a Δ*lexA* mutant (Table 1) (Jochmann *et al.*, 2009) strongly indicate LexA to be directly involved in CGP3 repression. Since the absence of the two SOS key players directly affects CGP3 induction, it was concluded that the SOS response can indeed cause CGP3 induction.

However, the fact that prophage induction in a Δ*recA* mutant was not completely abolished and the occurrence of the high SOS-independent SPI fraction (Helfrich *et al.*, 2015)

confirmed that prophage induction works not only according to the classic SOS-dependent model. Interestingly, a comparable *E. coli* experiment was conducted by Little and Michalowski, who showed, that in a Δ*recA* (λ^+) strain, λ still switches to the lytic cycle, but at a strongly reduced and nearly undetectable rate (<10^{-8}·generation^{-1}, for comparison: *C. glutamicum* ~10^{-3}-10^{-4}·generation^{-1}) (Little & Michalowski, 2010). The authors speculated about the high lysogenic stability since SOS-dependent λ SPI occurs ~1000 times more often (Little *et al.*, 1999) and therefore there is no need for the high intrinsic stability unless it would maximize the overall reproduction rate (Little & Michalowski, 2010). In contrast, for CGP3 in *C. glutamicum* the rate of SOS-dependent and –independent SPI is nearly on the same scale, indicating the SOS-independent SPI to have a distinct physiological consequence. In a few studies, the RecA-alternative mechanisms were elucidated. For example, Gosh et al demonstrated that λ induction in a Δ*recA E. coli* lysogen is triggered by quorum sensing signaling molecules of a *P. aeruginosa* strain (Ghosh *et al.*, 2009). Furthermore, in another study, the selective inhibition of the termination factor, Rho, caused an upregulation of viral genes or activation of a prophage (Menouni *et al.*, 2013). Even the XS H-NS was described to repress the excision of the cryptic prophage Rac in *E. coli* (Hong *et al.*, 2010). For the *C. glutamicum* strain ATCC 13032, induction of CGP3 in a cell-density or a Rho-dependent manner have so far not been reported, but the influence of an XS protein is of particular interest and is therefore discussed in the following section.

4.3 Virus-like elements are silenced by small NAPs

To identify regulatory proteins that are involved in CGP3 induction, DNA affinity chromatographies with the promoter region of the early phage promoter of the operon *alpAC* (Donovan *et al.*, 2015) were performed. By this means, the phage-encoded protein CgpS was strongly enriched (Pfeifer *et al.*, 2016). BLAST analysis revealed this protein to be a homolog of the Lsr2 protein of *Mycobacterium tuberculosis*, which in turn is an XS protein analogous to H-NS of *E. coli* (Gordon *et al.*, 2008). As Gordon et al. demonstrated for Lsr2, CgpS is also able to complement a Δ*hns* phenotype in *E. coli* (Gordon *et al.*, 2008) emphasizing its functions as an XS protein.

Several studies reported that XS proteins bind to AT-rich genomic regions and repress their gene expression (Navarre *et al.*, 2006, Ding *et al.*, 2015, Smits & Grossman, 2010, Li *et al.*, 2009). To identify targets of CgpS, we performed ChAP-Seq analysis that resulted in a genome-wide binding profile. The major CgpS target, identified by this approach, was the CGP3 region (~20.5% of the 187 kb region) (Pfeifer *et al.*, 2016). In addition, a gene of the defective prophage

CGP1 was also bound, besides various other targets, and, moreover, all bound DNA regions showed a high AT-content (in comparison to the average AT content of the host genome).

Although XS proteins usually prefer an association to AT-rich regions, the respective mechanisms can vary between distinct groups of XS. For H-NS and Lsr2, an AT-Hook (Q/RGR) motif was identified to mediate binding to AT rich sequences (Ali *et al.*, 2012) and for MvaT in *Pseudomonas* species an AT-pincer is responsible (Ding *et al.*, 2015). Based on multiple sequence alignments of Lsr2- and CgpS-related proteins, a similar sequence 'RGI' in CgpS was found instead of the prokaryotic AT-Hook Q/RGR motif. Studies with Lsr2 and H-NS muteins showed that a single exchange of the AT-hook sequence, e.g., an arginine to alanine, only reduced DNA binding instead of abolishing it (Gordon *et al.*, 2011). Hence, if this variation is a feasible mechanism to adjust DNA binding of XS, this may explain why H-NS associates to 740 targets in *S.* Typhimurium (Navarre *et al.*, 2007), why Lsr2 binds to >800 genes in *M. tuberculosis* (Gordon *et al.*, 2010) and why CgpS binds to only 90 regions in *C. glutamicum* (Pfeifer *et al.*, 2016).

The related mode of action, a similar domain organization and the presence of conserved amino acids (Pfeifer *et al.*, 2016, Gordon *et al.*, 2008, Navarre *et al.*, 2007) led to the conclusion that also a common mechanism to deactivate XS protein can be postulated in order to knockout the function of CgpS. Inspired by the work of Williamson and Free, who characterized H-NS-related proteins to function as H-NS antagonists by preventing correct multimerization (Williamson & Free, 2005), a similar approach was applied. Hence, by using a truncated variant of CgpS, consisting of its N-terminal oligomerization domain, CGP3 was artificially induced (Pfeifer *et al.*, 2016). Remarkably, a related type of counteracting strategy was even found in viruses, such as phage T7 (Ali *et al.*, 2011). Transcriptomic analysis performed under these counter silencing conditions revealed an upregulation of viral genes from the CGP3 and CGP1 regions (Pfeifer *et al.*, 2016) underlining the CgpS-mediated repression of cryptic prophages in *C. glutamicum*.

The results of this thesis emphasize the importance of XS proteins, such as CgpS, for the control of cryptic and degenerated prophage islands to regulate the expression of foreign genes. However, the natural triggers mediating the counteraction of CgpS are still unknown and will be addressed in further studies.

4.4 How does the molecular mechanism of CGP3 induction work?

From the viral or bacterial point of view prophage induction can be a detrimental and irreversible decision for which a stringent control is needed. While the SOS response is well-studied and known to be induced by DNA damage (Michel, 2005), various different mechanisms, including environmental stress (heat, osmotic, pH), regulatory proteins and alternative sigma factors, have been identified to counter act H-NS activity (Navarre, 2009), for example. However, in the current literature the SOS pathway was not mentioned to be involved directly in counteracting XS proteins and little has been reported about the context between XS and the SOS response.

For Lsr2, a protective role against ROS was described, which appeared to be based on its physical presence (Colangeli *et al.*, 2009). ROS are known to induce DNA damage and, thus, Lsr2 might be able to act as a sensor of DNA lesions. In this scenario, sufficient ssDNA or double-strand breaks would cause structural DNA changes, leading to the relief of the silenced nucleoprotein complexes. Similar releasing mechanisms are known for H-NS. In particular, it was shown that H-NS repression is relieved due to structural nucleoid alterations caused by temperature, osmolarity or pH variations (Navarre, 2009). Derived from these findings, it could be assumed that prophage induction might be triggered in an SOS- or rather RecA-independent manner. However, for the activation of CGP3, this suggested mechanism cannot be applied since it is unable to explain why the induction rate in Δ*recA* cells was not as high as in the wild type strain, when triggered by the DNA damaging agent MMC (Helfrich *et al.*, 2015).

An alternative option would call on the action of the SOS repressor LexA since SOS boxes are located upstream of prophage genes and over 47 CGP3 genes are upregulated in a Δ*lexA* mutant (Table 1) (Jochmann *et al.*, 2009). Based on these data and because ~85% of the ORFs in CGP3 are annotated as hypothetical or putative proteins, two basic mechanisms for CGP3 induction can be assumed. In the two cases, LexA binds and permanently represses prophage genes as was also found for phage GIL01 in *Bacillus thuringiensis*, for example (Fornelos *et al.*, 2011). Under conditions of DNA damage, activated RecA* protein triggers the autocatalytic cleavage of LexA and causes de-repression of phage genes. One option could be that the expression of ‘early’ phage genes triggers the progression of prophage induction independently from silencing proteins. A second possibility includes a counteractor, in which the respective gene is under the control of LexA and is assumed to be encoded on the prophage region. The counteractor mediates the release of CgpS repression, leading to CGP3 induction. The latter option is supported by several H-NS studies in which various regulatory proteins, e.g., Ler, LeuO, RovA and SlyA, were identified to counteract the H-NS protein (Stoebel *et al.*, 2008). Furthermore, as it was shown for the coliphage 186 and its protein Tum (Shearwin *et al.*, 1998), in the CGP3 region, a potential ‘antirepressor’ might be encoded, which gene is under the

control of LexA. However, since most of the LexA-regulated CGP3 genes encode hypothetical proteins (Table 1) and none of them seem to share similarities with regulatory proteins, it is therefore difficult to draw a conclusion or to find potential candidates for further studies. However, one exception might be represented by cg2064, which is an annotated protein that harbors high similarities to a DNA topoisomerase type IA (Topo IA) (IPR023405). Topo IA enzymes relax positively or negatively supercoiled DNA and require ATP, single-stranded DNA and metal ions for their function (crucial for replication, transcription, and recombination) (Champoux, 2002, Champoux, 2001). Interestingly, in one *E. coli* study, it was shown that the phenotype of a mutant lacking the Topo IA could be suppressed by an *hns*::Tn5 mutation, suggesting that Topo IA activity counteracts H-NS-mediated silencing in the wild type strain (Stewart *et al.*, 2005). Hence, the CGP3-encoded Topo IA might act in a similar way by counteracting CgpS. Another interesting hypothesis, that should be considered, includes the action of small, phage-encoded proteins (<10 kDa). This mechanism is based on a study, in which the ~5 kDa phage protein Mip inhibited the activity of MvaT, the *Pseudomonas* XS (Wagemans *et al.*, 2015). Hence, some of the small, CGP3-encoded and LexA-controlled proteins (with unknown function, e.g., cg1978) may act as counteracting proteins. Nonetheless, to confirm one of these hypotheses more experimental data are needed and therefore this mechanism will be addressed in future studies.

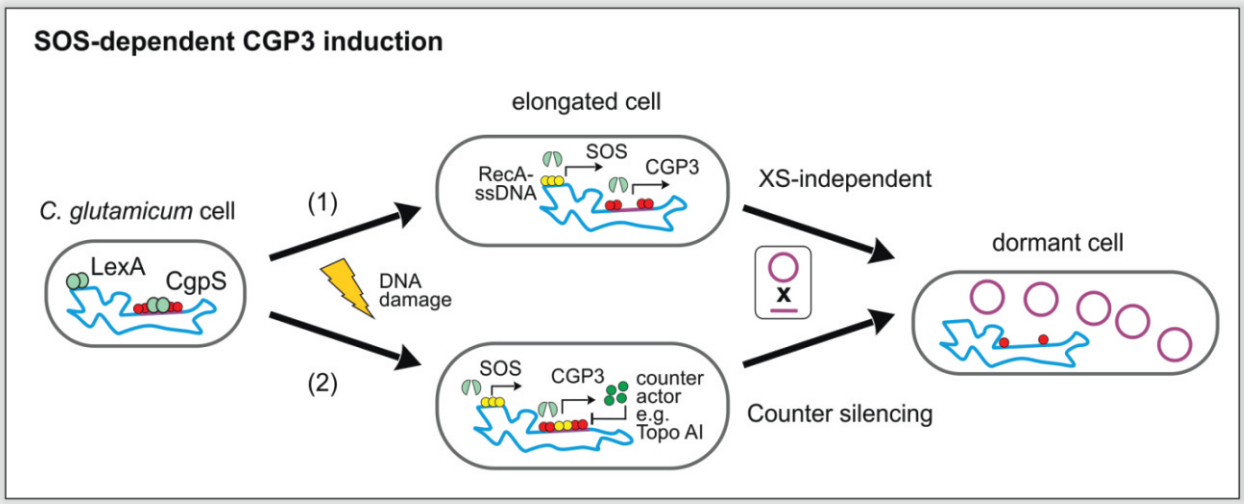


Figure 5: SOS-dependent activation of the cryptic prophage CGP3 in *C. glutamicum*. According to this model, LexA and CgpS bind and permanently repress expression of CPG3 genes. If ssDNA is formed, e.g., under DNA damaging conditions, the ubiquitous present protein RecA binds the ssDNA (RecA-ssDNA) and causes the autocatalytic cleavage of LexA. At this step two ways are conceivable to induce CGP3, either independently triggered of CgpS or due to counter silencing. (1) After LexA self-digestion, CGP3 excises itself from the genome and proceeds in the induction process without the need to overcome CgpS silencing. (2) After LexA cleavage, expression of a LexA-controlled counter actor gene is triggered and leads to CGP3 induction. A feasible candidate is cg2064 encoding a Topo IA enzyme that is described in the case of an *E. coli* to counteract the XS protein H-NS (Stewart *et al.*, 2005).

Table 1: List of CGP3 encoded genes that are upregulated in a *C. glutamicum* Δ*lexA* mutant. Experimental data were taken from Jochman et al. and Coryneregnet (Baumbach, 2007, Jochmann *et al.*, 2009).

Locus	Annotation	Locus	Annotation
cg1895	putative secreted protein	cg1993	hypothetical protein
cg1896	putative secreted protein	cg1996	Modification methylase
cg1897	putative secreted protein	cg1999	hypothetical protein
cg1904	putative membrane protein	cg2004	Similar to 232 protein-L*actobacillus* bacteriophage g1e
cg1911	putative secreted protein	cg2007	putative membrane protein
cg1917	hypothetical protein	cg2008	putative membrane protein
cg1918	putative secreted protein	cg2009	putative Clp-family ATP-binding-protease
cg1930	putative secreted hydrolase	cg2010	Permease of the major facilitator superfamily
cg1931	putative secreted protein	cg2011	putative membrane protein
cg1937	putative secreted protein	cg2014	hypothetical protein
cg1940	putative secreted protein	cg2016	hypothetical protein
cg1949	hypothetical protein	cg2017	hypothetical protein
cg1966	hypothetical protein	cg2018	putative membrane protein
cg1969	hypothetical protein	cg2019	putative membrane protein
cg1970	hypothetical protein	cg2020	putative membrane protein
cg1971	hypothetical protein predicted by Glimmer	cg2022	putative secreted protein
cg1974	putative lysin	cg2023	putative membrane protein
cg1975	hypothetical protein	cg2030	hypothetical protein predicted by Glimmer
cg1977	putative secreted protein	cg2032	putative membrane protein
cg1978	hypothetical protein	cg2034	hypothetical protein
cg1980	MoxR-like ATPase	cg2060	hypothetical protein
cg1981	hypothetical protein	cg2062	Similar to plasmid encoded-protein PXO2.09
cg1983	hypothetical protein	cg2063	putative membrane protein
cg1992	hypothetical protein predicted by Glimmer	cg2064	DNA topoisomerase I (omega protein)

The introduction of a single copy of *cgpS* (with its native promoter region) into different independent genomic loci was sufficient enough to significantly reduce the CGP3 SPI rate (data not shown). Furthermore, counter silencing studies revealed that the SPI rate was directly increased if a higher expression level of the truncated CgpS variant was used (data not shown). In addition, counteracting CgpS in an SOS-deficient strain (Δ*recA*) resulted in the same CGP3

induction rate as measured for the wild type (unpublished data, personally communicated with Hünnefeld, M). Even transcriptome analysis performed under the same conditions demonstrated that counter silencing has no direct influence on SOS genes (Pfeifer *et al.*, 2016). In contrast, using MMC to induce CGP3 in a Δ*recA* strain evinced almost no CGP3 activity (Helfrich *et al.*, 2015).

Thus, from these experiments, it can be concluded that (i) prophage induction works differently than has been described for the *E. coli*-λ system; (ii) besides the SOS response, further mechanisms feed into the activation of CGP3; and (iii) XS proteins, such as CgpS, affect the rate of SPI and may contribute to an adjustment of the SPI frequency in dependency of its physiological function.

4.5 Maintenance and impact of CGP3 on *C. glutamicum* populations

The cryptic prophage CGP3 constitutes 6% of the entire genome of *C. glutamicum* (Frunzke *et al.*, 2008, Kalinowski *et al.*, 2003). Efforts to delete this prophage region to generate a stable strain for metabolic engineering were successful if the CGP3-encoded restriction modification (RM) system was inactivated (Baumgart *et al.*, 2013). Thus, the question arose as to why and how CGP3 is maintained by this RM system. To explain this mechanism, it is suggested that the RM system would turn against *C. glutamicum* cells lacking the prophage region. In particular, after CGP3 deletion, newly synthesized DNA will no longer be methylated in the respective pattern, but the cytosol-located RM endonucleases are still intact (Figure 6). Hence, these nucleases cleave the unmodified DNA, thereby killing the daughter cells. This maintenance theory is based on the studies of Schäfer et al., who first characterized and proved the strong selectivity of the phage-encoded RM system against foreign DNA (Schäfer *et al.*, 1994). It was shown by intergeneric conjugation experiments that a *C. glutamicum* strain which contains more copies of the respective RM genes is no longer able to form transconjugants and, in addition, the conjugation efficiencies in a RM-deficient strain were strongly increased (Schäfer *et al.*, 1994, Schäfer *et al.*, 1990). Hence, the existence of CGP3 within *C. glutamicum* is therefore directly linked to the presence of RM-encoding genes (Baumgart *et al.*, 2013), which represents a persuasive explanation as to why the CGP3 region has been maintained in the ATCC 13032 strain.

Interestingly, similar reasons for the protection against foreign DNA and the sustainability of CGP3 are given by the CgpS-based silencing system. A maintenance principle can be derived from *cgpS* deletion experiments that resulted in CGP3-free colonies (Figure 6) (Pfeifer *et al.*, 2016). Since the counteraction of CgpS caused prophage induction and severe growth defects,

it was assumed that a *cgpS* knockout could only be compensated by recombination events leading to the loss of CGP3. Consistent with these assumptions, data from a *P. aeruginosa* study showed that null mutants of the XS *mvaT* and *mvaU* (members of H-NS family) were not viable unless the responsible phage genes were mutated (Castang & Dove, 2012). Moreover, a deletion of *hns* in *E. coli* resulted in strong growth defects (Yamada *et al.*, 1991), whereas Δ*hns Salmonella* Typhimurium strains needed to be counteracted by mutations in *rpoS* (general stress response) or in *phoP* (virulence gene regulator). Even more compensatory mutations in *Salmonella* were identified by an evolutional approach in *stpA*, encoding an H-NS paralog (Singh *et al.*, 2016). Because of this data, it can be concluded that XS proteins, such as CgpS, play crucial roles in the maintenance of a stable co-existence between foreign genetic elements and their host organisms.

An *E. coli* study presented by Wang et al. emphasized the beneficial impacts of nine cryptic prophages, since the loss of the entire proviral DNA (3.6% in *E. coli* K-12 BW251113) caused reduced fitness and less tolerance against osmotic and antibiotic stress (Wang *et al.*, 2010). To identify the respective conditions, comparative phenotypic microarrays were performed (Wang *et al.*, 2010). In a similar approach, the *C. glutamicum* wild type strain was compared with the MB001 strain in >1100 different conditions (covering different C-, N-, P-sources, antibiotics, osmotic, pH, etc.), however, the results have not revealed significant and reproducible differences. Under specific antibiotic concentrations such as for tetracycline, penimepicycline and rolitetracycline, the wild type surpassed the MB001 slightly (see appendix 6.2, PM12 & PM13). However, after several attempts (variation of different media, antibiotic concentrations, etc.), the competitive benefits were not reproducible (data not shown).

Co-cultivation experiments conducted with *C. glutamicum* wild type cells and the prophage-free variant MB001 revealed a slight growth advantage for the prophage-free strain on glucose minimal medium (Pfeifer *et al.*, 2017). Under MMC-inducing conditions, an even higher fitness advantage can be observed (Baumgart *et al.*, 2013). Similar results were obtained in genome reduction studies, where proviral elements were removed from the genomes of *P. aeruginosa* KT2440 and *Lactococcus lactis* NZ9000 (Zhu *et al.*, 2017, Martinez-Garcia *et al.*, 2015) (further interesting genome reduction examples are summarized in the review of Martinez-Garcia & de Lorenzo (Martinez-Garcia & de Lorenzo, 2016)). Improved fitness was obtained for the prophage-free variant of the *P. aeruginosa* KT2440 strain (under stress-inducing conditions; UV light) (Martinez-Garcia *et al.*, 2015) and *L. lactis* NZ9000 cells, which lacked four of the five prophages and showed directly enhanced growth in standard minimal media (Zhu *et al.*, 2017). Interestingly, the deletion of the last prophage region in the study by Zhu et al, was not possible (although not discussed in the article), resembling the initial problems with CGP3 (Baumgart *et al.*, 2013). However, improvements in growth under laboratory-defined or prophage-inducing

conditions are quite reasonable since the loss of dispensable DNA enables a faster replication. In addition, a prophage-free population is also spared of viral-associated toxic effects and cell death due to prophage induction.

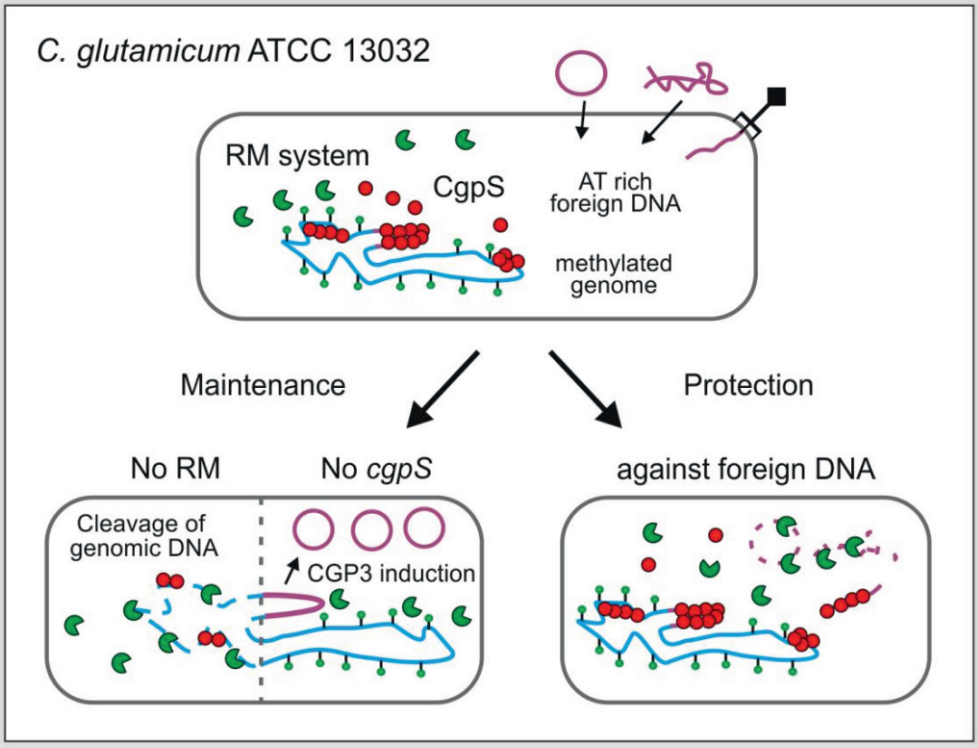


Figure 6: CGP3-encoded defense and maintenance systems. CgpS and the prophage encoded RM system are both described to protect *C. glutamicum* cell against foreign DNA by silencing or cleavage foreign DNA (Pfeifer *et al.*, 2016, Baumgart *et al.*, 2013, Schäfer *et al.*, 1990). Furthermore, the maintenance of CGP3 can be assigned to these two systems. Deletion of CGP3 was only possible, if the RM system was knocked out, suggesting that the remaining RM endonucleases would kill the daughter cells (since their genomes are incorrectly methylated) (Baumgart *et al.*, 2013). Deletion of *cgpS* was not possible without losing the CGP3 region (Pfeifer *et al.*, 2016). To explain this, it is suggested that, in the absence of CgpS, prophage induction leads to cell death and, therefore, only CGP3-free cells would form colonies.

Mobile elements such as prophages, are described to potentially influence bacterial genomes by rearranging the genomic composition due to insertions, disruptions, inversion, etc. (Darmon & Leach, 2014). Furthermore, several studies reported that H-NS promotes transpositions of mobile elements, especially found for Tn10 (Ward *et al.*, 2007, Wardle *et al.*, 2005), and, therefore, it was assumed that CgpS might have an influence on the transposition rate within the *C. glutamicum* genome. Stable genome dynamics and constant growth represent the basic requirements for platform strains, particularly if they need to be used in evolution experiments. For this reason a long-term evolution experiment (LTEE) (>600 generations) was conducted, representing the first LTEE, in which a prophage-free strain was used (Pfeifer *et al.*,

2017). Beside the impact of viral DNA on *C. glutamicum*, in a further interesting question, it was examined whether all prophages, especially CGP3, were maintained or got lost over the course of the experiment. Genome sequencing of several biological replicates with respect to different time points revealed that all cryptic prophages remained in the wild type genome. In addition, no significant differences in the mutations rates or transposon insertions were observed (Pfeifer *et al.*, 2017).

Based on these data, it can be concluded that the large DNA burden given by the ~190 kb CGP3 region and the associated metabolic costs (replication, transcription, protein production etc.) are not sufficient to negatively influence wild type cells or to force them to get rid of the provirus. Furthermore, these results provide strong evidence that two predominant mechanisms are required to stably maintain the CGP3 island in the genome of the ATCC 13032 strain, the RM system and the XS CgpS (Figure 6). Remarkably, both are encoded on the prophage itself.

4.6 XS-derived interactions in regulatory circuits of bacteria and virus-like elements

Xenogeneic silencing represents an elegant option to protect the host organism against uncontrolled expression of foreign genes that foster the incorporation of exogenous genetic elements into host regulatory circuits (Huttener *et al.*, 2015, Navarre *et al.*, 2007, Singh *et al.*, 2016). Therefore, it is quite reasonable to assume that the XS protein, CgpS, might facilitate the introduction of network pieces from the CGP3 ancestor phage into the regulatory circuits of *C. glutamicum*. Thus, the question arises: What are these intersections? The answers are directly connected to the impact of CGP3 or rather to the benefits and detriments that are discussed in the section above. In this section, the main focus is on the CgpS-associated regulon.

The genome wide-binding profile of CgpS delivered a detailed overview of the potential targets in *C. glutamicum*. Beside the CGP3 prophage, further genetic regions such as CGP1, genes within the HGT-acquired LCG island or transposase encoding genes (cg1951, cg2600) were also identified. Nonetheless, XS proteins bind and regulate also AT-rich host genomic regions as it was implied by the couple of hundred H-NS/Lsr2 targets (Gordon *et al.*, 2010, Navarre *et al.*, 2006). A non-foreign, but interesting CgpS-bound region, is the ferritin encoding gene *ftn*. Comparative transcriptomics revealed its upregulation under counter-silencing conditions (Pfeifer *et al.*, 2016). Remarkably, Lsr2 of *M. tuberculosis* and H-NS from *E. coli* were also identified to repress ferritin-encoded genes (Kurthkoti *et al.*, 2015, Nandal *et al.*, 2010). A counteraction of the two XS proteins are achieved by iron-dependent transcriptional regulators

IdeR (*M. tuberculosis*) and Fur of *E. coli* (Kurthkoti *et al.*, 2015, Nandal *et al.*, 2010). In *C. glutamicum*, the global iron regulator DtxR (homolog of IdeR) might counteract CgpS in a related way at the *ftn* promoter, which could explain why *ftn* is activated when sufficient iron is available and repressed under iron limitations (Wennerhold & Bott, 2006). The aforementioned examples describe an interesting mechanism regarding how a classical transcriptional regulator such as Fur or IdeR, may function as an activator by counteracting XS activity. Furthermore, it is quite astonishing that the two analogous XS proteins, H-NS and Lsr2, encounter at the iron homeostasis with the two also analogous regulator proteins, Fur and IdeR. However, since the exact mechanism and reasons why Lsr2 is involved in alternating iron levels remains unclear (Kurthkoti *et al.*, 2015), further studies are needed to identify the physiological relevance of XS in the iron homeostasis, especially of CgpS.

The highest peak within the genome-wide analysis outside of the CGP3 region was found at the cg0150 genetic locus (Pfeifer *et al.*, 2016). This gene encodes a protein, predicted to contain a fido domain (filamentation induced by cAMP/death on curing) and is described as a part of a TA-system (Garcia-Pino *et al.*, 2008). Overexpression of this gene in a *C. glutamicum* strain containing a prophage reporter led to higher CGP3-reporter activity and caused severe grow defects (unpublished). Hence, these results underline that the regulation of cg0150 needs to be tightly controlled. However, under counter-silencing conditions expression of cg0150 was not affected (Pfeifer *et al.*, 2016), indicating a further, unknown regulator is involved in repressing the transcription. This is also in line with the finding that the expression of cg0150 is not altered in the prophage- and *cgpS*-free strain MB001. However, the question regarding the advantages driving CgpS in occupying this genetic locus remain unanswered.

At long last, it is noteworthy that, besides the crucial silencing of foreign, potential toxic genes, XS proteins also act as global regulators by binding and repressing the expression of genes encoding transcriptional regulators. This has been shown by ChAP-seq/ChIP-on-Chip studies for H-NS in *Salmonella*, Lsr2 in *M. tuberculosis* and is also the case for CgpS in *C. glutamicum* (Table 2) (Gordon *et al.*, 2010, Navarre *et al.*, 2006, Pfeifer *et al.*, 2016). In addition, several studies, conducted with *E. coli* and *Vibrio cholera*, reported on the involvement of H-NS in the regulation of conserved transcriptional factors controlling pathogenicity, biofilm formation, acid stress response and even entire life-cycles (Ayala *et al.*, 2017, Stratmann *et al.*, 2008, Breddermann & Schnetz, 2017).

Thus, the modulation of transcriptional regulators that control the expression of genes involved in various different pathways strongly emphasizes both the high complexity of XS-dependent systems and at how many points XS proteins are possibly able to adjust expression levels. To add to this complexity, several copies of XS-like paralogues can be encoded on the same genome and these interactions may change the modulation of the respective genes.

Prominent examples are H-NS and StpA studied in *E. coli* (Müller *et al.*, 2006, Wolf *et al.*, 2006) or MvaT and MvaA from *Pseudomonas* species (Castang *et al.*, 2008, Tendeng *et al.*, 2003). A further noteworthy point to mention is that, besides the *cgpS* gene, no further Lsr2-like protein has been reported for *C. glutamicum* ATCC 13032, emphasizing this strain to be a suitable model for further investigations regarding the impact of phage-encoded XS proteins.

Table 2: CgpS bound genes encoding transcriptional regulators that are located outside of the CGP3 region (including elements of two-component systems).

Gene	Locus	Annotation
-	cg0725	putative transcriptional regulator, MarR-family
arnR	cg 1340	transcriptional regulator of *narKGHJI* and *hmp*
-	cg2462	putative transcriptional regulator, TetR-family
gntR1	cg2783	gluconate-responsive repressor 1, repressor of genes involved in gluconate catabolism and the pentose phosphate pathway, GntR-family
cgtS6	cg3060	two component sensor kinase

4.7 A step into the unknown – CgpS orthologues on phage genomes

A PSI-BLAST search revealed >600 CgpS orthologues, mainly encoded in Actinobacteria, which are predicted to harbor highly conserved structural similarities. (Pfeifer *et al.*, 2016). The genes were not only encoded on bacterial genomes, but were also present in 57 phage and 21 predicted prophage genomes. However, a closer examination of the provirus-encoded orthologues enlightened that not all respective genes are located within the predicted prophage regions, but seven of them instead are encoded on plasmids (Table 3). This incorrect allocation can be explained by the frequent mistakes that are made by prophage prediction tools due to very similar criteria of mobile elements. In addition, several examples of plasmids encoding XS proteins have already been discussed in the current literature (Doyle *et al.*, 2007, Yun *et al.*, 2015, Lang & Johnson, 2016). In particular, these plasmids are shown to harness XS proteins that could infiltrate new bacterial host cells, causing only a minimal impact on fitness, but driving a smooth horizontal transmission of genetic information (Doyle *et al.*, 2007). A further interesting aspect was given by Yun et al., who showed by comparative ChAP-Chip and transcriptome analysis that although the binding regions of plasmid and host-encoded XS proteins are very similar, their regulons differ, especially in the early stationary phase (Yun *et al.*, 2015). This finding again emphasizes the multi-faceted roles of XS proteins.

Table 3: 21 orthologues of CgpS (CgpS included) identified by a PSI-BLAST search (e-value ≤ 0.005) that are predicted to be encoded on prophages by PhySpy (Pfeifer *et al.*, 2016) or rather on plasmids.

Species	Annotation	Encoded on	Plasmid
Propionibacterium acnes TypeIA2 P.acn17	hypothetical protein TIA2EST22_01415	prophage	-
Propionibacterium acnes 266	putative Lsr2-like protein	prophage	-
Propionibacterium acnes HL096PA1	hypothetical protein PAGK_0304	prophage	unnamed, CP003294.1
Corynebacterium glutamicum ATCC 13032	CgpS, Lsr2-like protein	prophage	-
Tsukamurella paurometabola DSM 20162	protein Lsr2	prophage	pTpau01
Micromonospora aurantiaca ATCC 27029	putative Lsr2-like protein	prophage	-
Streptomyces cattleya DSM 46488	Lsr2-like protein	prophage	pSCATT
Streptomyces cattleya DSM 46488	Protein Lsr2	prophage	pSCATT
Gordonia polyisoprenivorans VH2	Lsr2-like protein	prophage	P174
Streptomyces hygroscopicus s ubsp. jinggangensis 5008	Lsr2-like protein	prophage	pSHJG1, pSHJG2
Nocardia brasiliensis ATCC 700358	Lsr2 protein	prophage	-
Nocardia brasiliensis ATCC 700358	Lsr2 protein	prophage	-
Frankia sp. EAN1pec	putative lysyl tRNA synthetase-like protein	prophage	-
Rhodococcus erythropolis PR4	conserved hypothetical protein	prophage	pREC1, pREC2, pREL1
Rhodococcus jostii RHA1	hypothetical protein RHA1_ro08754	plasmid, pRHL1	pRHL1, pRHL2, pRHL3
Rhodococcus jostii RHA1	lysyl tRNA synthetase-like protein	plasmid, pRHL3	pRHL1, pRHL2, pRHL3
Mycobacterium gilvum PYR-GCK	hypothetical protein Mflv_5452	plasmid, pMFLV01	Yes
Clavibacter michiganensis subsp. *michiganensis* NCPPB 382	hypothetical protein pCM2_0036	plasmid, pCM2	Yes
Rhodococcus opacus B4	hypothetical protein ROP_pROB01-05070	plasmid, pROB01	Yes
Rhodococcus opacus B4*b*	putative Lsr2-like protein	plasmid, pKNR	Yes
Verrucosispora maris AB-18-032	hypothetical protein VAB18032_30199	plasmid, pVMKU	Yes

Remarkably, most of the bacterial strains listed in Table 3 contain up to three plasmids, indicating a strong need for silencing proteins to presumably minimize plasmid-associated fitness costs. Furthermore, some of these strains, in particular, *Nocardia brasiliensis*, *Streptomyces cattleya* and *Rhodococcus jostii* and *R. opacus* harbor more than one gene encoding CgpS/ Lsr2-like proteins: However, the respective genes are never found on a single mobile element (prophage region or plasmid).

Unfortunately, several phage genomes have been uploaded more than one time under different NCBI accession numbers leading, to a false positive number of CgpS orthologues identified by the PSI-BLAST search (Pfeifer *et al.*, 2016). In a recent approach, a protein-protein BLAST analysis was performed with the CgpS sequence against the recently published database of actinobacteriophages (provided by the Hatfull group) (Russell & Hatfull, 2017) and led to over 224 hits (311 for Lsr2) in the uploaded phage genomes (Figure 7). Most of the CgpS- (and Lsr2-like) proteins are located on phages infecting the *Mycobacterium* species. This is because ~70% of the sequenced genomes are represented by mycobacteriophages (0.4% are corynebacteriophages).

To test whether CgpS/Lsr2 homologs are more frequent in virulent or temperate phages, ~2000 phages were classified (available in the database). They were allocated into the two respective groups according to the criteria that possible temperate phages encode at least one integrase, resolvase or recombinase. With this assignment, a set of 1581 virulent and 388 potential temperate phages were generated. Interestingly, approximately 51% of all CgpS-related proteins (~50% for Lsr2) were encoded on phage genomes of the larger, virulent group. By assuming a Gaussian-like phage distribution, it can be concluded that XS proteins, in particular, CgpS/Lsr2 orthologues, are four times more frequently encoded on (predicted) temperate rather than on virulent phages. Furthermore, since XS proteins are capable to repress proviral regions (Gordon *et al.*, 2010, Navarre *et al.*, 2006), one can assume that XS proteins might be able to replace repressor proteins of the respective prophages. Hence, overlaps of the temperate, XS-harboring phages were determined with genomes encoding and lacking repressor proteins (Figure 7). The comparison revealed that within the group of temperate viruses, phages that do not encode repressor proteins, but encode at least one CgpS orthologue are equally distributed as temperate phages that harbor both systems (similar distribution for Lsr2-like proteins).

In conclusion, distributions of CgpS and Lsr2 orthologues suggest that XS proteins are more favored by temperate phages. It can be hypothesized that XS proteins can even replace phage-repressing proteins. In addition, predicted virulent phages contain XS proteins, leading to the conclusion that XS proteins both enable a mutual adaption and presumably disturb the function of host-encoded XS proteins and other defense mechanisms.

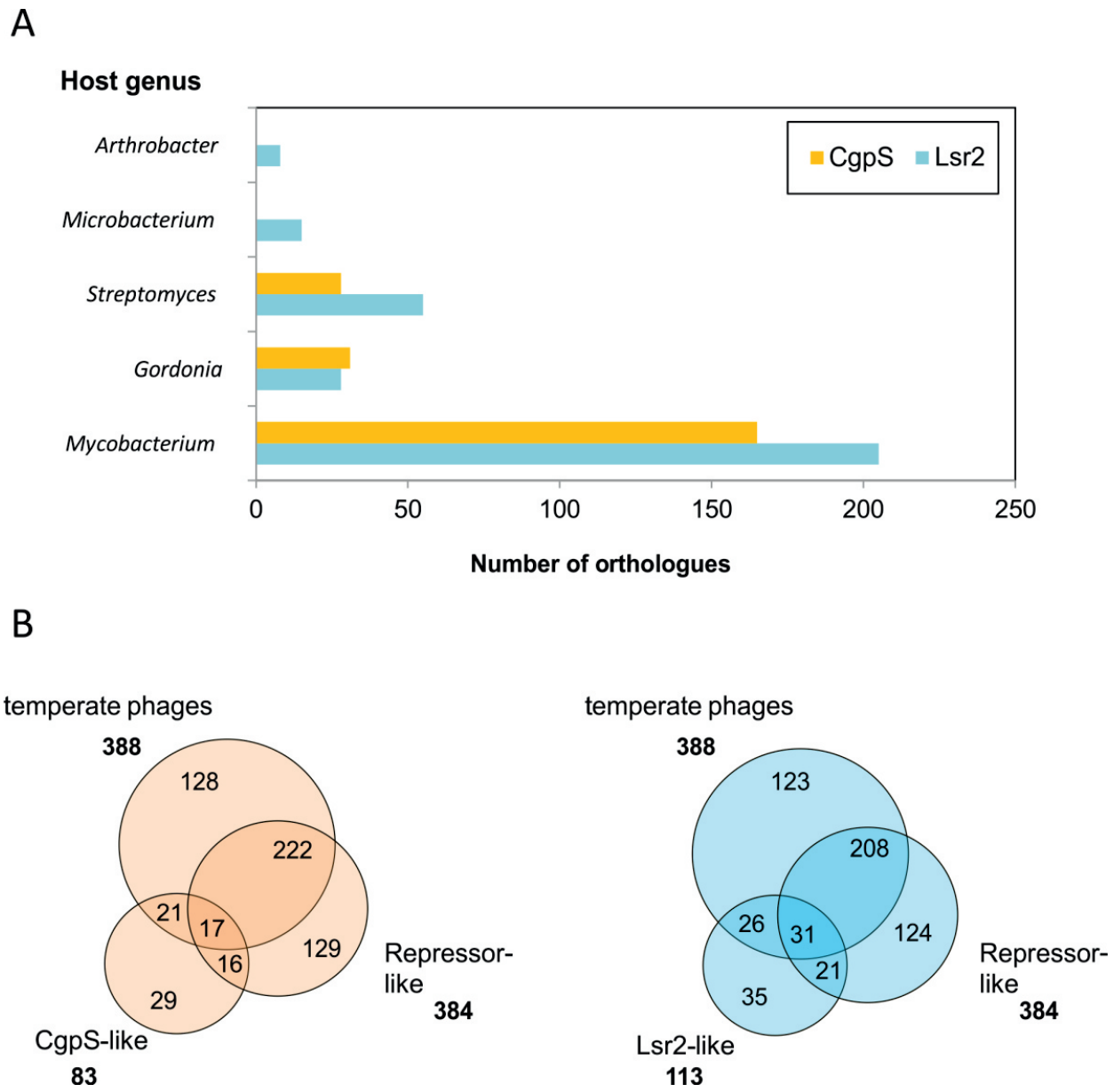


Figure 7: Distribution of CgpS and Lsr2 orthologues in actinobacteriophages. A: A BLASTp search (default option; BLOSUM62) with the amino acid sequences of CgpS and Lsr2 in the actinobacteriophage database (http://phagesdb.org/, ~10000 sequenced phages) (Russell & Hatfull, 2017) resulted in 224 orthologues for CgpS (orange) and 311 for Lsr2 of *M. tuberculosis* H37Rv (blue). Clustering was done according to host genus that is infected by the respective phages. **B.** Overlaps of phages encoding CgpS (#83) and Lsr2 (#113) orthologues with temperate (#388) and repressor (#384) harboring phages. Numbers of orthologues are less than mentioned in **A** since only 'phamerated' phage genomes (#1969) were available in the database (for comparing analysis). Temperate phages were defined as phages harboring at least an integrase, a recombinase or a resolvase.

This is, for example, supported by metagenomics data from Skennerton et al., who identified a phage-encoded H-NS-like protein with the potential to silence 6% of the hosts genome, including a CRISPR-Cas and an RM system (Skennerton *et al.*, 2011). Furthermore, in a bioinformatics study concerning the distribution of XS in prokaryotes, no different types of XS (e.g. H-NS- vs Lsr2-like) are found in the same species, suggesting an XS exclusion theory (Perez-Rueda & Ibarra, 2015). The hypothesis is further supported by the studies presented here, in which it was shown that the expression of the foreign XS gene *hns* in *C. glutamicum* caused induction of the prophage CGP3 and severe growth defects. Additionally, ChAP-Seq analysis revealed that CgpS binds two genetic regions encoding RM systems (Pfeifer *et al.*, 2016).

These data support that xenogeneic silencing is a widespread mechanism distributed not only among bacteria but also among phages (even plasmids), which underlines the close evolutionary connections between bacteria and phages shaped by horizontal gene transfer. Interestingly, the use of silencing systems in both domains seems to differ. While bacterial hosts use XS proteins as a rather mediative system allowing the step-wise acquisition of beneficial genes, (pro)-phages (and plasmids) harness XS to instead remain undetected in the host cell. This stealth-like strategy underlies the interests of prophages in providing a stable co-existence between bacteria and phages. In contrast, virulent phages harboring XS proteins have the potential to exploit the vulnerability of this system by disturbing host's defense mechanisms through the silencing of RM-encoding genes or the de-regulation of present XS targets. Therefore, this approach represents a further strategy in the arms-race between bacteria and phages.

5. References

Ali, S.S., E. Beckett, S.J. Bae & W.W. Navarre, (2011) The 5.5 protein of phage T7 inhibits H-NS through interactions with the central oligomerization domain. *J Bacteriol* **193**: 4881-4892.

Ali, S.S., B. Xia, J. Liu & W.W. Navarre, (2012) Silencing of foreign DNA in bacteria. *Curr Opin Microbiol* **15**: 175-181.

Asadulghani, M., Y. Ogura, T. Ooka, T. Itoh, A. Sawaguchi, A. Iguchi, K. Nakayama & T. Hayashi, (2009) The defective prophage pool of *Escherichia coli* O157: prophage-prophage interactions potentiate horizontal transfer of virulence determinants. *PLoS Pathog* **5**: e1000408.

Ayala, J.C., A.J. Silva & J.A. Benitez, (2017) H-NS: an overarching regulator of the *Vibrio cholerae* life cycle. *Res Microbiol* **168**: 16-25.

Baumbach, J., (2007) CoryneRegNet 4.0 - A reference database for corynebacterial gene regulatory networks. *BMC Bioinformatics* **8**: 429.

Baumgart, M., S. Unthan, C. Rückert, J. Sivalingam, A. Grünberger, J. Kalinowski, M. Bott, S. Noack & J. Frunzke, (2013) Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Appl Environ Microbiol* **79**: 6006-6015.

Bobay, L.M., M. Touchon & E.P. Rocha, (2014) Pervasive domestication of defective prophages by bacteria. *Proc Natl Acad Sci U S A* **111**: 12127-12132.

Breddermann, H. & K. Schnetz, (2017) Activation of *leuO* by LrhA in *Escherichia coli*. *Mol Microbiol*.

Browning, D.F., D.C. Grainger & S.J. Busby, (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Opin Microbiol* **13**: 773-780.

Brüssow, H., C. Canchaya & W.D. Hardt, (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* **68**: 560-602, table of contents.

Canchaya, C., F. Desiere, W.M. McShan, J.J. Ferretti, J. Parkhill & H. Brüssow, (2002) Genome analysis of an inducible prophage and prophage remnants integrated in the *Streptococcus pyogenes* strain SF370. *Virology* **302**: 245-258.

Canchaya, C., C. Proux, G. Fournous, A. Bruttin & H. Brüssow, (2003) Prophage genomics. *Microbiol Mol Biol Rev* **67**: 238-276.

Cardinale, C.J., R.S. Washburn, V.R. Tadigotla, L.M. Brown, M.E. Gottesman & E. Nudler, (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* **320**: 935-938.

Carrolo, M., M.J. Frias, F.R. Pinto, J. Melo-Cristino & M. Ramirez, (2010) Prophage spontaneous activation promotes DNA release enhancing biofilm formation in *Streptococcus pneumoniae*. *PLoS One* **5**: e15678.

Casjens, S., (2003) Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol* **49**: 277-300.

Casjens, S.R. & R.W. Hendrix, (2015) Bacteriophage lambda: Early pioneer and still relevant. *Virology* **479-480**: 310-330.

Castang, S. & S.L. Dove, (2012) Basis for the essentiality of H-NS family members in *Pseudomonas aeruginosa*. *J Bacteriol* **194**: 5101-5109.

Castang, S., H.R. McManus, K.H. Turner & S.L. Dove, (2008) H-NS family members function coordinately in an opportunistic pathogen. *Proc Natl Acad Sci U S A* **105**: 18947-18952.

Champoux, J.J., (2001) DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* **70**: 369-413.

Champoux, J.J., (2002) Type IA DNA topoisomerases: strictly one step at a time. *Proc Natl Acad Sci U S A* **99**: 11998-12000.

Colangeli, R., A. Haq, V.L. Arcus, E. Summers, R.S. Magliozzo, A. McBride, A.K. Mitra, M. Radjainia, A. Khajo, W.R. Jacobs, Jr., P. Salgame & D. Alland, (2009) The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proc Natl Acad Sci U S A* **106**: 4414-4418.

Colon, M.P., D. Chakraborty, Y. Pevzner & G.B. Koudelka, (2016) Mechanisms that Determine the Differential Stability of Stx(+) and Stx(-) Lysogens. *Toxins (Basel)* **8**: 96.

Darmon, E. & D.R. Leach, (2014) Bacterial genome instability. *Microbiol Mol Biol Rev* **78**: 1-39.

Daubin, V., E. Lerat & G. Perriere, (2003) The source of laterally transferred genes in bacterial genomes. *Genome Biol* **4**: R57.

Dillon, S.C. & C.J. Dorman, (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* **8**: 185-195.

Ding, P., K.A. McFarland, S. Jin, G. Tong, B. Duan, A. Yang, T.R. Hughes, J. Liu, S.L. Dove, W.W. Navarre & B. Xia, (2015) A Novel AT-Rich DNA Recognition Mechanism for Bacterial Xenogeneic Silencer MvaT. *PLoS Pathog* **11**: e1004967.

Donovan, C., A. Heyer, E. Pfeifer, T. Polen, A. Wittmann, R. Krämer, J. Frunzke & M. Bramkamp, (2015) A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria. *Nucleic Acids Res* **43**: 5002-5016.

Doyle, M., M. Fookes, A. Ivens, M.W. Mangan, J. Wain & C.J. Dorman, (2007) An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science* **315**: 251-252.

Engelberg-Kulka, H. & S. Kumar, (2015) Yet another way that phage lambda manipulates its *Escherichia coli* host: lambdarexB is involved in the lysogenic-lytic switch. *Mol Microbiol* **96**: 689-693.

Erental, A., Z. Kalderon, A. Saada, Y. Smith & H. Engelberg-Kulka, (2014) Apoptosis-like death, an extreme SOS response in *Escherichia coli*. *MBio* **5**: e01426-01414.

Feiner, R., T. Argov, L. Rabinovich, N. Sigal, I. Borovok & A.A. Herskovits, (2015) A new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nat Rev Microbiol* **13**: 641-650.

Fornelos, N., J.K. Bamford & J. Mahillon, (2011) Phage-borne factors and host LexA regulate the lytic switch in phage GIL01. *J Bacteriol* **193**: 6008-6019.

Fortier, L.C. & O. Sekulovic, (2013) Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence* **4**: 354-365.

Foster, P.L., (2007) Stress-induced mutagenesis in bacteria. *Crit Rev Biochem Mol Biol* **42**: 373-397.

Fothergill, J.L., E. Mowat, M.J. Walshaw, M.J. Ledson, C.E. James & C. Winstanley, (2011) Effect of antibiotic treatment on bacteriophage production by a cystic fibrosis epidemic strain of *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* **55**: 426-428.

Frunzke, J., M. Bramkamp, J.E. Schweitzer & M. Bott, (2008) Population Heterogeneity in *Corynebacterium glutamicum* ATCC 13032 caused by prophage CGP3. *J Bacteriol* **190**: 5111-5119.

Galhardo, R.S., P.J. Hastings & S.M. Rosenberg, (2007) Mutation as a stress response and the regulation of evolvability. *Critical Reviews in Biochemistry and Molecular Biology* **42**: 399-435.

Garcia-Pino, A., M. Christensen-Dalsgaard, L. Wyns, M. Yarmolinsky, R.D. Magnuson, K. Gerdes & R. Loris, (2008) Doc of prophage P1 is inhibited by its antitoxin partner Phd through fold complementation. *J Biol Chem* **283**: 30821-30827.

Ghosh, D., K. Roy, K.E. Williamson, S. Srinivasiah, K.E. Wommack & M. Radosevich, (2009) Acyl-homoserine lactones can induce virus production in lysogenic bacteria: an alternative paradigm for prophage induction. *Appl Environ Microbiol* **75**: 7142-7152.

Gordon, B.R., R. Imperial, L. Wang, W.W. Navarre & J. Liu, (2008) Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J Bacteriol* **190**: 7052-7059.

Gordon, B.R., Y. Li, A. Cote, M.T. Weirauch, P. Ding, T.R. Hughes, W.W. Navarre, B. Xia & J. Liu, (2011) Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc Natl Acad Sci U S A* **108**: 10690-10695.

Gordon, B.R., Y. Li, L. Wang, A. Sintsova, H. van Bakel, S. Tian, W.W. Navarre, B. Xia & J. Liu, (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **107**: 5154-5159.

Grünberger, A., C. Probst, S. Helfrich, A. Nanda, B. Stute, W. Wiechert, E. von Lieres, K. Nöh, J. Frunzke & D. Kohlheyer, (2015) Spatiotemporal microbial single-cell analysis using a high-throughput microfluidics cultivation platform. *Cytometry A* **87**: 1101-1115.

Hacker, J. & J.B. Kaper, (2000) Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**: 641-679.

Hatfull, G.F., (2014) Mycobacteriophages: windows into tuberculosis. *PLoS Pathog* **10**: e1003953.

Helfrich, S., E. Pfeifer, C. Krämer, C.C. Sachs, W. Wiechert, D. Kohlheyer, K. Nöh & J. Frunzke, (2015) Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations. *Mol Microbiol* **98**: 636-650.

Heyer, A., (2013) Characterization of a novel phage-encoded actin-like protein and the function of the ChrSA two-component system in *Corynebacterium glutamicum*. In. Research Center Juelich, Institute for Bio- and Geosciences (IBG)-1: Heinrich Heine Universität, pp. 113.

Higashi, K., T. Tobe, A. Kanai, E. Uyar, S. Ishikawa, Y. Suzuki, N. Ogasawara, K. Kurokawa & T. Oshima, (2016) H-NS Facilitates Sequence Diversification of Horizontally Transferred DNAs during Their Integration in Host Chromosomes. *PLoS Genet* **12**: e1005796.

Hong, S.H., X. Wang & T.K. Wood, (2010) Controlling biofilm formation, prophage excision and cell death by rewiring global regulator H-NS of *Escherichia coli*. *Microb Biotechnol* **3**: 344-356.

Howe, M.M. & E.G. Bade, (1975) Molecular biology of bacteriophage mu. *Science* **190**: 624-632.

Huttener, M., S. Paytubi & A. Juarez, (2015) Success in incorporating horizontally transferred genes: the H-NS protein. *Trends Microbiol* **23**: 67-69.

Ikeda, M. & S. Nakagawa, (2003) The *Corynebacterium glutamicum* genome: features and impacts on biotechnological processes. *Appl Microbiol Biotechnol* **62**: 99-109.

Janion, C., (2008) Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int J Biol Sci* **4**: 338-344.

Jochmann, N., A.K. Kurze, L.F. Czaja, K. Brinkrolf, I. Brune, A.T. Hüser, N. Hansmeier, A. Pühler, I. Borovok & A. Tauch, (2009) Genetic makeup of the *Corynebacterium glutamicum* LexA regulon deduced from comparative transcriptomics and in vitro DNA band shift assays. *Microbiology* **155**: 1459-1477.

Johnson, C.M. & A.D. Grossman, (2015) Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu Rev Genet* **49**: 577-601.

Juhas, M., J.R. van der Meer, M. Gaillard, R.M. Harding, D.W. Hood & D.W. Crook, (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* **33**: 376-393.

Kalinowski, J., B. Bathe, D. Bartels, N. Bischoff, M. Bott, A. Burkovski, N. Dusch, L. Eggeling, B.J. Eikmanns, L. Gaigalat, A. Goesmann, M. Hartmann, K. Huthmacher, R. Krämer, B. Linke, A.C. McHardy, F. Meyer, B. Mockel, W. Pfefferle, A. Pühler, D.A. Rey, C. Rückert, O. Rupp, H. Sahm, V.F. Wendisch, I. Wiegrabe & A. Tauch, (2003) The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J Biotechnol* **104**: 5-25.

Kamensek, S., Z. Podlesek, O. Gillor & D. Zgur-Bertok, (2010) Genes regulated by the *Escherichia coli* SOS repressor LexA exhibit heterogeneous expression. *BMC Microbiol* **10**: 283.

Kim, M. & S. Ryu, (2013) Antirepression system associated with the life cycle switch in the temperate podoviridae phage SPC32H. *J Virol* **87**: 11775-11786.

Kinoshita, S., S. Udaka & M. Shimono, (2004) Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms. *J Gen Appl Microbiol* **50**: 331-343.

Kneser, H., (1966) Repair of ultraviolet lesions and induction of lambda prophage. *Virology* **28**: 701-706.

Kurthkoti, K., P. Tare, R. Paitchowdhury, V.N. Gowthami, M.J. Garcia, R. Colangeli, D. Chatterji, V. Nagaraja & G.M. Rodriguez, (2015) The mycobacterial iron-dependent regulator IdeR induces ferritin (bfrB) by alleviating Lsr2 repression. *Mol Microbiol* **98**: 864-877.

Kuzminov, A., (1999) Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol Mol Biol Rev* **63**: 751-813.

L. Brown, T., J. Tucci, Z.A. Dyson, P. Lock, C.G. Adda & S. Petrovski, (2017) Dynamic interactions between prophages induce lysis in Propionibacterium acnes. *Research in Microbiology* **168**: 103-112.

Lang, K.S. & T.J. Johnson, (2016) Characterization of Acr2, an H-NS-like protein encoded on A/C2-type plasmids. *Plasmid* **87-88**: 17-27.

Lederberg, E.M. & J. Lederberg, (1953) Genetic Studies of Lysogenicity in *Escherichia Coli*. *Genetics* **38**: 51-64.

Li, C., H. Wally, S.J. Miller & C.D. Lu, (2009) The multifaceted proteins MvaT and MvaU, members of the H-NS family, control arginine metabolism, pyocyanin synthesis, and prophage activation in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* **191**: 6211-6218.

Little, J.W. & C.B. Michalowski, (2010) Stability and instability in the lysogenic state of phage lambda. *J Bacteriol* **192**: 6064-6076.

Little, J.W., D.P. Shepley & D.W. Wert, (1999) Robustness of a gene regulatory circuit. *EMBO J* **18**: 4299-4307.

Livny, J. & D.I. Friedman, (2004) Characterizing spontaneous induction of Stx encoding phages using a selectable reporter system. *Mol Microbiol* **51**: 1691-1704.

Lorenz, N., M. Reiger, M. Toro-Nahuelpan, A. Brachmann, L. Poettinger, L. Plener, J. Lassak & K. Jung, (2016) Identification and Initial Characterization of Prophages in *Vibrio campbellii*. *PLoS One* **11**: e0156010.

Lwoff, A., (1953) Lysogeny. *Bacteriol Rev* **17**: 269-337.

Makarova, K.S., Y.I. Wolf & E.V. Koonin, (2013) Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res* **41**: 4360-4377.

Martinez-Garcia, E. & V. de Lorenzo, (2016) The quest for the minimal bacterial genome. *Curr Opin Biotechnol* **42**: 216-224.

Martinez-Garcia, E., T. Jatsenko, M. Kivisaar & V. de Lorenzo, (2015) Freeing *Pseudomonas putida* KT2440 of its proviral load strengthens endurance to environmental stresses. *Environ Microbiol* **17**: 76-90.

McCool, J.D., E. Long, J.F. Petrosino, H.A. Sandler, S.M. Rosenberg & S.J. Sandler, (2004) Measurement of SOS expression in individual *Escherichia coli* K-12 cells using fluorescence microscopy. *Mol Microbiol* **53**: 1343-1357.

Menouni, R., S. Champ, L. Espinosa, M. Boudvillain & M. Ansaldi, (2013) Transcription termination controls prophage maintenance in *Escherichia coli* genomes. *Proc Natl Acad Sci U S A* **110**: 14414-14419.

Michel, B., (2005) After 30 years of study, the bacterial SOS response still surprises us. *PLoS Biol* **3**: e255.

Müller, C.M., U. Dobrindt, G. Nagy, L. Emody, B.E. Uhlin & J. Hacker, (2006) Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic *Escherichia coli*. *J Bacteriol* **188**: 5428-5438.

Mustard, J.A. & J.W. Little, (2000) Analysis of *Escherichia coli* RecA interactions with LexA, lambda CI, and UmuD by site-directed mutagenesis of *recA*. *J Bacteriol* **182**: 1659-1670.

Nanda, A.M., A. Heyer, C. Krämer, A. Grünberger, D. Kohlheyer & J. Frunzke, (2014) Analysis of SOS-induced spontaneous prophage induction in *Corynebacterium glutamicum* at the single-cell level. *J Bacteriol* **196**: 180-188.

Nanda, A.M., K. Thormann & J. Frunzke, (2015) Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *J Bacteriol* **197**: 410-419.

Nandal, A., C.C. Huggins, M.R. Woodhall, J. McHugh, F. Rodriguez-Quinones, M.A. Quail, J.R. Guest & S.C. Andrews, (2010) Induction of the ferritin gene (*ftnA*) of *Escherichia coli* by Fe(2+)-Fur is mediated by reversal of H-NS silencing and is RyhB independent. *Mol Microbiol* **75**: 637-657.

Navarre, W.W., (2009) H-NS as a defence system. In: Bacterial Chromatin. R.T. Dame & C.J. Dorman (eds). Springer Netherlands, pp.

Navarre, W.W., M. McClelland, S.J. Libby & F.C. Fang, (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev* **21**: 1456-1471.

Navarre, W.W., S. Porwollik, Y. Wang, M. McClelland, H. Rosen, S.J. Libby & F.C. Fang, (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* **313**: 236-238.

Ochman, H., J.G. Lawrence & E.A. Groisman, (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299-304.

Owen, S.V., N. Wenner, R. Canals, A. Makumi, D.L. Hammarlof, M.A. Gordon, A. Aertsen, N.A. Feasey & J.C. Hinton, (2017) Characterization of the Prophage Repertoire of African *Salmonella* Typhimurium ST313 Reveals High Levels of Spontaneous Induction of Novel Phage BTP1. *Front Microbiol* **8**: 235.

Pearl, S., C. Gabay, R. Kishony, A. Oppenheim & N.Q. Balaban, (2008) Nongenetic individuality in the host-phage interaction. *PLoS Biol* **6**: e120.

Pennington, J.M. & S.M. Rosenberg, (2007) Spontaneous DNA breakage in single living *Escherichia coli* cells. *Nat Genet* **39**: 797-802.

Perez-Rueda, E. & J.A. Ibarra, (2015) Distribution of putative xenogeneic silencers in prokaryote genomes. *Comput Biol Chem* **58**: 167-172.

Pfeifer, E., (2013) Untersuchungen zur Induktion und Infektiosität des Prophagen CGP3 in *Corynebacterium glutamicum*. In: Institute for Bio- and Geosciences 1. Jülich: Forschungszentrum Jülich, pp.

Pfeifer, E., C. Gätgens, T. Polen & J. Frunzke, (2017) Adaptive laboratory evolution of Corynebacterium glutamicum towards higher growth rates on glucose minimal medium. *Appl Environ Microbiol* **submitted**: 12.

Pfeifer, E., M. Hünnefeld, O. Popa, T. Polen, D. Kohlheyer, M. Baumgart & J. Frunzke, (2016) Silencing of cryptic prophages in *Corynebacterium glutamicum*. *Nucleic Acids Res* **44**: 10117-10131.

Rocha, E.P. & A. Danchin, (2002) Base composition bias might result from competition for metabolic resources. *Trends Genet* **18**: 291-294.

Rosenberg, S.M., C. Shee, R.L. Frisch & P.J. Hastings, (2012) Stress-induced mutation via DNA breaks in *Escherichia coli*: a molecular mechanism with implications for evolution and medicine. *Bioessays* **34**: 885-892.

Rossmann, F.S., T. Racek, D. Wobser, J. Puchalka, E.M. Rabener, M. Reiger, A.P. Hendrickx, A.K. Diederich, K. Jung, C. Klein & J. Huebner, (2015) Phage-mediated dispersal of biofilm and distribution of bacterial virulence genes is induced by quorum sensing. *PLoS Pathog* **11**: e1004653.

Rozanov, D.V., R. D'Ari & S.P. Sineoky, (1998) RecA-independent pathways of lambdoid prophage induction in *Escherichia coli*. *J Bacteriol* **180**: 6306-6315.

Russell, D.A. & G.F. Hatfull, (2017) PhagesDB: the actinobacteriophage database. *Bioinformatics* **33**: 784-786.

Schäfer, A., J. Kalinowski, R. Simon, A.H. Seep-Feldhaus & A. Pühler, (1990) High-frequency conjugal plasmid transfer from gram-negative *Escherichia coli* to various gram-positive coryneform bacteria. *J Bacteriol* **172**: 1663-1666.

Schäfer, A., A. Schwarzer, J. Kalinowski & A. Pühler, (1994) Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from *Corynebacterium glutamicum* ATCC 13032 and analysis of its role in intergeneric conjugation with *Escherichia coli*. *J Bacteriol* **176**: 7309-7319.

Sekulovic, O. & L.C. Fortier, (2016) Characterization of Functional Prophages in *Clostridium difficile*. *Methods Mol Biol* **1476**: 143-165.

Shearwin, K.E., A.M. Brumby & J.B. Egan, (1998) The Tum protein of coliphage 186 is an antirepressor. *J Biol Chem* **273**: 5708-5715.

Shimizu, T., Y. Ohta & M. Noda, (2009) Shiga Toxin 2 Is Specifically Released from Bacterial Cells by Two Different Mechanisms. *Infection and Immunity* **77**: 2813-2823.

Simmons, L.A., A.I. Goranov, H. Kobayashi, B.W. Davies, D.S. Yuan, A.D. Grossman & G.C. Walker, (2009) Comparison of responses to double-strand breaks between *Escherichia coli* and *Bacillus subtilis* reveals different requirements for SOS induction. *J Bacteriol* **191**: 1152-1161.

Singh, K., J.N. Milstein & W.W. Navarre, (2016) Xenogeneic Silencing and Its Impact on Bacterial Genomes. *Annu Rev Microbiol* **70**: 199-213.

Skennerton, C.T., F.E. Angly, M. Breitbart, L. Bragg, S. He, K.D. McMahon, P. Hugenholtz & G.W. Tyson, (2011) Phage encoded H-NS: a potential achilles heel in the bacterial defence system. *PLoS One* **6**: e20095.

Smits, W.K. & A.D. Grossman, (2010) The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in *Bacillus subtilis*. *PLoS Genet* **6**: e1001207.

Stevens, R.H., C.D.M.L. dos Santos, D. Zuanazzi, M.B.D. Mattos, D.F. Ferreira, S.C. Kachlany & E.M.B. Tinoco, (2013) Prophage induction in lysogenic Aggregatibacter actinomycetemcomitans cells co-cultured with human gingival fibroblasts, and its effect on leukotoxin release. *Microbial Pathogenesis* **54**: 54-59.

Stewart, N., J. Feng, X. Liu, D. Chaudhuri, J.W. Foster, M. Drolet & Y.C. Tse-Dinh, (2005) Loss of topoisomerase I function affects the RpoS-dependent and GAD systems of acid resistance in *Escherichia coli*. *Microbiology* **151**: 2783-2791.

Stoebel, D.M., A. Free & C.J. Dorman, (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154**: 2533-2545.

Stratmann, T., S. Madhusudan & K. Schnetz, (2008) Regulation of the *yjjQ-bglJ* operon, encoding LuxR-type transcription factors, and the divergent *yjjP* gene by H-NS and LeuO. *J Bacteriol* **190**: 926-935.

Tendeng, C., O.A. Soutourina, A. Danchin & P.N. Bertin, (2003) MvaT proteins in *Pseudomonas* spp.: a novel class of H-NS-like proteins. *Microbiology* **149**: 3047-3050.

Touchon, M., A. Bernheim & E.P. Rocha, (2016) Genetic and life-history traits associated with the distribution of prophages in bacteria. *ISME J* **10**: 2744-2754.

Van den Bergh, B., M. Fauvart & J. Michiels, (2017) Formation, physiology, ecology, evolution and clinical importance of bacterial persisters. *FEMS Microbiol Rev*.

Wagemans, J., A.S. Delattre, B. Uytterhoeven, J. De Smet, W. Cenens, A. Aertsen, P.J. Ceyssens & R. Lavigne, (2015) Antibacterial phage ORFans of Pseudomonas aeruginosa phage LUZ24 reveal a novel MvaT inhibiting protein. *Front Microbiol* **6**: 1242.

Wang, X., Y. Kim, Q. Ma, S.H. Hong, K. Pokusaeva, J.M. Sturino & T.K. Wood, (2010) Cryptic prophages help bacteria cope with adverse environments. *Nat Commun* **1**: 147.

Wang, X., Y. Kim & T.K. Wood, (2009) Control and benefits of CP4-57 prophage excision in *Escherichia coli* biofilms. *ISME J* **3**: 1164-1179.

Ward, C.M., S.J. Wardle, R.K. Singh & D.B. Haniford, (2007) The global regulator H-NS binds to two distinct classes of sites within the Tn10 transpososome to promote transposition. *Mol Microbiol* **64**: 1000-1013.

Wardle, S.J., M. O'Carroll, K.M. Derbyshire & D.B. Haniford, (2005) The global regulator H-NS acts directly on the transpososome to promote Tn10 transposition. *Genes Dev* **19**: 2224-2235.

Weinberger, A.D., Y.I. Wolf, A.E. Lobkovsky, M.S. Gilmore & E.V. Koonin, (2012) Viral diversity threshold for adaptive immunity in prokaryotes. *MBio* **3**: e00456-00412.

Wennerhold, J. & M. Bott, (2006) The DtxR regulon of *Corynebacterium glutamicum*. *J Bacteriol* **188**: 2907-2918.

Williamson, H.S. & A. Free, (2005) A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS antagonist. *Mol Microbiol* **55**: 808-827.

Williamson, S.J., L.A. Houchin, L. McDaniel & J.H. Paul, (2002) Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. *Applied and Environmental Microbiology* **68**: 4307-4314.

Wilson, W.H. & N.H. Mann, (1997) Lysogenic and lytic viral production in marine microbial communities. *Aquatic Microbial Ecology* **13**: 95-100.

Wolf, T., W. Janzen, C. Blum & K. Schnetz, (2006) Differential dependence of StpA on H-NS in autoregulation of *stpA* and in regulation of *bgl*. *J Bacteriol* **188**: 6728-6738.

Yamada, H., T. Yoshida, K. Tanaka, C. Sasakawa & T. Mizuno, (1991) Molecular analysis of the *Escherichia coli hns* gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. *Mol Gen Genet* **230**: 332-336.

Yun, C.S., Y. Takahashi, M. Shintani, T. Takeda, C. Suzuki-Minakuchi, K. Okada, H. Yamane & H. Nojiri, (2015) MvaT Family Proteins Encoded on IncP-7 Plasmid pCAR1 and the Host Chromosome Regulate the Host Transcriptome Cooperatively but Differently. *Appl Environ Microbiol* **82**: 832-842.

Zhu, D., Y. Fu, F. Liu, H. Xu, P.E. Saris & M. Qiao, (2017) Enhanced heterologous protein productivity by genome reduction in *Lactococcus lactis* NZ9000. *Microb Cell Fact* **16**: 1.

Zwir, I., W.S. Yeo, D. Shin, T. Latifi, H. Huang & E.A. Groisman, (2014) Bacterial nucleoid-associated protein uncouples transcription levels from transcription timing. *MBio* **5**: e01485-01414.

6. Appendix

6.1 A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria

Donovoan C.[1,2•], Heyer A.[3•], Pfeifer E.[3], Polen T.[3], Wittman A.[2], Krämer R.[2], Frunzke J*[3] and Bramkamp M.[1,2*].

[1]Department of Biology I, Ludwig-Maximilians-University Munich, Germany

[2]Institute for Biochemistry, University of Cologne, Germany

[3]IBG-1: Biotechnology, Forschungszentrum Jülich, Jülich, Germany

•These authors contributed equally to this work.

*Corresponding author

Name of Journal: *Nucleic Acids Research*

Impact factor: *9.202*

Author contributions

Own contribution to the work: 10 %

Own contribution to experiments, their evaluation and illustration

	Name	Experimental work	Evaluation	Processing & Preparation of figure/ table
Table S1	**Pfeifer E.**	10%	10%	10%
Figure S1	**Pfeifer E.**	100%	100%	100%

Nucleic Acids Research Advance Access published April 27, 2015

Nucleic Acids Research, 2015 **1**
doi: 10.1093/nar/gkv374

A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria

Catriona Donovan[1,2,†], Antonia Heyer[3,†], Eugen Pfeifer[3], Tino Polen[3], Anja Wittmann[2], Reinhard Krämer[2], Julia Frunzke[3,*] and Marc Bramkamp[1,2,*]

[1]Department of Biology I, Ludwig-Maximilians-University Munich, Großhaderner Str. 2–4, 82152 Planegg-Martinsried, Germany, [2]Institute for Biochemistry, University of Cologne, Zülpicherstr. 47, 50674 Cologne, Germany and [3]Institut für Bio- und Geowissenschaften, IBG-1: Biotechnologie, Forschungszentrum Jülich, D-52425 Jülich, Germany

Received January 30, 2015; Revised April 07, 2015; Accepted April 09, 2015

ABSTRACT

In host cells, viral replication is localized at specific subcellular sites. Viruses that infect eukaryotic and prokaryotic cells often use host-derived cytoskeletal structures, such as the actin skeleton, for intracellular positioning. Here, we describe that a prophage, CGP3, integrated into the genome of *Corynebacterium glutamicum* encodes an actin-like protein, AlpC. Biochemical characterization confirms that AlpC is a *bona fide* actin-like protein and cell biological analysis shows that AlpC forms filamentous structures upon prophage induction. The co-transcribed adaptor protein, AlpA, binds to a consensus sequence in the upstream promoter region of the *alpAC* operon and also interacts with AlpC, thus connecting circular phage DNA to the actin-like filaments. Transcriptome analysis revealed that *alpA* and *alpC* are among the early induced genes upon excision of the CGP3 prophage. Furthermore, qPCR analysis of mutant strains revealed that both AlpA and AlpC are required for efficient phage replication. Altogether, these data emphasize that AlpAC are crucial for the spatio-temporal organization of efficient viral replication. This is remarkably similar to actin-assisted membrane localization of eukaryotic viruses that use the actin cytoskeleton to concentrate virus particles at the egress sites and provides a link of evolutionary conserved interactions between intracellular virus transport and actin.

INTRODUCTION

DNA of viral origin, including fully functional prophages, cryptic (degenerated) prophage elements or phage morons, represents a common element of bacterial genomes (1,2). Upon prophage induction, e.g. triggered by the host SOS response, temperate phages enter the lytic pathway leading to the excision of the viral DNA from the genome, replication, virion assembly and lysis of the host cell (3,4). Recent reports revealed that viral replication in prokaryotes appears to be organized at specific intracellular locations and this process relies on the action of cytoskeletal proteins (5).

Cytoskeletal elements in eu- and prokaryotes are involved in a variety of different cellular processes (6,7). During mitotic segregation, eukaryotic chromosomes are moved by microtubules that attach to the centromere (6,8). In bacterial cells, mechanistically similar DNA segregation processes have been described (7,9–16). Best understood is the segregation of plasmid DNA via a tripartite *par*titioning system. The genetic organization of *par* loci is similar for both chromosome- and plasmid-encoded systems. In general, the *par* locus entails two *trans*-acting proteins encoded in an operon and *cis*-acting 'centromere-like' elements. The centromere-binding protein binds the centromere-like element forming a nucleoprotein complex (17–19). The segrosomes are recognized and segregated by the action of a partition protein, which, depending on the plasmid partitioning system, is either a Walker-A P-loop ATPase (ParA, Type I), an actin-like ATPase (ParM, Type II) or a tubulin-like GTPase (TubZ, Type III) (14,15,20–22). The precise partitioning mechanism remains under debate (23). Dynamic cytomotive filament pushing or pulling mechanisms have been supported for most actin and tubulin based segregation systems (21,24,25). Walker-A P-loop ATPase, on the other hand, have been proposed to employ a diffusion-ratchet model. In this model, the nucleotide bound from of the ATPase determines its affinity for the nucleoid. Combined with associated regulatory elements and intrinsic ATPase activity, gradients of the Walker-A ATPases form on the nucleoid. Rounds of nucleoid tethering and releasing lead to a directed movement of the DNA cargo (26,27). Active partitioning systems are imperative for efficient main-

*To whom correspondence should be addressed. Tel: +49 89 218074611; Fax: +49 89 218074621; Email: marc.bramkamp@lmu.de
Correspondence may also be addressed to Julia Frunzke. Tel: +49 2461 615430; Fax: +49 2461 612710; Email: j.frunzke@fz-juelich.de
†These authors contributed equally to the paper as first authors.

© The Author(s) 2015. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

tenance of low copy plasmids, such as P1 prophage from *Escherichia coli* (28–30).

Actin-like proteins also play a major role in cell growth and shape determination. MreB is the archetype of the bacterial cytoskeletal proteins (31–33). The MreB structure revealed homology to actin (34). MreB and its homologs are now known to be involved in the positioning of cell wall synthesizing complexes (35–38). Remarkably, MreB was also shown to be involved in viral replication. Replication of the *Bacillus subtilis* phage φ29 depends on the presence of all three MreB isoforms (MreB, Mbl and MreBH) (39–41). MreB apparently interacts with a phage-encoded membrane protein p16.7. Similarly, replication of the *E. coli* phage PRD1 is reduced upon inhibition of MreB. Recently, a tubulin-like protein, PhuZ, from bacteriophage 201φ2-1 was described to form spindle-like filaments *in vivo* thereby positioning the phage DNA within the cell (42,43).

In this study, we identified and characterized a novel actin-like protein encoded by the cryptic prophage CGP3 located in the genome of *Corynebacterium glutamicum* strain ATCC 13032. The genome of this important industrial platform organism harbors three cryptic prophages, CGP1–3, of which only CGP3 has been shown to replicate extra-chromosomally in a circularized form (44). CGP3 encompasses with 187 kb almost 6% of the entire *C. glutamicum* genome and belongs to the largest phage elements with known sequence (45). A cluster of tRNA genes is found on the left periphery of the CGP3 phage, while the right border encodes a putative phage integrase. The element is flanked by conserved attachment sites (44). Spontaneous induction of the CGP3 phage in a subpopulation of cells (1–3%) has been reported previously (44). However, CGP3 appeared to be inactive in terms of cell lysis and virion production and is therefore referred to as a cryptic prophage, which likely became trapped in the genome in the course of evolution.

Here, we describe that the first open reading frame in the CGP3 prophage encodes an actin-like protein, AlpC, and adjacent a phage DNA-binding protein, AlpA. Both AlpC and AlpA are necessary for efficient phage replication *in vivo*. Further, we show that AlpA binds a consensus sequence on the phage DNA molecules. The actin-like protein AlpC assembles into filaments that interact with the AlpA bound CGP3 DNA, which may function as a scaffold for the organization efficient viral replication. *Corynebacterium glutamicum* does not encode an MreB homlog, thus it would seem advantageous that the *C. glutamicum* CGP3 prophage encodes its own cytoskeletal element. Our data suggest that bacterial phages use an actin-based transport system, analogous to vertebrate viruses such as the herpesvirus which use host cell derived cytoskeletal elements (46,47).

MATERIALS AND METHODS

Recombinant DNA work

Standard methods like PCR, restriction or ligation were carried out according to established protocols (48,49). Oligonucleotide synthesis and DNA sequencing was performed by Eurofins MWG Operon (Ebersfeld, Germany). Strains, plasmids and oligonucleotides are listed in Supplementary Table S1. Strain construction is described in the Supplemental Materials and Methods.

Determination of circular phage DNA using quantitative PCR

The relative amount of circular phage DNA was determined *via* quantitative PCR (qPCR). Therefore, *C. glutamicum* wild type, the *alpC* deletion strain, and the *alpA* deletion strain were grown in 5 ml BHI (Brain Heart Infusion, Difco) for about 6 h at 30°C. A second precultivation was performed in CGXII minimal medium containing 4% glucose as carbon source. From each preculture two main cultures were inoculated to an OD_{600} of 1 in CGXII minimal medium. At an OD_{600} of 3 mitomycin C (final concentration of 0.6 μM) was added to one culture to induce the SOS response; the second, untreated culture served as reference. Five milliliters samples were harvested by centrifugation (4000 x *g*, 10 min and 4°C) at different time points after induction (0–24 h). The preparation of genomic DNA was performed as described previously (50).

The determination of the relative amount of circular DNA of the phage CGP3 by quantitative PCR was conducted according to the protocol described in Frunzke *et al.* (44). Briefly, each sample contained 1 μg total DNA as a template for amplification. Amplification was performed using the DyNAmo Capillary SYBR Green qPCR Kit (Finnzymes Oy, Vantaa) and a LightCycler type 1.0 (Roche Diagnostics). The *ddh* gene (cg2900), which is present in one copy in the *C. glutamicum* genome, served as reference gene for normalization; the oligonucleotides ddh-LC-for and ddh-LC-rev were used for amplification. For the detection of circular CGP3 DNA, oligonucleotides (Phage-LC-for and Phage-LC-rev) were designed which anneal to the left and right boarder of CGP3, pointing into divergent directions. This arrangement was used to specifically amplify a 150 bp DNA fragment covering the attachment site (*attP*) of the excised circular CGP3 DNA molecule. For the calculation of the relative amount of the circular phage DNA, the amount of the circular phage DNA in induced cells was normalized with the amount of circular phage DNA in uninduced cells of the same strain at each time point.

DNA microarrays

For transcriptome analysis cells were cultivated as described for the determination of circular CGP3 DNA (previous paragraph). Transcriptomes of *C. glutamicum* ATCC 13032 were comparisons of cells treated with or without 0.6 mM mitomycin C after 1, 3, 6 and 9 h after addition of mitomycin C. The custom DNA microarrays were obtained from Agilent Technologies (Waldbronn, Germany). Agilent's eArray platform was used to design oligonucleotide probes with the best probe methodology and assemble custom 4×44K 60mer microarray designs (https://earray.chem.agilent.com/earray/). The custom design included oligonucleotides for the annotated protein-coding genes and structural RNA genes of the four bacterial genomes from *C. glutamicum*, *Escherichia coli*, *Gluconobacter oxydans* and *B. subtilis* for genome-wide gene expression analysis. For *C. glutamicum*, the genome annotation NC_006958 from NCBI was used listing 3057 protein coding genes and 80 structural tRNA and ribosomal RNA genes (45). In the custom design, *C. glutamicum* genes are represented by one, two or three oligonucleotides which were used to

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

determine relative RNA levels. The custom array design also included the Agilent's control spots. Purified cDNA samples to be compared were pooled and prepared two-color samples were hybridized on 4×44K arrays at 65°C for 17 h using Agilent's gene expression hybridization kit, Agilent's hybridization chamber and Agilent's hybridization oven. After hybridization the arrays were washed using Agilent's wash buffer kit according to the manufacturer's instructions. Fluorescence of hybridized DNA microarrays was determined at 532 nm (Cy3-dUTP) and 635 nm (Cy5-dUTP) at 5 μm resolution with a GenePix 4000B laser scanner and GenePix Pro 6.0 software (Molecular Devices, Sunnyvale, CA, USA). Fluorescence images were saved to raw data files in TIFF format (GenePix Pro 6.0). Quantitative TIFF image analysis was carried out using GenePix image analysis software and the Agilent's gene array list (GAL) file. The results were saved as GPR-file (GenePix Pro 6.0). For background correction of spot intensities, ratio calculation and ratio normalization, GPR-files were processed using the BioConductor R-packages limma and marray (http://www.bioconductor.org). For further analysis, the processed and lowess-normalized data as well as detailed experimental information according to MIAME (51) were stored in the in house DNA microarray database (52). To search the data for differentially expressed genes by the processed Cy5/Cy3 ratio reflecting the relative RNA level, the criteria flags ≥0 (GenePix Pro 6.0) and signal/noise ≥5 for Cy5 (F635Median/B635Median) or Cy3 (F532Median/B532Median) were used. Array data were deposited in the GEO database (ncbi.nlm.nih.gov/geo) under accession number GSE45907.

Heterologous protein expression and purification

His_{10}-AlpC was heterogolously overproduced in *E. coli* BL21 (DE3) pLysS. Expression was induced with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were collected by centrifugation at 5000 × *g* (4°C) for 10 min, resuspended in buffer A (100 mM Tris/HCl, pH 7.5, 150 mM KCl, 150 mM NaCl, 10% glycerol and 10 mM imidazole), supplemented with DNase I and protease inhibitor. The cleared cell lysate was applied to a 1 ml HisTrap™ FF column, washed with 10 column volumes of buffer A and subsequently eluted by a step gradient of buffer B (buffer A supplemented with 490 mM imidazole). The affinity purified AlpC protein was further applied to a Superdex™ 200 10/300 gel filtration column and eluted in buffer C (buffer A lacking imidazole). Eluted fractions were analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting. Heterologous overexpression and purification of $AlpC^{D301A}$ was identical.

The production of His_{10}-AlpA in *E. coli* BL21 (DE3) pLysS was induced with 0.5 mM IPTG and harvested after 4 h of expression as described for His_{10}-AlpC. The disruption of the cells and purification of the protein with a Ni^{2+}-NTA column (nickel-nitriloacetic acid) (Qiagen, Hilden) were performed as described previously (53). The elution of AlpA was conducted with TNI400 buffer (20 mM Tris/HCl, 300 mM NaCl, 400 mM imidazole). The fractions of eluted protein were pooled and the buffer was exchanged to binding buffer (20 mM Tris/HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 5% (v/v) glycerol, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 0.005% (w/v) Triton X-100) with a PD10 desalting column (GE Healthcare) for DNA–protein binding studies.

Electrophoretic mobility shift assays

Studies of the DNA–protein binding of AlpA and potential target DNA were performed according to (54). The DNA (500 bp) of the upstream region of *alpA* and downstream region of *alpC* were amplified by PCR. The promoter region of cg2036, a gene encoded by the prophage CGP3 in *C. glutamicum*, was tested as control fragment. 100 ng DNA per lane were incubated with different molar ratios of purified AlpA (0–500-fold molar excess) for 20 min before loading to a nondenaturing 10% polyacrylamide gel.

Nucleotide hydrolysis assay

ATPase and GTPase activity was measured in a coupled enzyme assay constantly regenerating ATP/GTP, allowing monitoring of ATP hydrolysis over time (55). The regeneration of ATP/GTP is coupled to the oxidation of nicotinamide adenine dinucleotide phophate (NADH). When ADP/GDP is converted back to ATP/GTP pyruvate kinase converts phosphoenolpyruvate (PEP) to pyruvate. Lactate dehydrogenase then converts the pyruvate to lactate, resulting in the oxidation of one NADH molecule. The decrease of NADH absorbance at 340 nm is monitored over time.

Nucleotide hydrolysis rate of 1 μM protein was performed in buffer containing 50 mM Tris/HCl pH 7.5, 150 mM NaCl, 10% glycerol. The reaction was performed in a total volume of 100 μl containing varying nucleotide concentration, equimolar $MgSO_4$, 1 mM PEP, 0.6 mM NADH, 20 U/ml pyruvate kinase and 20 U/ml lactate dehydrogenase. The samples were set-up in a 96-well microtiter plate and the absorbance of NADH was monitored at 340 nm for 1 h at 30°C (Tecan Plate Reader, software: I-control). All data were obtained from triplicate determination and corrected for nucleotide autohydrolysis.

Sedimentation assay

Prior to experimental setup, protein samples were subjected to centrifugation at 120 000 × *g* for 10 min to remove any aggregated protein. Purified protein (2 μM) was mixed with nucleotide (2 μM) in the presence or absence of 2 μM Mg^{2+} or 4 μM EDTA. Volumes were adjusted to 100 μl with buffer C. Reaction mixtures were incubated at 30°C for 30 min. Higher-ordered AlpC protein complexes were then sedimentated by centrifugation at 120 × 000 *g* for 10 min. Supernatant and pellet fractions were separated and quantitatively analyzed by immunoblot.

For quantitative western blot analyses His_{10}-AlpC and His_{10}-$AlpC^{D301A}$ from the sedimentation assays, samples were blotted onto an Immobilon-FL PVDF membrane (Millipore) and probed with anti-His antibody (Qiagen), followed by IRDye 800-conjugated goat anti-mouse IgG (H + L) antibodies (LI-COR) as secondary antibody. The IR fluorescence signals were quantified with the OdysseyTM IR fluorescence scanning system (LI-COR). Background

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

values were automatically subtracted using the Odyssey software (LICOR) (Median Top/Bottom method). Signals were normalized to a protein only sample.

Co-sedimentation assay

To test for interaction of AlpA and AlpC, co-sedimentation assays with purified proteins were performed. AlpC and AlpA were mixed in equimolar ratio (5 μM, respectively, in a reaction volume of 20 μl) in EMSA buffer (20 mM Tris/HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 5% (v/v) glycerol, 0.5 mM EDTA and 0.005% (w/v) Triton X-100). The mixture was incubated at 30°C for 30 min in the presence or absence of 1 mM ATP and/or 100 ng DNA (500 bp fragment of the upstream region of *alpA*). The reaction mixture was centrifuged at 16 100 × *g* for 1 h. The supernatant was withdrawn and the sedimented protein was resuspended in the same volume. Both fractions were separated on a 12% SDS-PAGE and visualized by Coomassie staining. The amount of sedimented AlpA protein was quantified by using ImageQuant TL software (GE Healthcare).

Fluorescence microscopy

For microscopic examination, expression from the pEKEx2 plasmid was induced by addition of 0.5 mM IPTG to the growth medium, for ~60–90 min. For co-visualization of eCFP-AlpC filament and induced CGP3 prophage, cells were grown in CGXII medium supplemented with 4% glucose. Excision of the CGP3 prophage was induced by addition of mitomycin C (final concentration 0.6 μg/ml or as indicated in text). Approximately 30 min prior to microscopic examination, 0.1 mM IPTG was added to induce synthesis of YFP-TetR. AlpC-mCherry expressed from its native promoter was induced by addition of 5 μM mitomycin C.

For phase contrast and fluorescence microscopy, 1–3 μl of a culture sample was placed on a microscope slide coated with a thin 1% agarose layer and covered by a cover slip. Images were taken on a Zeiss AxioImager M1 equipped with a Zeiss AxioCam HRm camera or on a Zeiss AxioImager M2 equipped with a Zeiss AxioCam MRm camera. GFP fluorescence was monitored using filter set 38 HE eGFP, BG-430 fluorescence and CFP (eCFP) were monitored using filter 47 HE CFP, red fluorescence (membrane stain) was monitored by using filter 43 HE Cy3 or filter 63He and DAPI / Hoechst fluorescence was examined with filter set 49. An EC Plan-Neofluar 100x/1.3 Oil Ph3 objective was used. Digital images were acquired and analyzed with the AxioVision 4.6 software (Carl Zeiss). Final image preparation was done using Adobe Photoshop 6.0 (Adobe Systems Incorporated).

Time lapse and FRAP analysis was carried out using a Delta Vision Elite (GE Healthcare, Applied Precision) equipped with an Insight SSI™ illumination, an X4 laser module and a CoolSnap HQ2 CCD camera. Images were taken with a 100× oil PSF U-Plan S-Apo 1.4 NA objective.

FRAP analysis

Fluorescence recovery after photobleaching experiments were performed using a *C. glutamicum* strain expressing AlpC-mCherry under control of the native promoter, but encoded on a plasmid (pEC-XC99E-P_{alpA}-*alpC-mcherry*). Induction of AlpC expression was initiated with 5 μM mitomycin C addition to exponentially growing cells. Cells were mounted on BHI agar pads at 30°C. Cells were imaged using an DeltaVision Elite system (GE Healthcare) using the SSI illumination system with the following settings: mCherry excitation 50% SSI with 0,2 s exposure. Bleaching was done with a 561 nm laser (50 mW) at 10% power and a 100x Oil PSF Objective (U-PLAN S-APO 100X Oil, 1.4NA, 0.12 WD). The duration of the laser pulse was 0.01 s. Time lapse image series were taken every 10 s and the laser event was placed after the first image. Dark-state reversal of the mCherry fluorophore was controlled for by fixing cells with 1% formaldehyde prior to image analysis (30 min room temperature). Images were analyzed with Fiji and values were normalized for both bleaching and cytoplasmic background fluorescence. Relative values were used in order to allow comparison between cells, where a value of 1 corresponds to the fluorescence intensity before the bleaching event and a value of 0 corresponds to the cytoplasmic background fluorescence. The cytoplasmic background fluorescence value is the average between the cytoplasmic fluorescence of three cells selected from the analyzed microscope image. Subtraction of cytoplasmic fluorescence from the original fluorescence value highlights filament dependent fluorescent signals.

$$\text{Filament fluorescence} = \text{total fluorescence} - (\text{cytoplasmic background fluorescence average} * \text{cell area})$$

Bleaching due to imaging was calculated by averaging the decrease of relative fluorescence in three independent areas of the microscope image not affected by the laser bleaching event. Images were assembled in Adobe Photoshop and graphs plotted with Microsoft Excel.

RESULTS

A prophage-encoded cytoskeletal protein

In previous work, our labs have shown that the *C. glutamicum* ATCC 13032 prophage CGP3 is able to excise from the chromosome and replicate autonomously (44). The first open reading frame in the prophage, cg1890, shows low sequence identity to bacterial actins (Figure 1). Recently, a phylogenetic analysis identified Cg1890 as an actin-like protein (56). Cg1890 shares considerably low sequence identity (mostly <30%) with actin and other actin-like proteins, however the typical actin signature motif is conserved (Figure 1B). Hence, we renamed Cg1890 to AlpC (designated AlpC, for Actin-Like Protein *Corynebacterium*). The actin signature motifs are involved in binding nucleotide in the presence of a divalent cation and catalyzing the transfer of a phosphoryl group from ATP to a hydroxyl group (57,58). Amide residues from the conserved loops of the phosphate 1 and phosphate 2 motifs are involved in hydrogen bonding with the β- and γ-phosphates of ATP (58). The connect 1 sequence makes interactions with the metal ion complexed to ATP.

A phylogenetic search revealed that among actin-like proteins AlpC and homologous proteins are closer related to

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

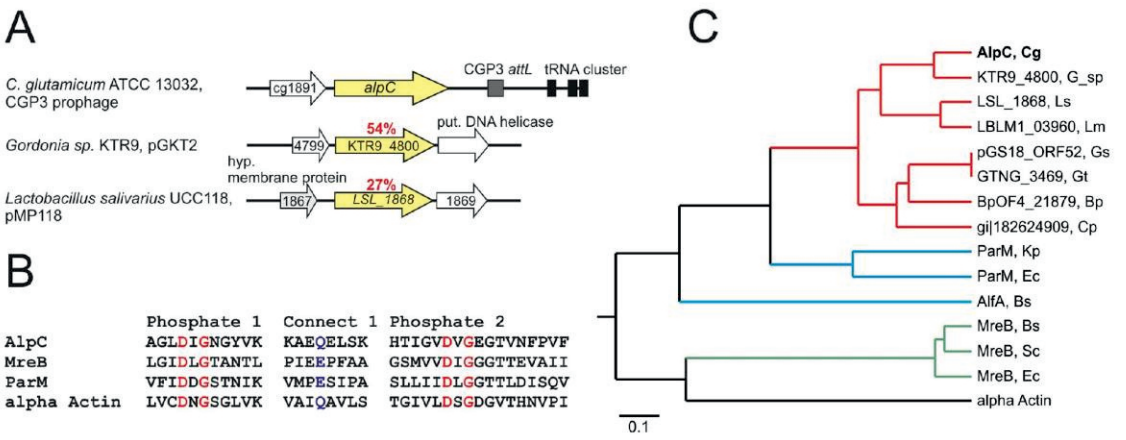


Figure 1. Genetic organization and phylogenetic analysis of AlpC. (**A**) Genetic organization of *alpC* and homologous proteins (yellow). (**B**) Alignment of the phosphate 1 (red), connect 1 (blue) and phosphate 2 (red) regions (57). The residues of *C. glutamicum* ATCC 13032 AlpC, *B. subtilis* MreB, *E. coli* ParM and human alpha actin are shown. Highlighted residues correspond to alpha actin D13 and G15 (red), Q139 (blue), and D156 and G158 (red). (**C**) Phylogenetic tree visualizing the relationship of AlpC homologs to bacterial actin-like proteins and human alpha actin. Branches shown in green indicate MreB proteins, branches colored in blue indicate plasmid partitioning ParM and AlfA proteins. AlpC and homologous proteins (red) are closer related to ParM. Amino acid sequences were aligned using the neighbor-joining method (ClustalW): AlpC Cg, *C. glutamicum* ATCC 13032 (gil62390556), KTR9_4800 G_sp, Gordonia sp. KTR9 (gil301321491), LSL_1868 Ls, *Lactobacillus salivarius* UCC118 (gil90962843), LBLM1_03960 Lm, *Lactobacillus mucosae* LM1 (gil377831150), pGS18_ORF52 Gs, *Geobacillus stearothermophilus* (gil169636508) (plasmid-encoded), GTNG_3469 Gt, *Geobacillus thermodenitrificans* NG80-2 (gil138898362), BpOF4_21879 Bp, *Bacillus pseudofirmus* OF4 (gil288557196), Cp, *Clostridium perfringens* D str. JGS1721 (gil182624909), ParM Kp, *Klebsiella pneumonia* (gil146150982), ParM Ec, *E. coli* (gil385721336), MreB Bs, *B. subtilis* (gil255767642), MreB Sc, *Streptomyces coelicolor* (gil21221069), MreB Ec, *E. coli* (gil377940072), alpha actin, *Homo sapiens* (gil178029).

plasmid partitioning systems (ParM) than to MreB-type cytoskeletal proteins (Figure 1C). Homologous proteins of *C. glutamicum* AlpC were found to be, in most cases, plasmid-encoded and were identified in a variety of Gram-positive bacteria such as *Lactobacillus* and *Bacillus* species as well as the pathogen *Clostridium perfringes* (Figure 1C). While the phylogenetic analysis carried out by Derman and coworkers identified of a large number of actin-like proteins (56), the role of phage-encoded actin-like proteins has not been studied in further detail, yet.

AlpC hydrolyzes nucleotides *in vitro*

A distinguishing feature of actin and actin related proteins is the ability to hydrolyze nucleotides. In order to analyze nucleotide hydrolysis of AlpC, a recombinant His-tagged variant of AlpC and a mutant AlpC variant lacking the conserved aspartic acid of the phosphate 2 motif ($AlpC^{D301A}$) were overproduced in the heterogolous bacterium *E. coli* and purified. In a nucleotide hydrolysis assay, the activity of 1 μM of purified protein was measured with increasing concentrations of ATP or GTP (Figure 2A). Akin to other characterized actin-like proteins, AlpC can hydrolyze both ATP and GTP, and exhibits Michaelis–Menten kinetics. In the presence of ATP, the maximum turnover rate is reached at slightly lower substrate concentrations compared to GTP. In the presence of GTP, AlpC has a V_{max} of 4.3 mM min^{-1}, compared to 3.75 mM min^{-1} for ATP (Figure 2A). The measured K_m is low and does not differ significantly between ATP and GTP, 0.2 mM and 0.43 mM, respectively. *In vivo*, the more abundant ATP is probably the favored substrate. Nucleotide hydrolysis is abolished in the $AlpC^{D301A}$ mutant (Figure 2A).

AlpC assembles into higher-ordered oligomeric complexes, *in vitro* and *in vivo*

One of the defining properties of actin and related cytoskeletal proteins is the ability to polymerize into filamentous structures *in vivo*. To determine if AlpC assembles into higher-ordered oligomeric complexes *in vitro*, sedimentation assays were performed (Figure 2B). After removal of potential protein aggregates by centrifugation, 2 μM of purified protein was incubated for 30 min at 30°C in the presence or absence of nucleotide, Mg^{2+} and EDTA as indicated in Figure 2B. The higher-ordered oligomeric AlpC complexes were isolated by centrifugation. Sedimentation of AlpC was dependent on the presence of nucleotide (ATP or GTP) and Mg^{2+}. Addition of the metal chelator EDTA abolished sedimentation of AlpC, ruling out sedimentation of protein aggregates. The inactive mutant $AlpC^{D301A}$ did not exhibit a nucleotide-dependent sedimentation behavior (Figure 2B). Sedimentation of $AlpC^{D301A}$ was observed when incubated in the presence of ATP only or GTP and Mg^{2+}, however assembly into higher-ordered $AlpC^{D301A}$ complexes was not observed when incubated in the presence of nucleotide, Mg^{2+} and EDTA

In order to study the behavior of AlpC at physiological concentration, the native *alpC* locus was replaced with an *ecfp-alpC* allele (CDC020, Supplementary Table S1). Under conditions of prophage induction using mitomycin C, eCFP-AlpC (CDC020) readily assembled into filaments (Figure 3A). AlpC filaments or foci were observed in the

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

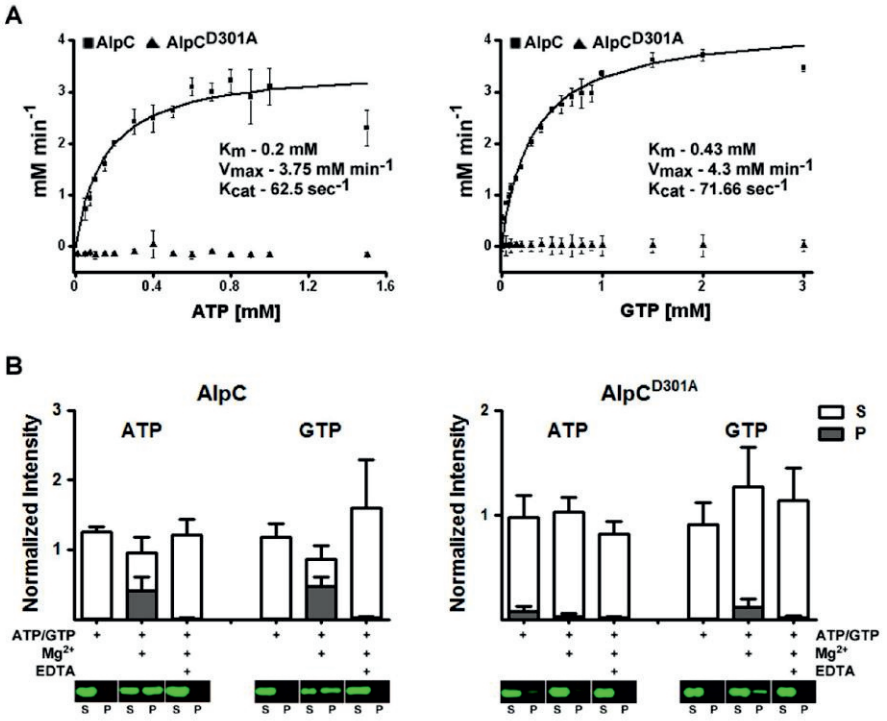


Figure 2. AlpC hydrolyses ATP and GTP, *in vitro*. (**A**) Nucleotide hydrolysis was assayed using a coupled assay in which ATP/GTP was continuously regenerated from ADP/GDP (see Materials and Methods). The nucleotide hydrolysis rate was assayed in the presence of 1 μM protein with increasing concentration of ATP/GTP. Plots are fitted for Michaelis–Menten kinetics. The hydrolysis kinetics of AlpC is similar for both ATP and GTP. Mutation of the conserved aspartic acid of the phosphate 2 motif of AlpC (AlpC^{D301A}) abrogates nucleotide hydrolysis. Error bars indicate the standard error of three independent experiments. (**B**) AlpC assembles into higher-ordered protein complexes. Polymerization of 2 μM AlpC or AlpC^{D301A} was assayed at 30°C in the presence of 2 mM nucleotide (ATP or GTP), 2 mM Mg^{2+} and/or 4 mM EDTA, as indicated, and detected by means of sedimentation assays and quantitative Western blot. All fluorescence intensities are normalized against AlpC or AplC^{D301A} protein samples lacking nucleotide, Mg^{2+} or EDTA. Sedimentation of AlpC is dependent on the presence of nucleotide and Mg^{2+}. In the presence of EDTA sedimentation of AlpC is abolished. Sedimentation of the hydrolytic inactive mutant (AlpC^{D301A}) was observed (in the presence of ATP only or GTP and Mg^{2+}), however, significantly reduced compared to the wild-type AlpC protein. In the presence of EDTA AlpC^{D301A} did not sediment. A montage of the immunoblotted samples is shown in the lower part of the figure. S, supernatant; P, pellet. Error bars indicate the standard error of three independent experiments.

vast majority of cells. The filaments were varying in length, straight and mostly found pointing to the cell membrane at different angles. These results suggest that AlpC expression and filament formation occurs in response to induction of the CGP3 prophage.

The low level of eCFP-AlpC (CDC020) produced when expressed at physiological concentration made it unsuitable for further analysis, such as analysis of filament distribution or time lapse. Consequently, *alpC-cfp*, encoding a fusion protein of AlpC and CFP, was expressed from a pEKEx2 plasmid (CDC021, Supplementary Table S1), bringing the fusion gene under control of an IPTG inducible promoter. In the majority of cells (97.2%, $n = 870$), AlpC-CFP (CDC021) assembled into numerous long and curved filaments, often extending from one pole to the other (Figure 3B). However, some cells contained either a combination of filaments and foci or only foci. To gain more insight into the distribution of the AlpC-CFP filaments within the cell, Z-stacks were acquired. As shown in Figure 3C (and Movie S1), a number of long, curved filaments extending the length of the cell were observed, some filaments appearing to elongate close to the cell membrane. Thus, akin to actin and related cytoskeletal proteins AlpC can readily polymerizes forming filamentous structures *in vivo*. The catalytically inactive mutant AlpC^{D301A} also assembled into filaments *in vivo* when overproduced (CDC022, Supplementary Table S1) (Figure 3B). However, the frequency of filament formation was drastically reduced (1.6%, $n = 750$) and the morphology of the AlpC^{D301A} filaments was different from the wild type protein.

AlpC filaments are dynamic

In vivo, filaments of the actin-like protein ParM involved in plasmid partition were shown to be extremely dynamic, undergoing bursts of rapid growth and catastrophic decay (59). Although such dynamic instability is not a distinguishing feature of actin and actin-like proteins, filament dynamics is essential for protein function. To ascertain if *C. glutam-*

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

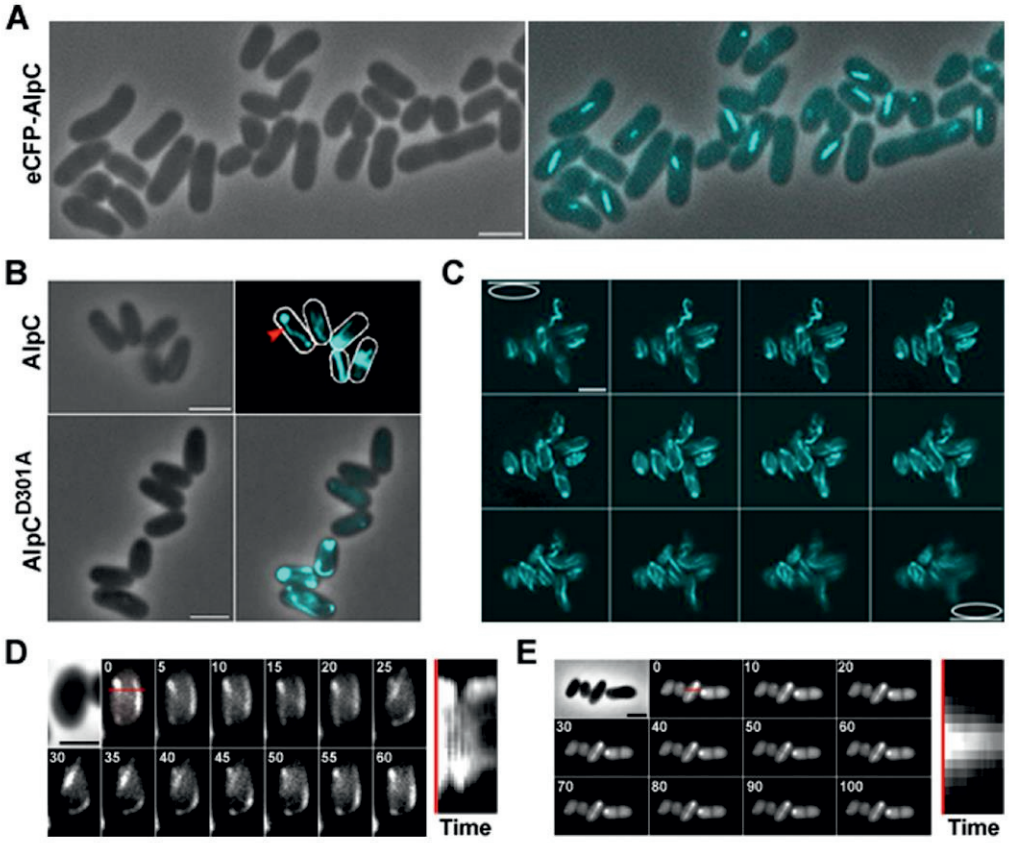


Figure 3. AlpC assembles into filaments in *C. glutamicum*. (**A**) eCFP-AlpC (CDC020) assembles into short, straight filaments when expressed at physiological concentration. Cells were grown in CGXII containing 4% glucose. Excision of the CGP3 prophage was induced with 0.6 μg/ml mitomycin C for 3 h prior to microscopic analysis. (**B**) AlpC-CFP (CDC021) assembles into long, curved filaments when overproduced in *C. glutamicum*, red arrowhead. Cells were grown in BHI and induced by addition of 0.5 mM IPTG 1 h prior to microscopic analysis. Mutation of the conserved aspartic acid of the phosphate 2 motif of AlpC reduces filament assembly (AlpC^{D301A}-CFP, CDC022) (lower panel). (**C**) Z-stack analysis of the subcellular distribution of AlpC-CFP (CDC021) filaments. Cells overproducing AlpC-CFP contain numerous filaments of various orientation, curvature and length (Movie S1). (**D**) AlpC-CFP filaments are dynamic. Images were acquired at 5 s intervals (top left) for 1 min (Movie S2). The red line denotes the position in the cell used to generate the kymograph (right). (**E**) AlpC^{D301A}-CFP filaments are static. Images were acquired at 5 s intervals (top left) for 100 s (Movie S3). As in (C), a kymograph is shown on the right. Scale bar, 2 μm

icum AlpC is also dynamic, time-lapse analysis was carried out. For this purpose a strain harboring a plasmid-encoded AlpC-CFP translational fusion was used (CDC021, Supplementary Table S1). Cells were grown in BHI medium, induced with 0.5 mM IPTG for 1 h prior to microscopic analysis. Cells were placed on a microscopic slide coated with 1% agarose in BHI. Images were acquired at 5 s intervals for a total of 60 s. The time lapse analysis revealed that AlpC-CFP (CDC021) filaments are indeed dynamic (Figure 3D and Movie S2). While some filaments appear to move along the membrane, other filaments curl into the cytoplasm. *In vivo*, filament assembly of the catalytically inactive mutant (AlpC^{D301A}-CFP, CDC022) is not completely abolished, however the frequency of filament formation is greatly reduced (Figure 3B). The AlpC^{D301A}-CFP filaments are more stable and, hence less dynamic than wild-type AlpC filaments (Figure 3E, Movie S3). Filament dynamics were further studied using fluorescence recovery after photobleaching (FRAP) experiments. For this purpose, we constructed a strain harboring a plasmid-encoded AlpC-mCherry under control of the native promoter. This system allows for moderate AlpC-mCherry overproduction. The AlpC-mCherry strain was necessary for FRAP experiments because the natural level of AlpC leads to shorter filaments which will be bleached almost entirely and hence do not allow for FRAP analysis. AlpC-mCherry induction was achieved by addition of 5 μM mitomycin C. Readily, cells produced long, curved AlpC-mCherry filaments (Supplementary Figure S1). Parts of these filaments were bleached using the 561 nm laser and a subsequent time lapse series with image acquisition at 10 s intervals revealed rapid recovery of the AlpC-mCherry (Figure 4A and B, Movie S4).

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

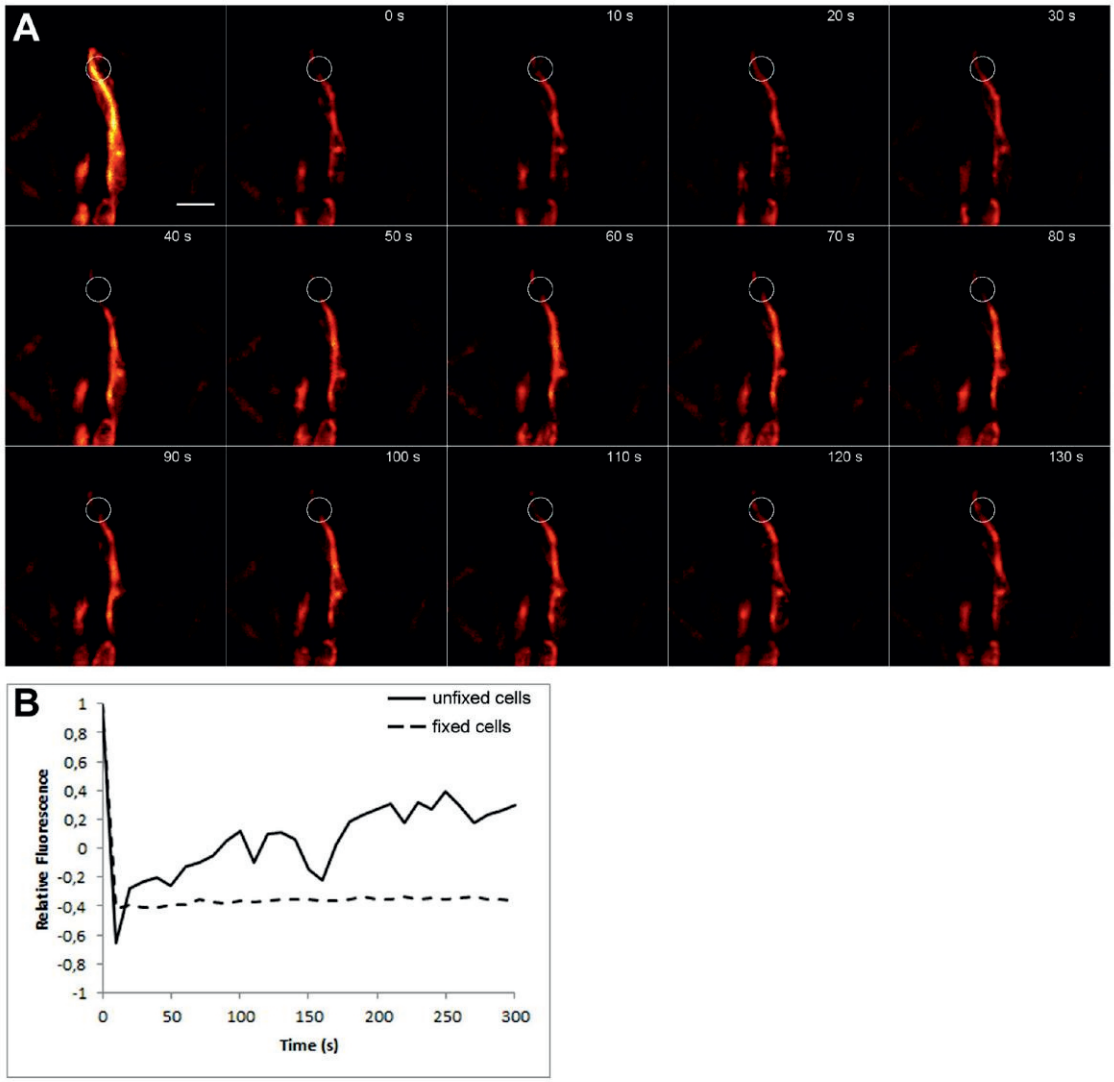


Figure 4. FRAP analysis of AlpC-mCherry. Cells encoding AlpC-mCherry were induced using 5 μM mitomycin C and imaged after 60 min of induction. Cells were grown on a BHI agar pad at 30°C. Part of the AlpC-mCherry filament was bleached using a 561 nm laser as described in Material and Methods section. Images were taken every 10 s. An assembly of still images acquired for 120 s is shown in (**A**). The full movie can be seen in the supplemental material (Movie S4). In a control experiment cells were fixed prior to bleaching and no recovery was observed (Supplementary Figure S2 and Movie S5). (**B**) FRAP curves for the AlpC-mCherry strain without fixing and with fixing prior to bleaching (fixed) are shown. Values are normalized for background fluorescence and intracellular mCherry fluorescence. The curve indicates recovery that is due to incorporation of unbleached AlpC-mCherry.

We controlled for dark-state-reversal of the mCherry fluorophore by fixing the cells with 1% formaldehyde. Subsequent FRAP experiments did not reveal any recovery of the mCherry fluorophore (Supplementary Figure S2, Movie S5), suggesting that the fast initial recovery is not a result of photoswitching, but reflects the true dynamics of the AlpC protein.

Identification of an AlpC adaptor protein

Actin-like proteins encoded on plasmids are often co-transcribed with an adaptor protein, which connects the actin filament with the plasmid DNA. In many cases, the adaptor protein stimulates or stabilizes the actin filaments (20,24,60,61). AlpC is encoded in a putative operon together with Cg1891, a protein of unknown function (Figure 1A). We postulated that Cg1891 could be an adaptor

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

that couples phage DNA to AlpC filaments. Hence, we renamed cg1891 to *alpA* (A for adaptor). To test whether AlpA binds to specific DNA regions, purified His_{10}-AlpA protein was analyzed by electrophoretic mobility shift assays (EMSA). We tested DNA fragments covering the up- and downstream region of the putative *alpA–alpC* operon (Figure 5A). The upstream region was specifically bound by AlpA, *in vitro* (Figure 5B). A smear of shifted DNA rather than a clear band was observed, indicating an oligomerization of the protein along the DNA, presumably at more than one binding site (62). The downstream region and the control DNA fragment (upstream of the CGP3 gene cg2040) showed no binding of AlpA.

Analysis of the *alpAC* promoter region revealed the presence of 24 conserved DNA repeats (designated as *alpS*) (Figure 6A). EMSA studies with purified AlpA were carried out to determine if AlpA binds to DNA fragments covering these repeats (Figure 6B). As outlined in Figure 6A, DNA fragments covering various portions of the *alpAC* promoter region were used for EMSA analysis. Both sub-fragments 1 and 2, which covered the majority of the conserved DNA repeats of the promoter region, displayed AlpA binding. The third sub-fragment and the control DNA, both of which lack DNA repeats, do not appear to be bound by AlpA. Using the MEME suite software a consensus motif was derived (TTAAnnG), which revealed five highly conserved bases within the *alpS* motif (Figure 6C, and Supplementary Figure S3). However, this DNA motif is not unique to the *alpAC* operon. In comparison to the whole genome, the *alpS* motif is more concentrated at the *alpAC* operon. These results suggest that AlpA binds to conserved DNA repeats, *alpS*, of the promoter region, however analysis of the exact mechanism will require further studies.

Co-localization of AlpA and AlpC *in vivo*

To support our hypothesis that AlpA is an adaptor protein linking CGP3 DNA to the AlpC filament, localization of both proteins was studied *in vivo*. For this purpose, an in-frame deletion of *alpA* was constructed in the background of strain CDC020, in which *alpC* was replaced by allelic exchange with e*cfp-alpC*. Deletion of *alpA* had no influence on the formation of eCFP-AlpC filaments. AlpC also assembles into filaments when expressed in *E. coli* in the absence of AlpA (data not shown). To generate a strain that allows co-visualization of both AlpA and AlpC the *alpA* gene was fused to *eyfp* and cloned under control of the native promoter in the Δ*alpA alpC*::*ecfp-alpC* background (ATCC 13032 *ΔalpA alpC::ecfp-alpC* P_{alpA}-*alpA-eyfp*, Supplementary Table S1). Fluorescence microscopy revealed foci of AlpA-eYFP, which formed upon induction of CGP3 by addition of mitomycin C. Frequently, one to two foci of AlpA-eYFP per cell were observed at various positions; in a few cases up to four foci were detected in single cells. Remarkably, AlpA-eYFP foci were observed at the tips or aligned with AlpC filaments (Figure 5C). In 78% ($n = 37$) of cells at least one AlpA focus was found to be associated with an AlpC filament. Co-localization was further confirmed with plasmid-encoded variants of AlpA-eYFP and AlpC-mCherry. In agreement with the genomic replacement, co-localization of AlpA-eYFP foci and AlpC-mCherry filaments was observed in 91% of the cells ($n = 116$, Supplementary Figure S1). In both experiments, the angle of the filaments and the position of the putative adaptor protein foci were, however, variable. The co-localization of AlpA and AlpC suggested that AlpA might move along the preformed AlpC filament track. To further reinforce this idea and that AlpA is AlpC-CGP3 adaptor protein, live cell time lapse co-visualization of AlpA-eYFP and AlpC-mCherry was carried out. Again, prior to microscopic analysis, cells were treated with mitomycin C to induce the CGP3 prophage. Time lapse analysis revealed that the AlpA-eYFP foci move along the AlpC-mCherry filament (Figure 7B and MovieS6).

In vivo co-visualization of the CGP3 viral DNA and AlpC filaments

The CGP3 prophage and eCFP-AlpC (CDC020), where AlpC is expressed at the physiological level, were co-visualized *in vivo* by fluorescence microscopy. Visualization of CGP3 DNA was carried out as described previously (44). Basically, an array of *tetO* operator regions of transposon Tn10 was integrated into an intergenic region within the CGP3 prophage region. Co-expression of plasmid-encoded *yfp-tetR* allowed direct visualization of the prophage region. Prophage induction was triggered by mitomycin C addition. In many cells an increased number of foci corresponding to CGP3 DNA were visible (44). In these cells, eCFP-AlpC filaments or foci were readily observed. The eCFP-AlpC foci often co-localized with the CGP3 prophage foci (33%, $n = 56$) (Figure 7A). In many cases cells containing an AlpC filament had CGP3 DNA at the filament end. We noticed that most AlpC filaments point in an angle to the membrane and seem to push the phage DNA towards the cell membrane. Indeed, phage DNA is mostly found at the cell membrane (Figure 7A).

AlpA interacts with AlpC *in vitro*

The binding of AlpA to phage DNA as well as the co-localization of AlpA foci with AlpC filaments, suggest a direct interaction of AlpAC, coupling the CGP3 DNA with the AlpC filament. To test this possibility further, we carried out co-sedimentation assays with AlpA and AlpC. While AlpC exhibits a nucleotide dependent sedimentation behavior, AlpA did not sediment after incubation with ATP, DNA or a combination of ATP and DNA (Supplementary Figure S4). When AlpA and AlpC were incubated in the presence of ATP or ATP and DNA, a significant fraction (30% and 38%, respectively) of AlpA protein co-sedimented with AlpC (Figure 5D). BSA served as a control protein, which did not co-sediment with AlpC (Supplementary Figure S4). Further support for a direct interaction between AlpA and AlpC filaments is provided by a super shift experiment where oligomerization of AlpC in the presence of ATP led to a super shift of AlpA-bound *alpS* DNA (Supplementary Figure S5).

AlpA and AlpC are required for efficient CGP3 replication

To investigate the physiological role of the actin-like protein AlpC and the adaptor protein AlpA, and respective roles in

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

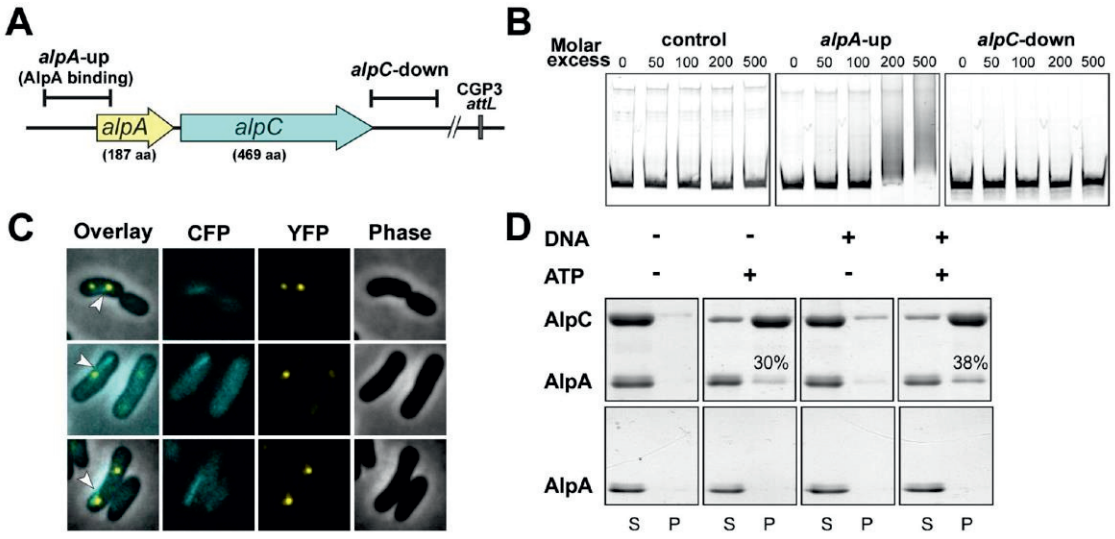


Figure 5. Characterization of the putative adaptor protein AlpA (cg1891) (**A**) Genomic localization of the AlpA binding region. The DNA fragment tested for AlpA-binding is located upstream of *alpA;* 110 bp downstream of the transcriptional start site of *alpA*. (**B**) EMSA studies of AlpA with potential DNA target regions. For DNA–protein interaction studies, DNA fragments (500 bp) of the up- and downstream region of the putative operon *alpA–alpC* were incubated with different molar ratios of AlpA (0-, 200- and 500-fold molar excess of AlpA). The promoter region of cg2036 served as control. (**C**) Co-visualization studies of eCFP-AlpC and AlpA-eYFP. The expression of both proteins was induced with mitomycin C for 2 h prior to localization analysis (**D**) AlpA interacts with AlpC *in vitro*. Purified AlpA and AlpC were incubated with ATP, DNA or both ATP and DNA. The supernatant and pellet fractions were separated by centrifugation. Both fractions were analyzed by SDS-PAGE and visualized by Coomassie staining. In the absence of AlpC, AlpA does not sediment, irrespective of the presence of ATP and/or DNA (lower panel). AlpA co-sedimentes with AlpC in the presence of ATP (30%) or ATP and DNA (38%). S, supernatant; P, pellet. The amount of sedimented AlpA protein was quantified by using ImageQuant TL software (GE Healthcare).

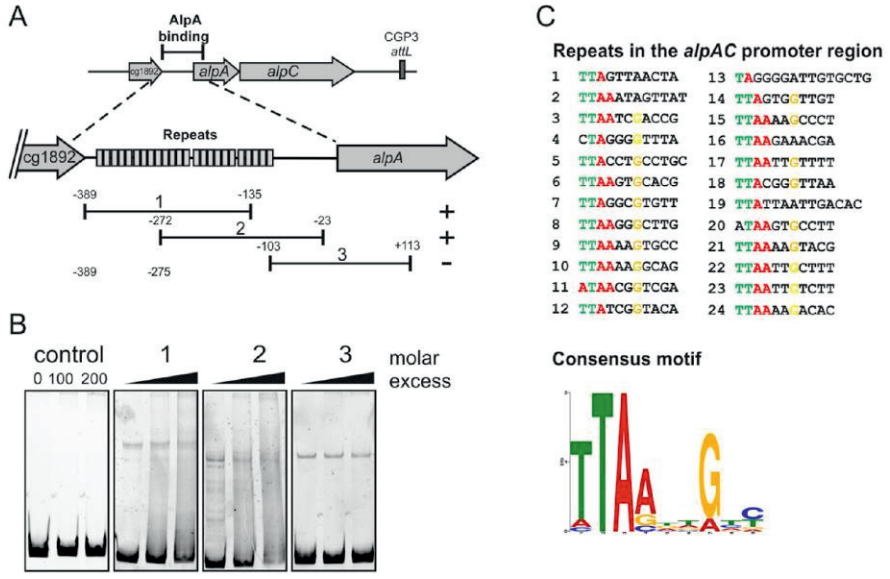


Figure 6. AlpA binds to conserved repeats in the *alpAC* promoter region. (**A**) Overview of the tested sub-fragments covering the upstream promoter region of *alpAC*. The conserved repeats are indicated as boxes. (**B**) Corresponding EMSA study to (A). To test for AlpA–DNA interaction, the DNA fragments were incubated with varying molar excess of AlpA (0-, 200- and 500-fold molar excess of AlpA; see supplemental material). (**C**) Conserved repeats in the *alpAC* promoter region. A consensus motif was derived using the MEME suite software (http://meme.nbcr.net/meme/).

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

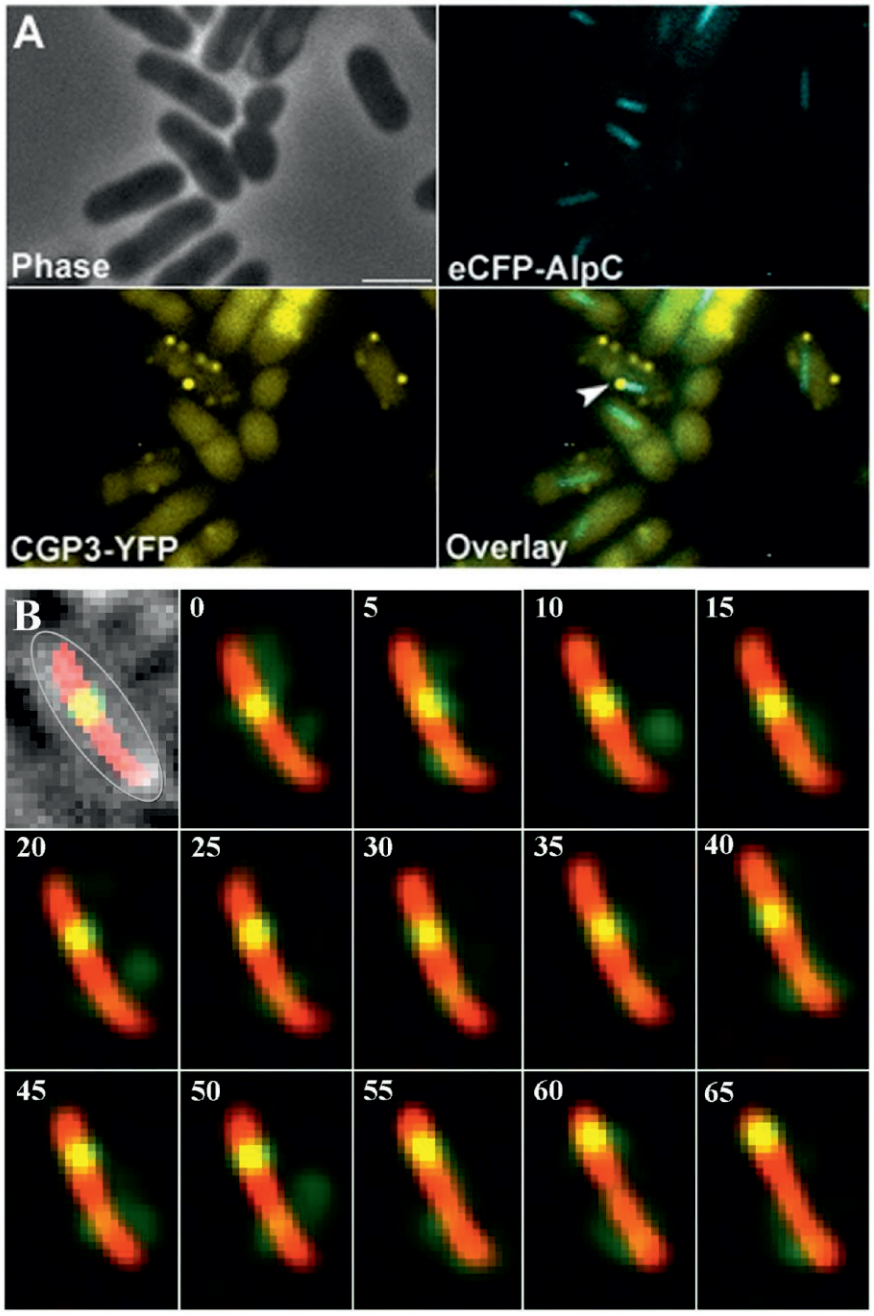


Figure 7. Co-visualization of eCFP-AlpC and induced CGP3 prophage (CGP3- YFP). (**A**). Cells expressing eCFP-AlpC as a single copy from the native promoter and a YFP labeled CGP3 DNA were treated with mitomycin C for 3 h to induce prophage excision. In the examples shown, the tip of an AlpC filament is connected to a CGP3 phage DNA molecule (arrowheads). (**B**) Time lapse analysis of AlpA-eYFP and AlpC-mCherry. Cells were treated with mitomycin C prior to microscope analysis. Images were acquired every 5 s for a total of 65 s (as indicated on the upper left corners). The first inset shows an overlay of the phase contrast, yellow and red channels, and the cell is outlined in gray. Within the time frame of the time lapse analysis AlpA-eYFP moves from the midcell region along the AlpC-mCherry filament to the cell pole (see also Movie S6). Scale bar, 2 μm.

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

the replication and/or segregation of the prophage CGP3, in-frame deletion mutants lacking the *alpC* and *alpA* gene were constructed. Deletion of either *alpC* or *alpA* has no recognizable growth phenotype. In order to analyze the impact of AlpC and AlpC on CGP3 replication, *C. glutamicum* was treated with mitomycin C to induce prophage excision and replication. The intracellular amount of circular phage DNA significantly increased and reached ∼10-fold induction in mitomycin C treated cells after 6 h. Quantification of the intracellular amount of circular CGP3 DNA by qPCR in wild type and Δ*alpC* strains revealed a similar progression (Figure 8A). However, the maximal amount of circular CGP3 DNA was ∼2-fold reduced in Δ*alpC* cells compared to wild type. As observed for Δ*alpC*, the maximal amount of phage DNA was ∼2-fold reduced in Δ*alpA*. Deletion of the CGP3 encoded cg2040, a Cro/CI type regulator, did not alter the amount of circular phage DNA (data not shown). A complementation strain encoding AlpC-CFP had no significant difference to wild type, indicating that the fusion construct is functional (data not shown). The number of induced CGP3 phages was also quantified by means of fluorescent micrsocopy, *in vivo*. Thereby, the *alpC* gene was deleted in the ATCC 13032::pLAU44-CGP3-Spec strain background. In the absence of *alpC* (strain CDC024), the average number of CGP3 per cell was reduced compared to WT cells (1.9% and 2.9%, respectively ($n \geq 210$)) (Supplementary Figure S6). These results indicate that AlpC is directly influencing CGP3 replication *in vivo*.

Transcriptome analysis revealed that both genes, *alpC* and *alpA,* are among the early genes of CGP3 induced upon treatment with mitomycin C (Figure 8B). The expression of *alpC* and *alpA* was 3–4-fold up-regulated 1 h after induction, whereas the majority of CGP3 genes showed an increased mRNA level after 3–6 h (Figure 8C). Further genes showing a >4-fold upregulation after 1 h include cg1962 and cg1977, encoding hypothetical proteins, as well as cg1996–1998 encoding a restriction modification system (*cglIMR*). Remarkably, cg1977 and the operon cg1996–1998 are both under direct control of LexA itself, explaining their high induction upon addition of mitomycin C. Altogether, early expression of *alpC* and *alpA* in the course of CGP3 induction and the reduced level of circular CGP3 DNA in the two deletion mutants emphasizes the participation of AlpC and AlpA in the early stages of phage induction and/or synthesis, such as replication of phage DNA.

DISCUSSION

Virus particles can hijack the host cell cytoskeleton for active movement within the infected cells. Crucial for this is a dynamic assembly of a scaffold that can either serve as a molecular track or generate an intrinsic movement. Examples of such cytomotive filaments are prokaryotic actin-like proteins. Actin and actin-related proteins share limited sequence and structural similarity (33,34,63,64). The formation of dynamic filamentous structures *in vivo* is common to both eukaryotic and prokaryotic actin homologs. In *C. glutamicum,* AlpC readily assembles into long curved filaments when overproduced (Figure 3B and C). At physiological concentration, AlpC assembled into short straight filaments, in addition to formation of compact foci (Figures 5C and 7A). The dynamics of actin-like filaments varies and is often linked to the mode of subunit assembly, stability and function of the protein in question (65). ParM, for example, exhibits extreme dynamic instability, displaying bursts of growth followed by rapid decay (24,59,66). Stabilization of the ParM filament, which comprises of two antiparallel filaments, requires that one end of each ParM filament, is capped with a ParR bound plasmid (60,67). Thus, the dynamic instability of the ParM filament is intrinsic to the mode of action in plasmid segregation. Similarly, the mode of action of Alp7A filaments in plasmid segregation requires that both ends of the filament are capped with a plasmid (56). The actin-like protein AlfA assembles into stable long-lived filaments by addition of subunits to one end of the filament only (61). Although the mechanism by which Alp7A segregates plasmids is not well understood, it differs from ParM mediated segregation (20,61). AlpC filaments exhibited a dynamic behavior (Figure 4A and Movie S4); however, unlike ParM and Alp7A dynamic instability was not observed (24,56,59,66).

Initially, we speculated that AlpC might function to actively segregate excised prophage DNA, akin to plasmid segregation systems. However, the genetic organization of the *alpAC* operon differs (Figure 1A) from known chromosome/plasmid segregation loci (ParM, ParR), which encode a motor protein upstream of the DNA-binding adaptor protein. An example of a similar genetic organization is found in Gordonia sp. KTR8, where two gene homologs to *alpA* and *alpC* are found on the plasmid pGKT2 (Figure 1A). In this case, the two proteins might play a role in plasmid segregation. In cells that contained multiple CGP3 foci and AlpC filaments, interaction between AlpC filaments and at least one phage DNA molecule was observed (Figure 7A and B), arguing against a role in segregation of two phage genomes into daughter cells. In the light of recent studies, which indicate that induction of the cryptic prophage CGP3 leads to cell death (Pfeifer & Frunzke, submitted for publication) a function of AlpAC in the segregation of CGP3 DNA into dividing daughter cells appears unlikely. Although CGP3 represents a cryptic, degenerated prophage element several functions, including the excision from the host genome, the SOS-dependent control of the lysogenic switch and the killing of the host cell, represent conserved processes of a lytic cycle (68).

The DNA of the induced prophage often accumulated at the cell membrane (Figure 7A and B). There is an increasing amount of evidence suggesting that plasmid and phage replication occurs at a specific subcellular localization, e.g. at the membrane (5,39,41,69). The *B. subtilis* phage φ29 serves as a model for the study of phage replication. During the early stages of infection phage replication takes place at the bacterial nucleoid and later in the infection cycle replication is redirected to the bacterial membrane (5,39,41,69). Here, we speculate that excised CGP3 phage DNA is directed to the membrane where additional replication occurs. Based on our data, we propose that circular CGP3 DNA is connected to AlpC filaments *via* the adaptor protein AlpA, which interacts with *alpS* repeats upstream of the *alpAC* promoter (Figures 5 and 6). Consistent with this idea, cells lacking AlpC or AlpA have a 2-fold decrease in phage copy

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

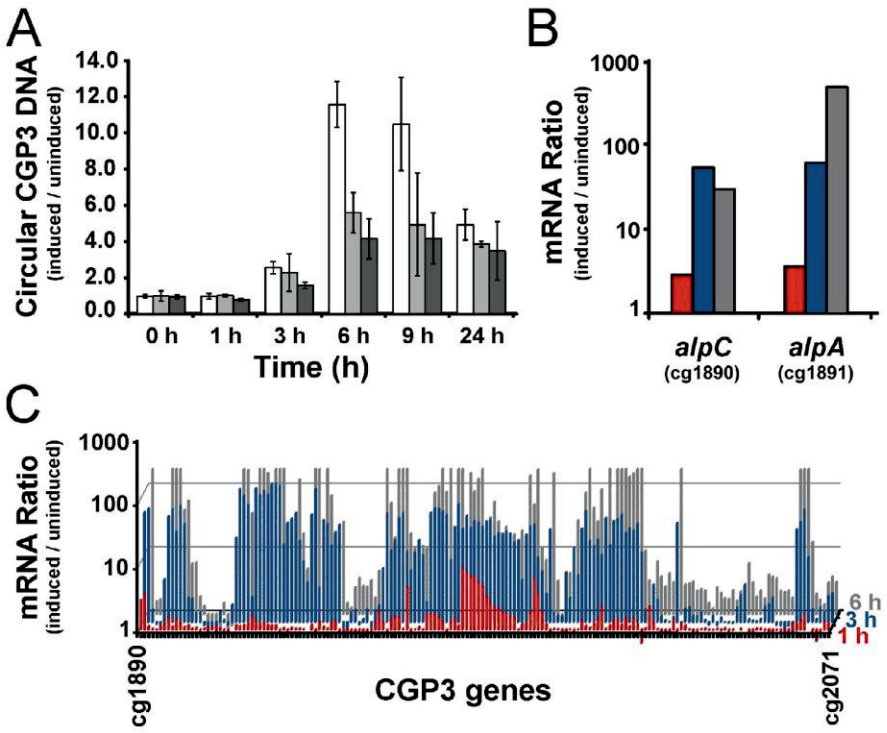


Figure 8. Impact of AlpC on CGP3 replication. (**A**) The relative amount of circular phage CGP3 DNA was quantified by qPCR in *C. glutamicum* wild type (white) and Δ*alpC* (light gray), and Δ*alpA (dark gray)* upon treatment with 0.6 μM mitomycin C. Shown are average values with standard deviation of three independent biological replicates. (**B**) Time course of *alpC* and *alpA* expression upon prophage induction triggered by the addition of 0.6 μM mitomycin C. Shown is the mRNA ratio of cells treated with mitomycin C *versus* untreated cells 1 (red), 3 (blue) and 6 h (gray) after mitomycin C addition analyzed by DNA microarrays. (**C**) Time course of the mRNA ratio of the whole CGP3 gene region after addition of mitomycin C (as described in B).

number (Figure 8), which phenocopies the impact of a *phuZ* deletion on the replication of phage 201φ2-1 DNA (43) and, thus, clearly suggests a concerted role of AlpAC in CGP3 phage replication.

Recently, reports on phage-encoded tubulin homologs have been published (43,70). Both reports describe that phage-encoded PhuZ (43) and TubZ (70) are required for correct phage placement in the infected host. Interestingly, the two reports arrive at different conclusions. PhuZ is encoded in the *Pseudomonas* phage 201φ2-1 and forms GTP hydrolysis dependent dynamic filaments *in vivo* and *in vitro* (43). These filaments traverse the entire cell and are thought to assemble phage particles in the center of the cell. However, PhuZ was expressed from an inducible promoter. Here, we have observed a similar polymerization behavior of AlpC when it was overproduced in *C. glutamicum*, but got a different picture when *alpC* was expressed from the native promoter. Phage replication was also impaired in a PhuZ mutant background, suggesting that proper segregation is important for phage replication. The authors arrive at the conclusion that PhuZ forms a spindle-like apparatus that positions the virus particles in the cell center (43). Simultaneously, a TubZ homolog in the *Clostridum botulinum* phage c-st was shown to behave like a classical type II segregation system. TubZ binds to the centromeric region of the phage DNA *via* the adaptor protein TubR (70). Superficially, the c-st TubZ and AlpC share a similar mechanism, in particular since both are encoded by prophages that replicate as plasmid-like entities. However, close inspection of the results that we describe here, make it unlikely that the short AlpC filaments observed after induction of phage replication would be suited to segregate plasmid DNA into separating daughter cells. Rather, our data are consistent with AlpC guiding CGP3 DNA to the cell membrane, where replication likely occurs. Fluorescent colocalization of AlpC filaments and viral DNA show that phage DNA foci are often found in numerous copies at the membrane, while AlpC filaments only attach to single CGP3 molecules. In this aspect AlpC might resemble the role of eukaryotic F-actin in the anterograde transport of herpes-like viruses (46,47,71). The transport brings virus particles to the cell membrane. Actin plays, in fact, a prominent role as a host factor for viral replication in eukaryotes. Actin is not only required for anterograde transport, but also involved in uptake, retrograde transport to the nucleus, replication and long-range spread (47). This is in line with observations that link the *B. subtilis* actin homolog MreB with replication of phage φ29 (41). In the absence of MreB,

Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

phage polymerase and the putative membrane anchor p16.7 are delocalized. Consequently, phage replication was reduced. It may, therefore, be advantageous for phages replicating in bacteria lacking an endogenous actin cytoskeleton to encode this cytoskeletal element within their own genome. A unifying theme is that phages with large genomes seem to rely on cytoskeletal filaments for intracellular movement. Thus, it is plausible that the connection of actin and viral replication is ancient (72) and that the general principle is well conserved even between bacteria and phages.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

The authors are thankful to Giacomo Giacomelli (LMU Munich) for help with the analysis of FRAP data.

FUNDING

Deutsche Forschungsgemeinschaft (priority program SPP 1617) [FR 2759/2-1]; Helmholtz Association (Young Investigator) [VH-NG-716]; Deutsche Forschungsgemeinschaft [BR 2915/6-1]. Funding for open access charge: Deutsche Forschungsgemeinschaft (priority program SPP 1617) [FR 2759/2-1]; Helmholtz Association (Young Investigator) [VH-NG-716]; Deutsche Forschungsgemeinschaft [BR 2915/6-1].
Conflict of interest statement. None declared.

REFERENCES

1. Canchaya,C., Fournous,G. and Brussow,H. (2004) The impact of prophages on bacterial chromosomes. *Mol. Microbiol.*, **53**, 9–18.
2. Casjens,S. (2003) Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.*, **49**, 277–300.
3. Oppenheim,A.B., Kobiler,O., Stavans,J., Court,D.L. and Adhya,S. (2005) Switches in bacteriophage lambda development. *Annu. Rev. Genet.*, **39**, 409–429.
4. Węgrzyn,G., Licznerska,K. and Węgrzyn,A. (2012) Phage lambda—new insights into regulatory circuits. *Adv. Virus Res.*, **82**, 155–178.
5. Erb,M.L. and Pogliano,J. (2013) Cytoskeletal proteins participate in conserved viral strategies across kingdoms of life. *Curr. Opin. Microbiol.*, **16**, 786–789.
6. Bloom,K. and Joglekar,A. (2010) Towards building a chromosome segregation machine. *Nature*, **463**, 446–456.
7. Löwe,J. and Amos,L.A. (2009) Evolution of cytomotive filaments: the cytoskeleton from prokaryotes to eukaryotes. *Int. J. Biochem. Cell. Biol.*, **41**, 323–329.
8. Yanagida,M. (2005) Basic mechanism of eukaryotic chromosome segregation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **360**, 609–621.
9. Ptacin,J.L., Lee,S.F., Garner,E.C., Toro,E., Eckart,M., Comolli,L.R., Moerner,W.E. and Shapiro,L. (2010) A spindle-like apparatus guides bacterial chromosome segregation. *Nat. Cell. Biol.*, **12**, 791–798.
10. Gerdes,K., Howard,M. and Szardenings,F. (2010) Pushing and pulling in prokaryotic DNA segregation. *Cell*, **141**, 927–942.
11. Schumacher,M.A. (2008) Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. *Biochem. J.*, **412**, 1–18.
12. Hayes,F. and Barilla,D. (2006) The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat. Rev. Microbiol.*, **4**, 133–143.
13. Ghosh,S.K., Hajra,S., Paek,A. and Jayaram,M. (2006) Mechanisms for chromosome and plasmid segregation. *Annu. Rev. Biochem.*, **75**, 211–241.
14. Ebersbach,G. and Gerdes,K. (2005) Plasmid segregation mechanisms. *Annu. Rev. Genet.*, **39**, 453–479.
15. Gerdes,K., Møller-Jensen,J., Ebersbach,G., Kruse,T. and Nordstrom,K. (2004) Bacterial mitotic machineries. *Cell*, **116**, 359–366.
16. Fogel,M.A. and Waldor,M.K. (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev.*, **20**, 3269–3282.
17. Funnell,B.E. (1988) Mini-P1 plasmid partitioning: excess ParB protein destabilizes plasmids containing the centromere *parS*. *J. Bacteriol.*, **170**, 954–960.
18. Salje,J., Gayathri,P. and Löwe,J. (2010) The ParMRC system: molecular mechanisms of plasmid segregation by actin-like filaments. *Nat. Rev. Microbiol.*, **8**, 683–692.
19. Schumacher,M.A. and Funnell,B.E. (2005) Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature*, **438**, 516–519.
20. Becker,E., Herrera,N.C., Gunderson,F.Q., Derman,A.I., Dance,A.L., Sims,J., Larsen,R.A. and Pogliano,J. (2006) DNA segregation by the bacterial actin AlfA during *Bacillus subtilis* growth and development. *EMBO J.*, **25**, 5919–5931.
21. Larsen,R.A., Cusumano,C., Fujioka,A., Lim-Fong,G., Patterson,P. and Pogliano,J. (2007) Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in *Bacillus thuringiensis*. *Genes Dev.*, **21**, 1340–1352.
22. Donovan,C., Schwaiger,A., Krämer,R. and Bramkamp,M. (2010) Subcellular localization and characterization of the ParAB system from *Corynebacterium glutamicum*. *J. Bacteriol.*, **192**, 3441–3451.
23. Kiekebusch,D. and Thanbichler,M. (2014) Plasmid segregation by a moving ATPase gradient. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 4741–4742.
24. Garner,E.C., Campbell,C.S., Weibel,D.B. and Mullins,R.D. (2007) Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. *Science*, **315**, 1270–1274.
25. Møllller-Jensen,J., Jensen,R.B., Löwe,J. and Gerdes,K. (2002) Prokaryotic DNA segregation by an actin-like filament. *EMBO J.*, **21**, 3119–3127.
26. Vecchiarelli,A.G., Han,Y.W., Tan,X., Mizuuchi,M., Ghirlando,R., Biertumpfel,C., Funnell,B.E. and Mizuuchi,K. (2010) ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Mol. Microbiol.*, **78**, 78–91.
27. Vecchiarelli,A.G., Hwang,L.C. and Mizuuchi,K. (2013) Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E1390–E1397.
28. Prentki,P., Chandler,M. and Caro,L. (1977) Replication of prophage P1 during the cell cycle of *Escherichia coli*. *Mol. Gen. Genet.: MGG*, **152**, 71–76.
29. Abeles,A.L., Friedman,S.A. and Austin,S.J. (1985) Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. *J. Mol. Biol.*, **185**, 261–272.
30. Li,Y., Dabrazhynetskaya,A., Youngren,B. and Austin,S. (2004) The role of Par proteins in the active segregation of the P1 plasmid. *Mol. Microbiol.*, **53**, 93–102.
31. Chastanet,A. and Carballido-López,R. (2012) The actin-like MreB proteins in *Bacillus subtilis*: a new turn. *Front. Biosci.*, **4**, 1582–1606.
32. Carballido-López,R. and Formstone,A. (2007) Shape determination in *Bacillus subtilis*. *Curr. Opin. Microbiol.*, **10**, 611–616.
33. van den Ent,F., Amos,L. and Löwe,J. (2001) Bacterial ancestry of actin and tubulin. *Curr. Opin. Microbiol.*, **4**, 634–638.
34. van den Ent,F., Amos,L.A. and Löwe,J. (2001) Prokaryotic origin of the actin cytoskeleton. *Nature*, **413**, 39–44.
35. Jones,L.J., Carballido-López,R. and Errington,J. (2001) Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. *Cell*, **104**, 913–922.
36. Daniel,R.A. and Errington,J. (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. *Cell*, **113**, 767–776.
37. Garner,E.C., Bernard,R., Wang,W., Zhuang,X., Rudner,D.Z. and Mitchison,T. (2011) Coupled, circumferential motions of the cell wall synthesis machinery and MreB filaments in *B. subtilis*. *Science*, **333**, 222–225.

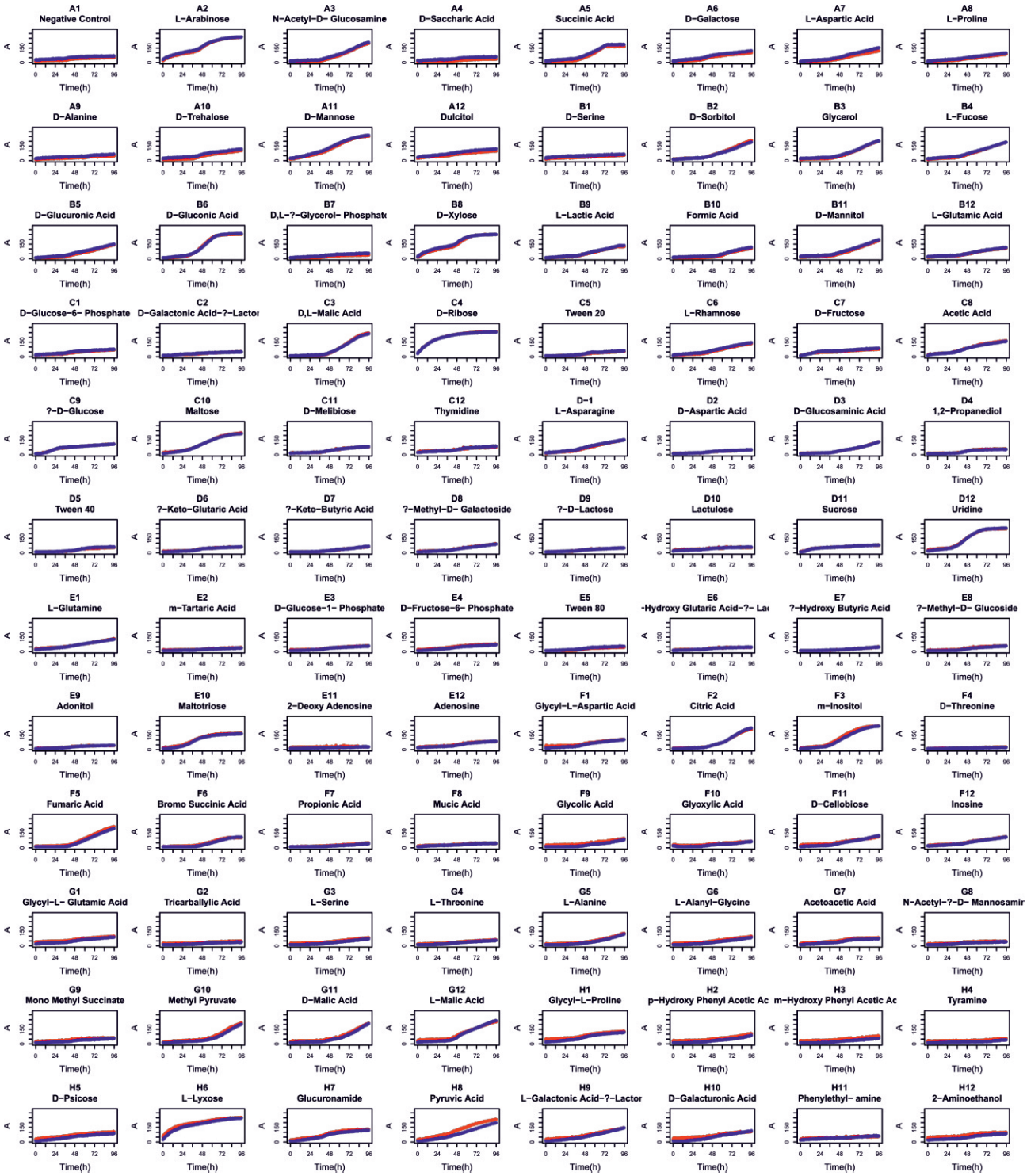
Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

38. Domínguez-Escobar,J., Chastanet,A., Crevenna,A.H., Fromion,V., Wedlich-Soldner,R. and Carballido-López,R. (2011) Processive movement of MreB-associated cell wall biosynthetic complexes in bacteria. *Science*, **333**, 225–228.
39. Muñoz-Espín,D., Serrano-Heras,G. and Salas,M. (2012) Role of host factors in bacteriophage phi29 DNA replication. *Adv. Virus Res.*, **82**, 351–383.
40. Muñoz-Espín,D., Holguera,I., Ballesteros-Plaza,D., Carballido-López,R. and Salas,M. (2010) Viral terminal protein directs early organization of phage DNA replication at the bacterial nucleoid. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 16548–16553.
41. Muñoz-Espín,D., Daniel,R., Kawai,Y., Carballido-López,R., Castilla-Llorente,V., Errington,J., Meijer,W.J. and Salas,M. (2009) The actin-like MreB cytoskeleton organizes viral DNA replication in bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **106** 13347–13352.
42. Erb,M.L., Kraemer,J.A., Coker,J.K., Chaikeeratisak,V., Nonejuie,P., Agard,D.A. and Pogliano,J. (2014) A bacteriophage tubulin harnesses dynamic instability to center DNA in infected cells. *eLife*, **3**, doi:10.7554/eLife.03197.
43. Kraemer,J.A., Erb,M.L., Waddling,C.A., Montabana,E.A., Zehr,E.A., Wang,H., Nguyen,K., Pham,D.S., Agard,D.A. and Pogliano,J. (2012) A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. *Cell*, **149**, 1488–1499.
44. Frunzke,J., Bramkamp,M., Schweitzer,J.E. and Bott,M. (2008) Population heterogeneity in *Corynebacterium glutamicum* ATCC 13032 caused by prophage CGP3. *J. Bacteriol.*, **190**, 5111–5119.
45. Kalinowski,J., Bathe,B., Bartels,D., Bischoff,N., Bott,M., Burkovski,A., Dusch,N., Eggeling,L., Eikmanns,B.J., Gaigalat,L. *et al.* (2003) The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.*, **104**, 5–25.
46. Smith,G.A., Gross,S.P. and Enquist,L.W. (2001) Herpesviruses use bidirectional fast-axonal transport to spread in sensory neurons. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 3466–3470.
47. Taylor,M.P., Koyuncu,O.O. and Enquist,L.W. (2011) Subversion of the actin cytoskeleton during viral infection. *Nat. Rev. Microbiol.*, **9**, 427–439.
48. Sambrook,J., Fritsch,E. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
49. Kirchner,O. and Tauch,A. (2002) Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. *J. Biotechnol.*, **104**, 287–299.
50. Eikmanns,B.J., Thum-Schmitz,N., Eggeling,L., Lüdtke,K.U. and Sahm,H. (1994) Nucleotide sequence, expression and transcriptional analysis of the *Corynebacterium glutamicum gltA* gene encoding citrate synthase. *Microbiology*, **140**, 1817–1828.
51. Brazma,A. (2009) Minimum information about a microarray experiment (MIAME)—successes, failures, challenges. *ScientificWorld J.*, **9**, 420–423.
52. Polen,T. and Wendisch,V.F. (2004) Genomewide expression analysis in amino acid-producing bacteria using DNA microarrays. *Appl. Biochem. Biotechnol.*, **118**, 215–232.
53. Frunzke,J., Gätgens,C., Brocker,M. and Bott,M. (2011) Control of heme homeostasis in *Corynebacterium glutamicum* by the two-component system HrrSA. *J. Bacteriol.*, **193**, 1212–1221.
54. Heyer,A., Gätgens,C., Hentschel,E., Kalinowski,J., Bott,M. and Frunzke,J. (2012) The two-component system ChrSA is crucial for haem tolerance and interferes with HrrSA in haem-dependent gene regulation in *Corynebacterium glutamicum*. *Microbiology*, **158**, 3020–3031.
55. Ingerman,E., Perkins,E.M., Marino,M., Mears,J.A., McCaffery,J.M., Hinshaw,J.E. and Nunnari,J. (2005) Dnm1 forms spirals that are structurally tailored to fit mitochondria. *J. Cell Biol.*, **170**, 1021–1027.
56. Derman,A.I., Becker,E.C., Truong,B.D., Fujioka,A., Tucey,T.M., Erb,M.L., Patterson,P.C. and Pogliano,J. (2009) Phylogenetic analysis identifies many uncharacterized actin-like proteins (Alps) in bacteria: regulated polymerization, dynamic instability and treadmilling in Alp7A. *Mol. Microbiol.*, **73**, 534–552.
57. Bork,P., Sander,C. and Valencia,A. (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 7290–7294.
58. Kabsch,W., Mannherz,H.G., Suck,D., Pai,E.F. and Holmes,K.C. (1990) Atomic structure of the actin:DNase I complex. *Nature*, **347**, 37–44.
59. Garner,E.C., Campbell,C.S. and Mullins,R.D. (2004) Dynamic instability in a DNA-segregating prokaryotic actin homolog. *Science*, **306**, 1021–1025.
60. Salje,J. and Löwe,J. (2008) Bacterial actin: architecture of the ParMRC plasmid DNA partitioning complex. *EMBO J.*, **27**, 2230–2238.
61. Polka,J.K., Kollman,J.M., Agard,D.A. and Mullins,R.D. (2009) The structure and assembly dynamics of plasmid actin AlfA imply a novel mechanism of DNA segregation. *J. Bacteriol.*, **191**, 6219–6230.
62. Derman,A.I., Nonejuie,P., Michel,B.C., Truong,B.D., Fujioka,A., Erb,M.L. and Pogliano,J. (2012) Alp7R regulates expression of the actin-like protein Alp7A in *Bacillus subtilis*. *J. Bacteriol.*, **194**, 2715–2724.
63. Daniel,R.A. and Errington,J. (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. *Cell*, **113**, 767–776.
64. van den Ent,F., Møller-Jensen,J., Amos,L.A., Gerdes,K. and Löwe,J. (2002) F-actin-like filaments formed by plasmid segregation protein ParM. *EMBO J.*, **21**, 6935–6943.
65. Cabeen,M.T. and Jacobs-Wagner,C. (2010) The bacterial cytoskeleton. *Annu. Rev. Genet.*, **44**, 365–392.
66. Gayathri,P., Fujii,T., Namba,K. and Löwe,J. (2013) Structure of the ParM filament at 8.5A resolution. *J. Struct. Biol.*, **184**, 33–42.
67. Gayathri,P., Fujii,T., Møller-Jensen,J., van den Ent,F., Namba,K. and Löwe,J. (2012) A bipolar spindle of antiparallel ParM filaments drives bacterial plasmid segregation. *Science*, **338**, 1334–1337.
68. Nanda,A.M., Heyer,A., Krämer,C., Grunberger,A., Kohlheyer,D. and Frunzke,J. (2014) Analysis of SOS-induced spontaneous prophage induction in *Corynebacterium glutamicum* at the single-cell level. *J. Bacteriol.*, **196**, 180–188.
69. Firshein,W. and Kim,P. (1997) Plasmid replication and partition in *Escherichia coli*: is the cell membrane the key? *Mol. Microbiol.*, **23**, 1–10.
70. Oliva,M.A., Martin-Galiano,A.J., Sakaguchi,Y. and Andreu,J.M. (2012) Tubulin homolog TubZ in a phage-encoded partition system. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 7711–7716.
71. Greber,U.F. and Way,M. (2006) A superhighway to virus infection. *Cell*, **124**, 741–754.
72. Villarreal,L.P. (2004) Are viruses alive? *Sci. Am.*, **291**, 100–105.

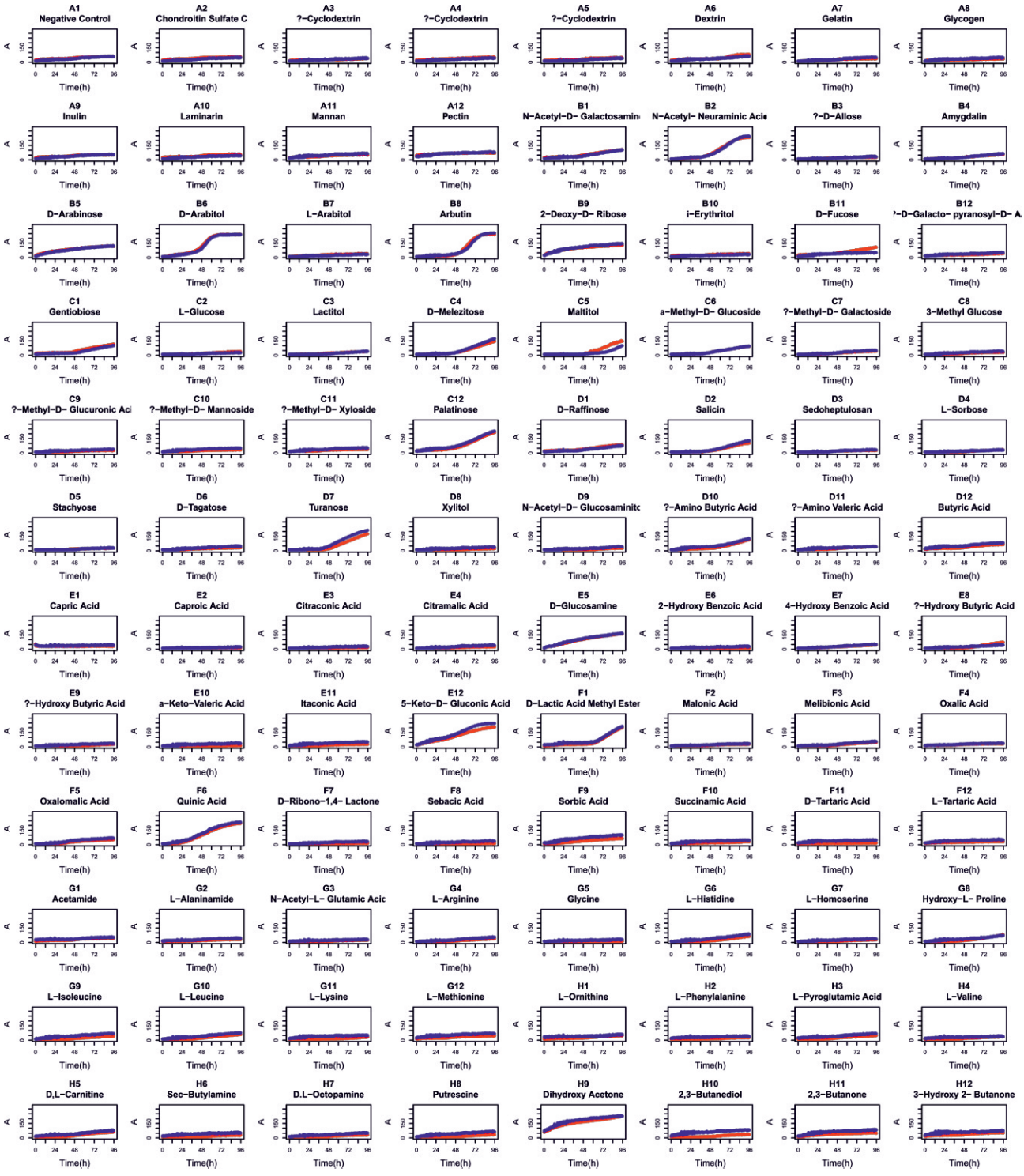
Downloaded from http://nar.oxfordjournals.org/ at Forschungszentrum Juelich GmbH, Zentralbibliothek on April 30, 2015

6.2 Results of phenotypic micro array experiment

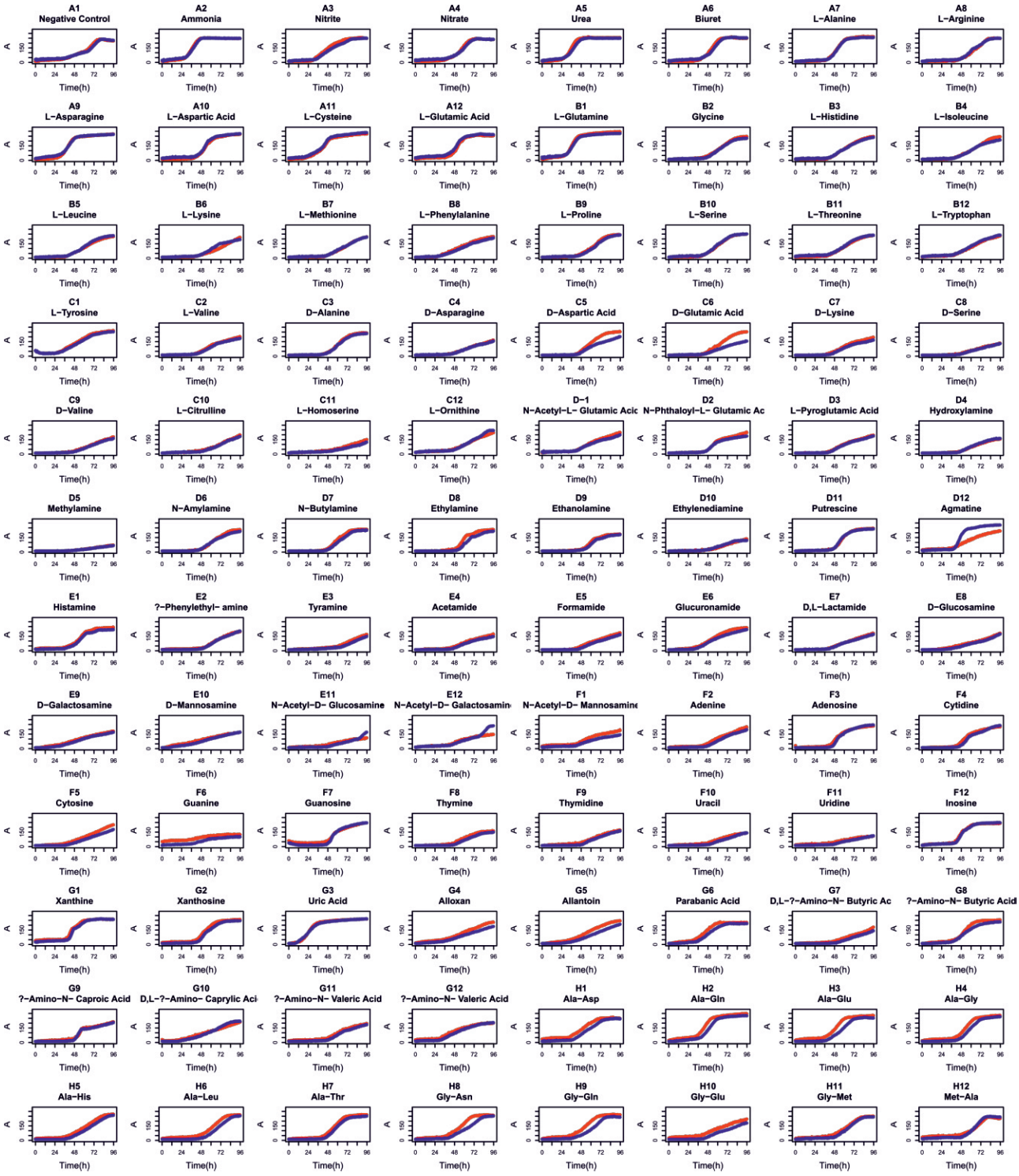
PM1 – Carbon sources I



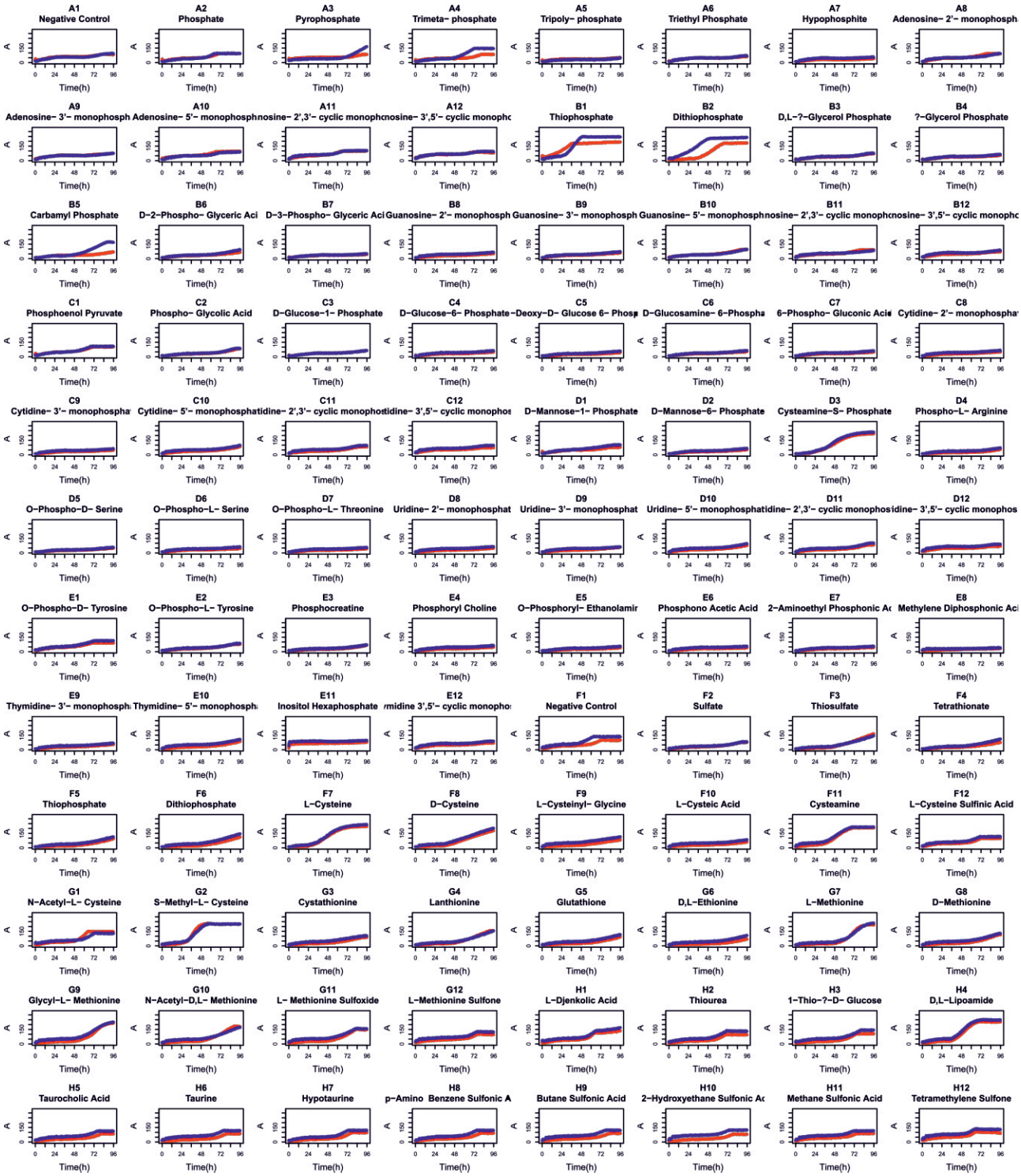
PM2 – Carbon sources II



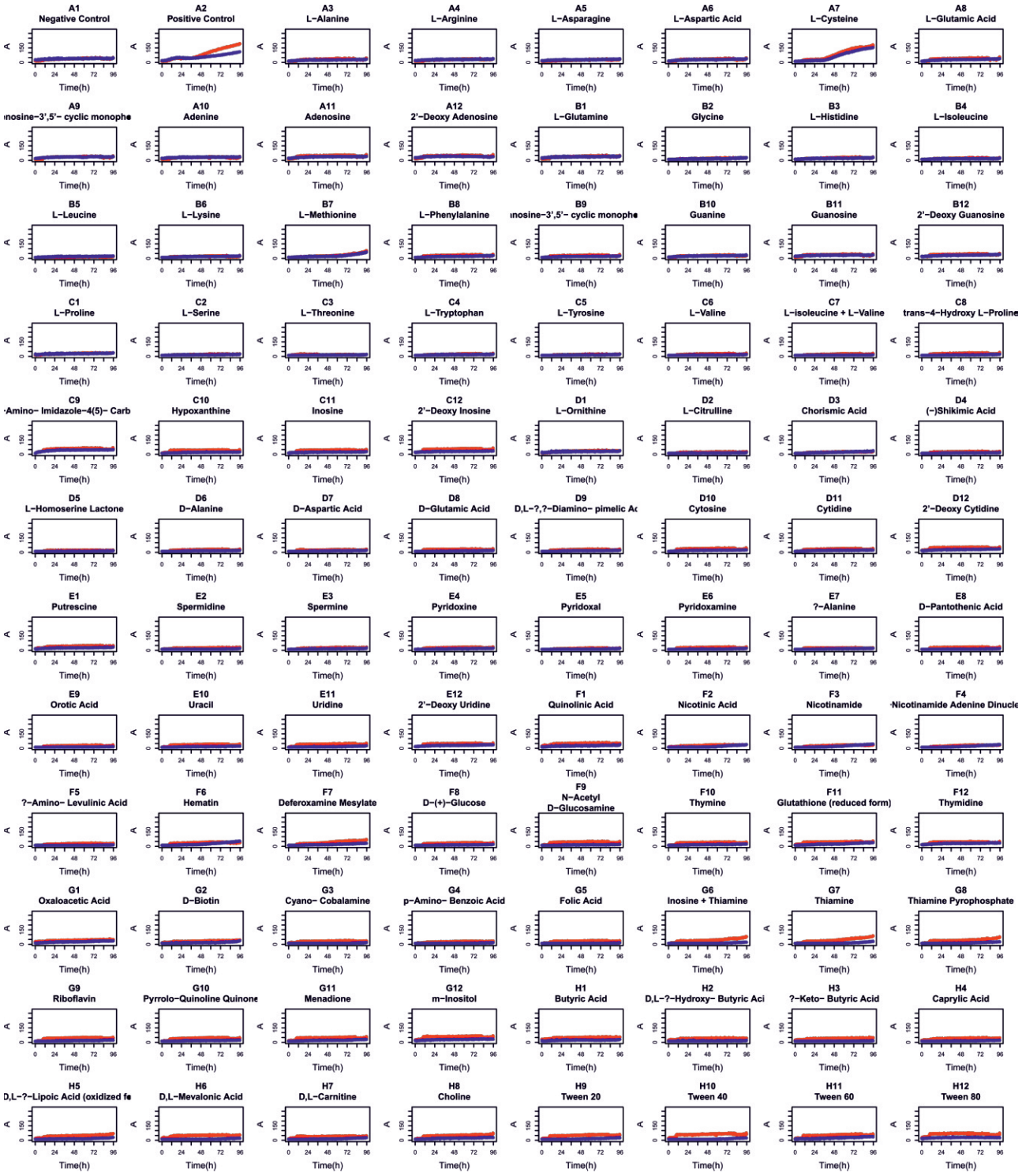
PM3 – Nitrogen sources



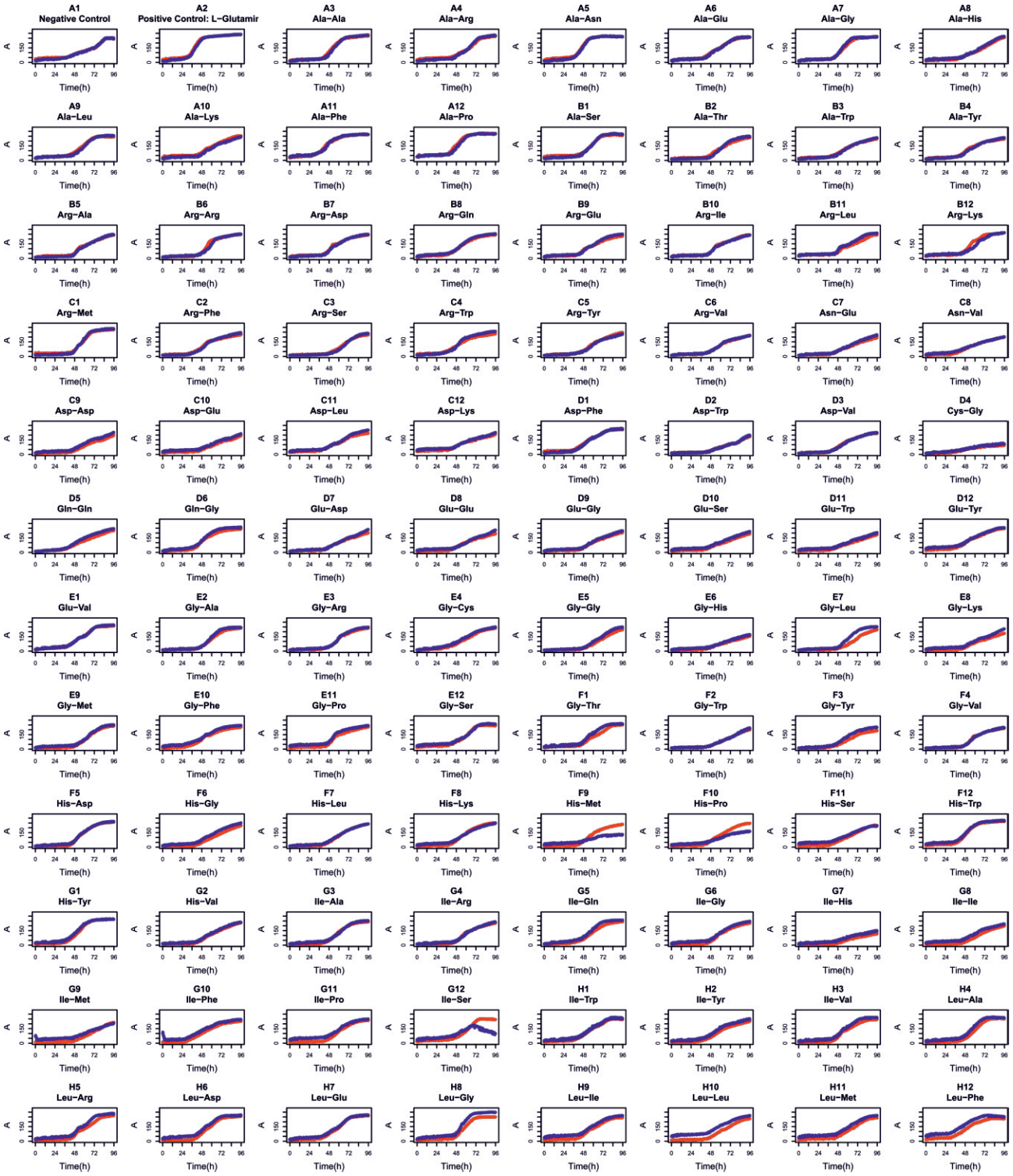
PM4 - Phosphorus and sulfur sources



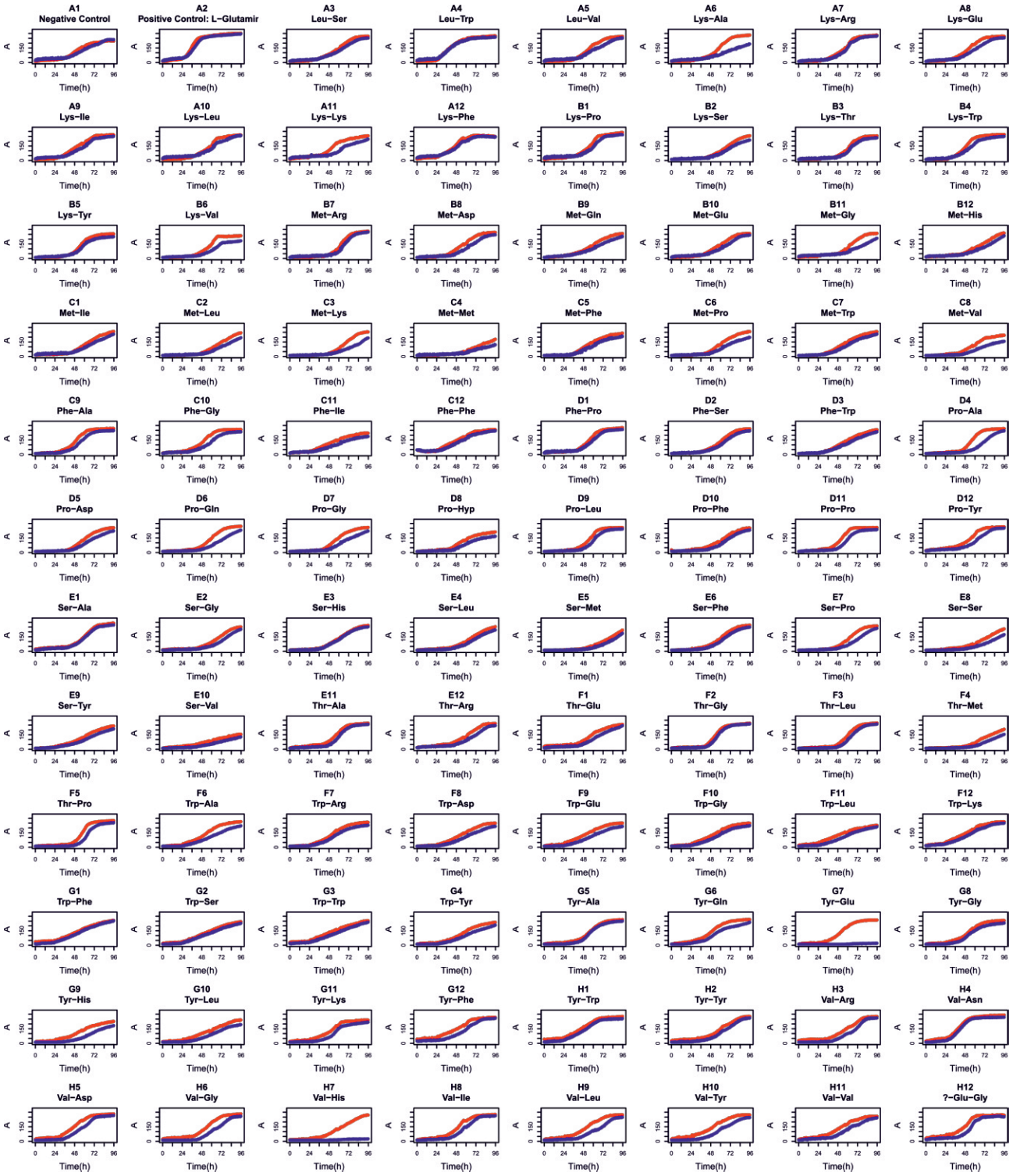
PM5 - Nutrient supplements



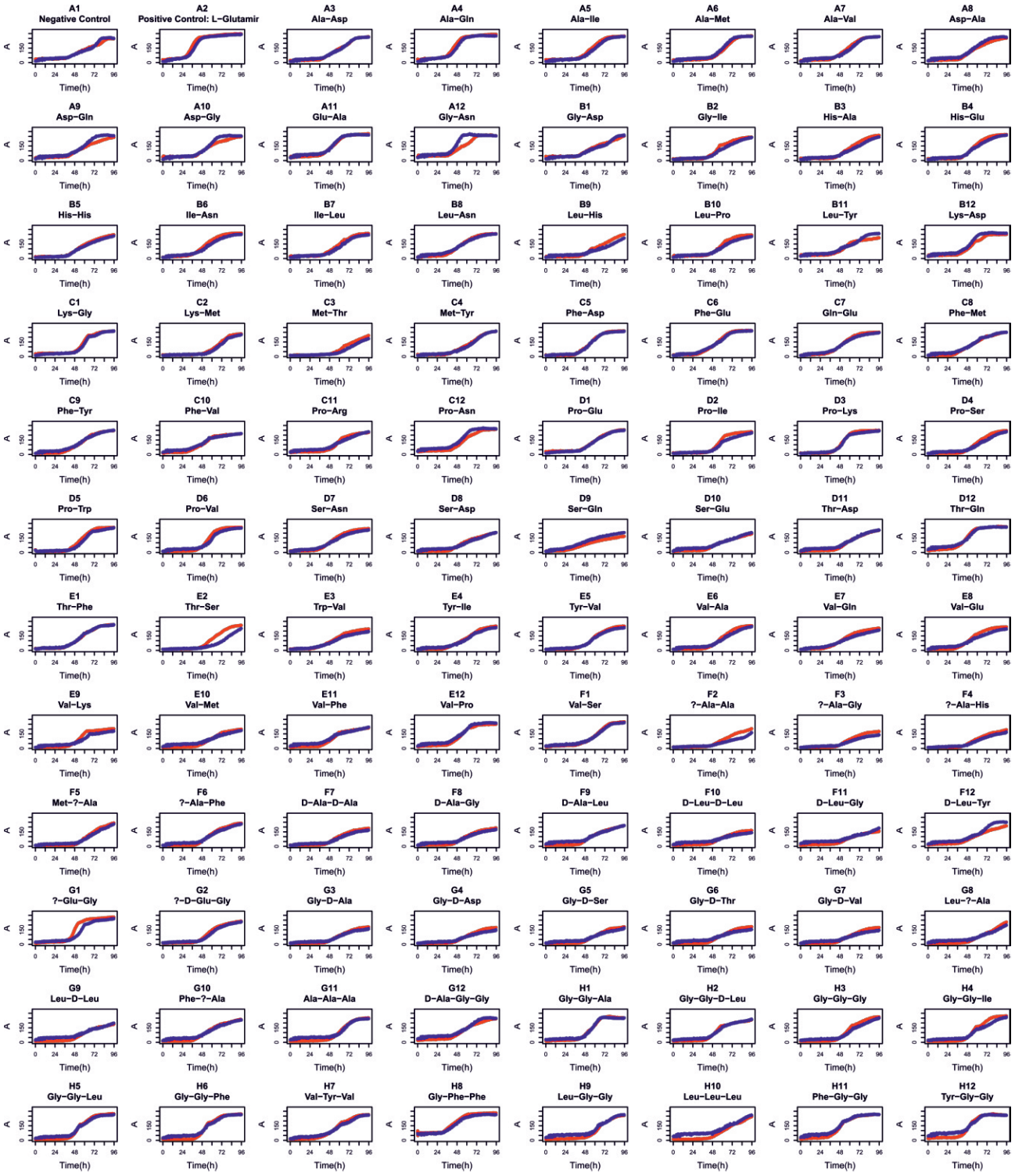
PM6 - Peptide nitrogen sources I



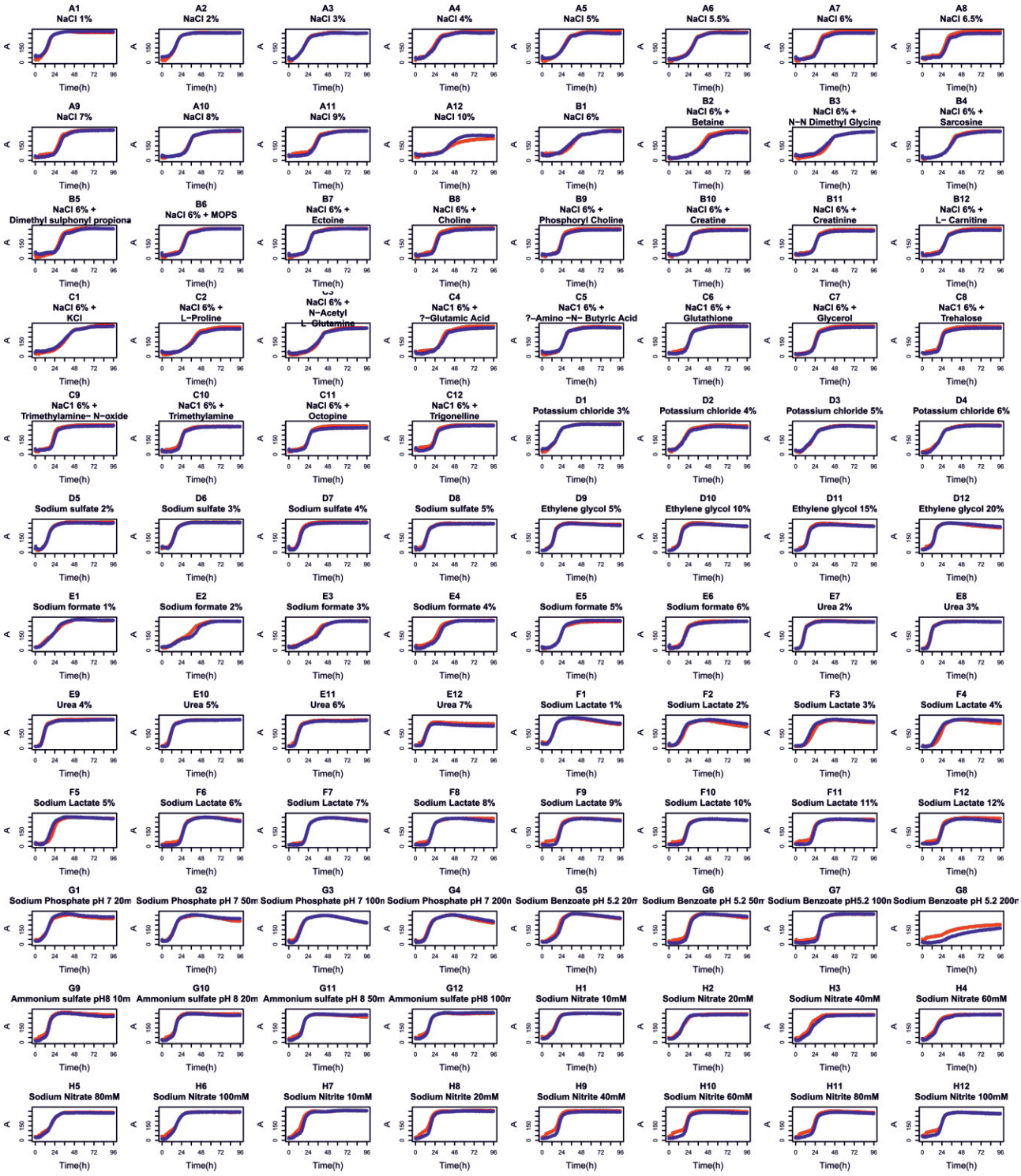
PM7 - Peptide nitrogen sources II



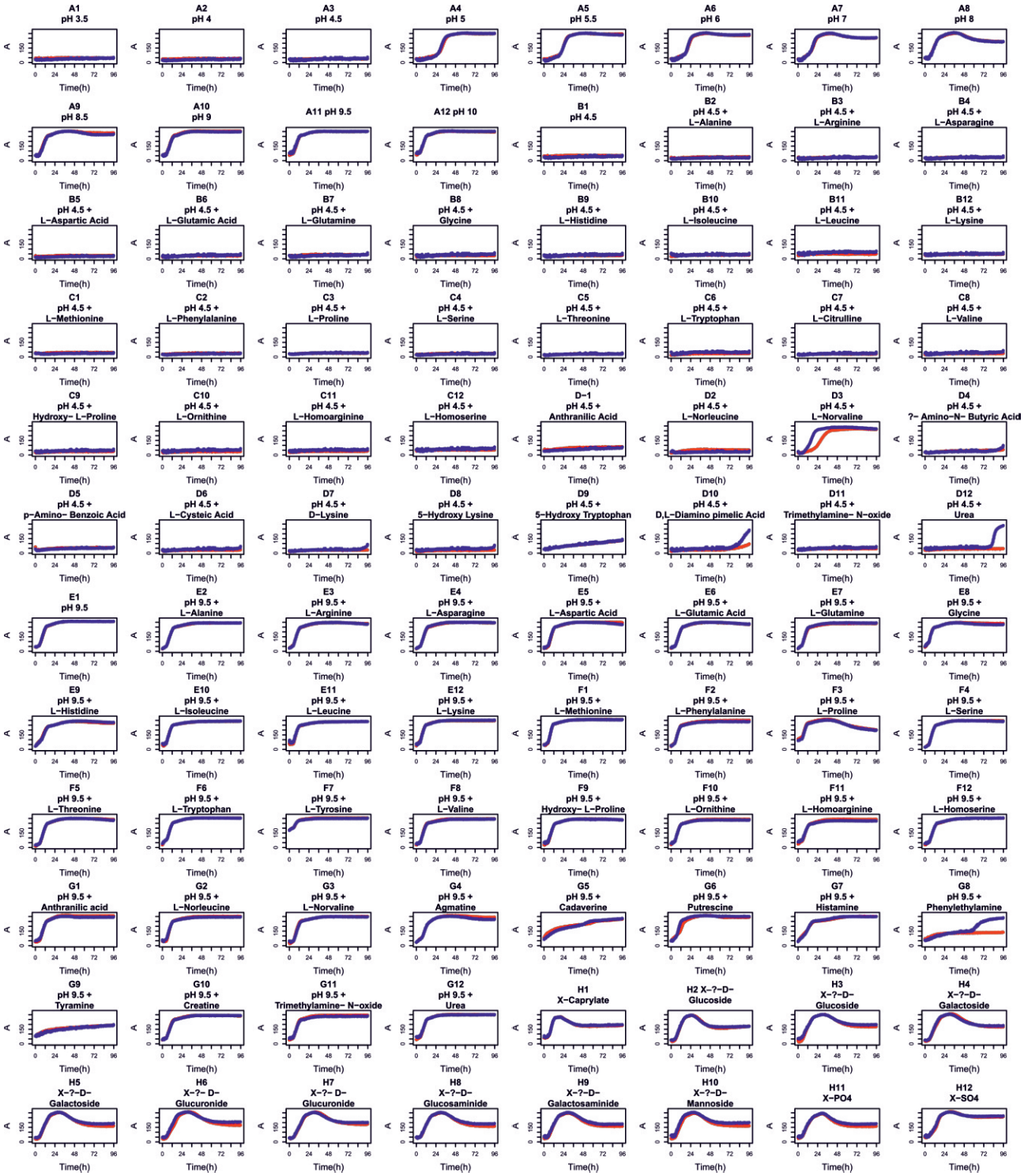
PM8 - Peptide nitrogen sources III



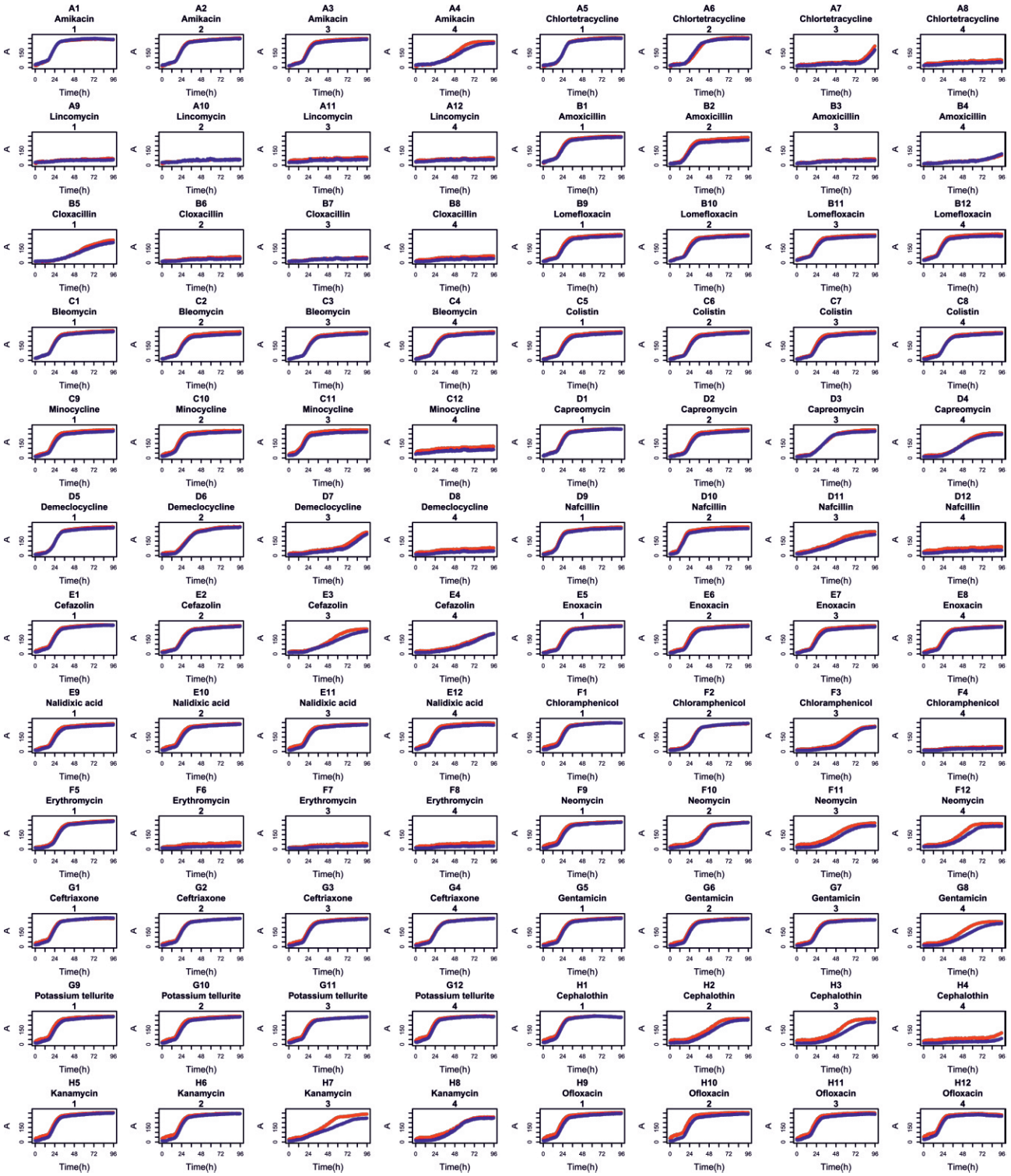
PM9 - Osmolytes



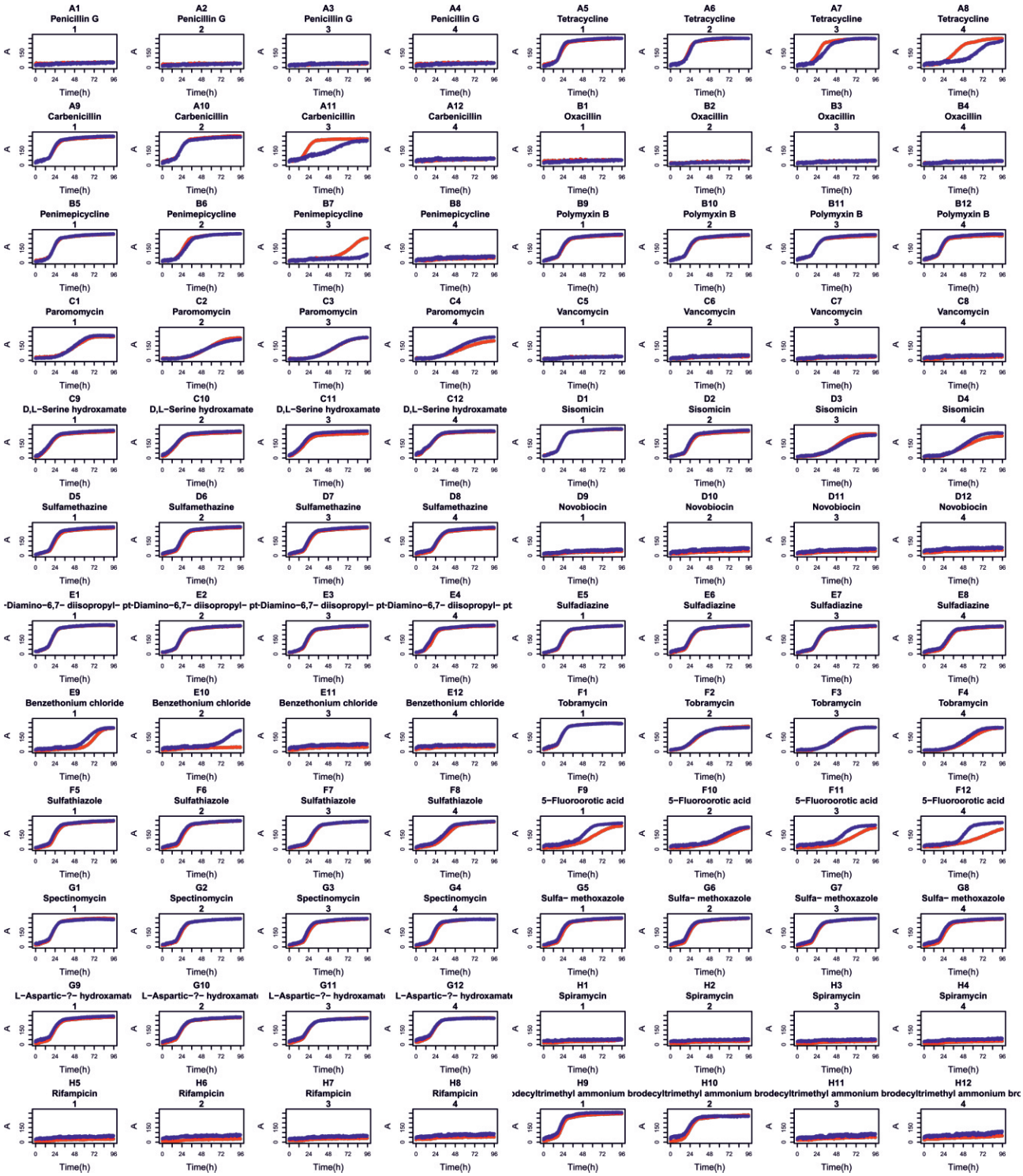
PM10 - pH


A1 pH 3.5
A2 pH 4
A3 pH 4.5
A4 pH 5
A5 pH 5.5
A6 pH 6
A7 pH 7
A8 pH 8
A9 pH 8.5
A10 pH 9
A11 pH 9.5
A12 pH 10
B1 pH 4.5
B2 pH 4.5 + L−Alanine
B3 pH 4.5 + L−Arginine
B4 pH 4.5 + L−Asparagine
B5 pH 4.5 + L−Aspartic Acid
B6 pH 4.5 + L−Glutamic Acid
B7 pH 4.5 + L−Glutamine
B8 pH 4.5 + Glycine
B9 pH 4.5 + L−Histidine
B10 pH 4.5 + L−Isoleucine
B11 pH 4.5 + L−Leucine
B12 pH 4.5 + L−Lysine
C1 pH 4.5 + L−Methionine
C2 pH 4.5 + L−Phenylalanine
C3 pH 4.5 + L−Proline
C4 pH 4.5 + L−Serine
C5 pH 4.5 + L−Threonine
C6 pH 4.5 + L−Tryptophan
C7 pH 4.5 + L−Citrulline
C8 pH 4.5 + L−Valine
C9 pH 4.5 + Hydroxy− L−Proline
C10 pH 4.5 + L−Ornithine
C11 pH 4.5 + L−Homoarginine
C12 pH 4.5 + L−Homoserine
D−1 pH 4.5 + Anthranilic Acid
D2 pH 4.5 + L−Norleucine
D3 pH 4.5 + L−Norvaline
D4 pH 4.5 + ?− Amino−N− Butyric Acid
D5 pH 4.5 + p−Amino− Benzoic Acid
D6 pH 4.5 + L−Cysteic Acid
D7 pH 4.5 + D−Lysine
D8 pH 4.5 + 5−Hydroxy Lysine
D9 pH 4.5 + 5−Hydroxy Tryptophan
D10 pH 4.5 + D,L−Diamino pimelic Acid
D11 pH 4.5 + Trimethylamine− N−oxide
D12 pH 4.5 + Urea
E1 pH 9.5
E2 pH 9.5 + L−Alanine
E3 pH 9.5 + L−Arginine
E4 pH 9.5 + L−Asparagine
E5 pH 9.5 + L−Aspartic Acid
E6 pH 9.5 + L−Glutamic Acid
E7 pH 9.5 + L−Glutamine
E8 pH 9.5 + Glycine
E9 pH 9.5 + L−Histidine
E10 pH 9.5 + L−Isoleucine
E11 pH 9.5 + L−Leucine
E12 pH 9.5 + L−Lysine
F1 pH 9.5 + L−Methionine
F2 pH 9.5 + L−Phenylalanine
F3 pH 9.5 + L−Proline
F4 pH 9.5 + L−Serine
F5 pH 9.5 + L−Threonine
F6 pH 9.5 + L−Tryptophan
F7 pH 9.5 + L−Tyrosine
F8 pH 9.5 + L−Valine
F9 pH 9.5 + Hydroxy− L−Proline
F10 pH 9.5 + L−Ornithine
F11 pH 9.5 + L−Homoarginine
F12 pH 9.5 + L−Homoserine
G1 pH 9.5 + Anthranilic acid
G2 pH 9.5 + L−Norleucine
G3 pH 9.5 + L−Norvaline
G4 pH 9.5 + Agmatine
G5 pH 9.5 + Cadaverine
G6 pH 9.5 + Putrescine
G7 pH 9.5 + Histamine
G8 pH 9.5 + Phenylethylamine
G9 pH 9.5 + Tyramine
G10 pH 9.5 + Creatine
G11 pH 9.5 + Trimethylamine− N−oxide
G12 pH 9.5 + Urea
H1 X−Caprylate
H2 X−?−D− Glucoside
H3 X−?−D− Glucoside
H4 X−?−D− Galactoside
H5 X−?−D− Galactoside
H6 X−?− D− Glucuronide
H7 X−?− D− Glucuronide
H8 X−?−D− Glucosaminide
H9 X−?−D− Galactosaminide
H10 X−?−D− Mannoside
H11 X−PO4
H12 X−SO4
A
Time(h)

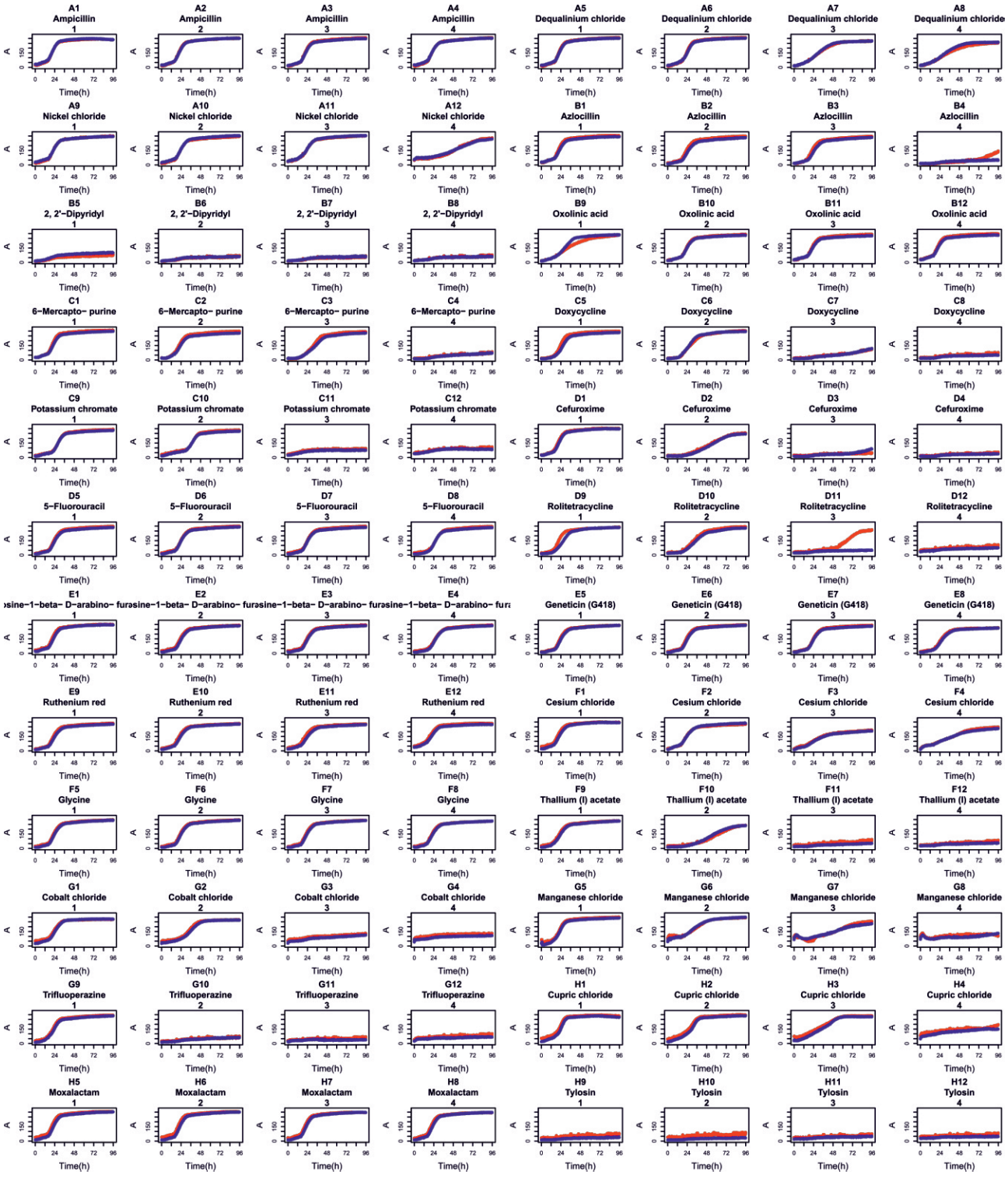
PM11


A1 Amikacin 1
A2 Amikacin 2
A3 Amikacin 3
A4 Amikacin 4
A5 Chlortetracycline 1
A6 Chlortetracycline 2
A7 Chlortetracycline 3
A8 Chlortetracycline 4
A9 Lincomycin 1
A10 Lincomycin 2
A11 Lincomycin 3
A12 Lincomycin 4
B1 Amoxicillin 1
B2 Amoxicillin 2
B3 Amoxicillin 3
B4 Amoxicillin 4
B5 Cloxacillin 1
B6 Cloxacillin 2
B7 Cloxacillin 3
B8 Cloxacillin 4
B9 Lomefloxacin 1
B10 Lomefloxacin 2
B11 Lomefloxacin 3
B12 Lomefloxacin 4
C1 Bleomycin 1
C2 Bleomycin 2
C3 Bleomycin 3
C4 Bleomycin 4
C5 Colistin 1
C6 Colistin 2
C7 Colistin 3
C8 Colistin 4
C9 Minocycline 1
C10 Minocycline 2
C11 Minocycline 3
C12 Minocycline 4
D1 Capreomycin 1
D2 Capreomycin 2
D3 Capreomycin 3
D4 Capreomycin 4
D5 Demeclocycline 1
D6 Demeclocycline 2
D7 Demeclocycline 3
D8 Demeclocycline 4
D9 Nafcillin 1
D10 Nafcillin 2
D11 Nafcillin 3
D12 Nafcillin 4
E1 Cefazolin 1
E2 Cefazolin 2
E3 Cefazolin 3
E4 Cefazolin 4
E5 Enoxacin 1
E6 Enoxacin 2
E7 Enoxacin 3
E8 Enoxacin 4
E9 Nalidixic acid 1
E10 Nalidixic acid 2
E11 Nalidixic acid 3
E12 Nalidixic acid 4
F1 Chloramphenicol 1
F2 Chloramphenicol 2
F3 Chloramphenicol 3
F4 Chloramphenicol 4
F5 Erythromycin 1
F6 Erythromycin 2
F7 Erythromycin 3
F8 Erythromycin 4
F9 Neomycin 1
F10 Neomycin 2
F11 Neomycin 3
F12 Neomycin 4
G1 Ceftriaxone 1
G2 Ceftriaxone 2
G3 Ceftriaxone 3
G4 Ceftriaxone 4
G5 Gentamicin 1
G6 Gentamicin 2
G7 Gentamicin 3
G8 Gentamicin 4
G9 Potassium tellurite 1
G10 Potassium tellurite 2
G11 Potassium tellurite 3
G12 Potassium tellurite 4
H1 Cephalothin 1
H2 Cephalothin 2
H3 Cephalothin 3
H4 Cephalothin 4
H5 Kanamycin 1
H6 Kanamycin 2
H7 Kanamycin 3
H8 Kanamycin 4
H9 Ofloxacin 1
H10 Ofloxacin 2
H11 Ofloxacin 3
H12 Ofloxacin 4
A
Time(h)
0 24 48 72 96
0 150

PM12


A1 Penicillin G 1
A2 Penicillin G 2
A3 Penicillin G 3
A4 Penicillin G 4
A5 Tetracycline 1
A6 Tetracycline 2
A7 Tetracycline 3
A8 Tetracycline 4
A9 Carbenicillin 1
A10 Carbenicillin 2
A11 Carbenicillin 3
A12 Carbenicillin 4
B1 Oxacillin 1
B2 Oxacillin 2
B3 Oxacillin 3
B4 Oxacillin 4
B5 Penimepicycline 1
B6 Penimepicycline 2
B7 Penimepicycline 3
B8 Penimepicycline 4
B9 Polymyxin B 1
B10 Polymyxin B 2
B11 Polymyxin B 3
B12 Polymyxin B 4
C1 Paromomycin 1
C2 Paromomycin 2
C3 Paromomycin 3
C4 Paromomycin 4
C5 Vancomycin 1
C6 Vancomycin 2
C7 Vancomycin 3
C8 Vancomycin 4
C9 D,L−Serine hydroxamate 1
C10 D,L−Serine hydroxamate 2
C11 D,L−Serine hydroxamate 3
C12 D,L−Serine hydroxamate 4
D1 Sisomicin 1
D2 Sisomicin 2
D3 Sisomicin 3
D4 Sisomicin 4
D5 Sulfamethazine 1
D6 Sulfamethazine 2
D7 Sulfamethazine 3
D8 Sulfamethazine 4
D9 Novobiocin 1
D10 Novobiocin 2
D11 Novobiocin 3
D12 Novobiocin 4
E1 ·Diamino−6,7− diisopropyl− pt
E2 Diamino−6,7− diisopropyl− pt
E3 Diamino−6,7− diisopropyl− pt
E4 Diamino−6,7− diisopropyl− pt
E5 Sulfadiazine 1
E6 Sulfadiazine 2
E7 Sulfadiazine 3
E8 Sulfadiazine 4
E9 Benzethonium chloride 1
E10 Benzethonium chloride 2
E11 Benzethonium chloride 3
E12 Benzethonium chloride 4
F1 Tobramycin 1
F2 Tobramycin 2
F3 Tobramycin 3
F4 Tobramycin 4
F5 Sulfathiazole 1
F6 Sulfathiazole 2
F7 Sulfathiazole 3
F8 Sulfathiazole 4
F9 5−Fluoroorotic acid 1
F10 5−Fluoroorotic acid 2
F11 5−Fluoroorotic acid 3
F12 5−Fluoroorotic acid 4
G1 Spectinomycin 1
G2 Spectinomycin 2
G3 Spectinomycin 3
G4 Spectinomycin 4
G5 Sulfa− methoxazole 1
G6 Sulfa− methoxazole 2
G7 Sulfa− methoxazole 3
G8 Sulfa− methoxazole 4
G9 L−Aspartic−?− hydroxamat
G10 L−Aspartic−?− hydroxamat
G11 L−Aspartic−?− hydroxamat
G12 L−Aspartic−?− hydroxamat
H1 Spiramycin 1
H2 Spiramycin 2
H3 Spiramycin 3
H4 Spiramycin 4
H5 Rifampicin 1
H6 Rifampicin 2
H7 Rifampicin 3
H8 Rifampicin 4
H9 ›decyltrimethyl ammonium br
H10 decyltrimethyl ammonium br
H11 decyltrimethyl ammonium br
H12 decyltrimethyl ammonium br
A
Time(h)
0 24 48 72 96
0 150

PM13


A1 Ampicillin 1
A2 Ampicillin 2
A3 Ampicillin 3
A4 Ampicillin 4
A5 Dequalinium chloride 1
A6 Dequalinium chloride 2
A7 Dequalinium chloride 3
A8 Dequalinium chloride 4
A9 Nickel chloride 1
A10 Nickel chloride 2
A11 Nickel chloride 3
A12 Nickel chloride 4
B1 Azlocillin 1
B2 Azlocillin 2
B3 Azlocillin 3
B4 Azlocillin 4
B5 2, 2'−Dipyridyl 1
B6 2, 2'−Dipyridyl 2
B7 2, 2'−Dipyridyl 3
B8 2, 2'−Dipyridyl 4
B9 Oxolinic acid 1
B10 Oxolinic acid 2
B11 Oxolinic acid 3
B12 Oxolinic acid 4
C1 6−Mercapto− purine 1
C2 6−Mercapto− purine 2
C3 6−Mercapto− purine 3
C4 6−Mercapto− purine 4
C5 Doxycycline 1
C6 Doxycycline 2
C7 Doxycycline 3
C8 Doxycycline 4
C9 Potassium chromate 1
C10 Potassium chromate 2
C11 Potassium chromate 3
C12 Potassium chromate 4
D1 Cefuroxime 1
D2 Cefuroxime 2
D3 Cefuroxime 3
D4 Cefuroxime 4
D5 5−Fluorouracil 1
D6 5−Fluorouracil 2
D7 5−Fluorouracil 3
D8 5−Fluorouracil 4
D9 Rolitetracycline 1
D10 Rolitetracycline 2
D11 Rolitetracycline 3
D12 Rolitetracycline 4
E1 ɔsine−1−beta− D−arabino− furɔsine−1−beta− 1
E2 D−arabino− furɔsine−1−beta− 2
E3 D−arabino− furɔsine−1−beta− 3
E4 D−arabino− furɔsine−1−beta− D−arabino− furɔ 4
E5 Geneticin (G418) 1
E6 Geneticin (G418) 2
E7 Geneticin (G418) 3
E8 Geneticin (G418) 4
E9 Ruthenium red 1
E10 Ruthenium red 2
E11 Ruthenium red 3
E12 Ruthenium red 4
F1 Cesium chloride 1
F2 Cesium chloride 2
F3 Cesium chloride 3
F4 Cesium chloride 4
F5 Glycine 1
F6 Glycine 2
F7 Glycine 3
F8 Glycine 4
F9 Thallium (I) acetate 1
F10 Thallium (I) acetate 2
F11 Thallium (I) acetate 3
F12 Thallium (I) acetate 4
G1 Cobalt chloride 1
G2 Cobalt chloride 2
G3 Cobalt chloride 3
G4 Cobalt chloride 4
G5 Manganese chloride 1
G6 Manganese chloride 2
G7 Manganese chloride 3
G8 Manganese chloride 4
G9 Trifluoperazine 1
G10 Trifluoperazine 2
G11 Trifluoperazine 3
G12 Trifluoperazine 4
H1 Cupric chloride 1
H2 Cupric chloride 2
H3 Cupric chloride 3
H4 Cupric chloride 4
H5 Moxalactam 1
H6 Moxalactam 2
H7 Moxalactam 3
H8 Moxalactam 4
H9 Tylosin 1
H10 Tylosin 2
H11 Tylosin 3
H12 Tylosin 4
A
Time(h)

6.3 Supplemental material to: Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations

Supplementary Material to:

Spatiotemporal analysis of SOS and prophage dynamics in *Corynebacterium glutamicum* populations

Stefan Helfrich[1], Eugen Pfeifer[1], Christina Krämer[1], Christian Carsten Sachs, Wolfgang Wiechert, Dietrich Kohlheyer*, Katharina Nöh* and Julia Frunzke*

Supplementary Tables

Table S1. Annotation update of the CGP3 region (cg1890–cg2071) in *C. glutamicum* ATCC 13032.

Locus Tag Cg N°	Locus Tag NCgl N°	Gene name	Annotation, additional information
cg1890	NCgl1611	*alpC*	actin-like protein, (Donovan *et al.*, 2015)
cg1891	NCgl1612	*alpA*	putative phage DNA adapter protein, (Donovan et al., 2015)
cg1892	NCgl1613		hypothetical protein
cg1893	NCgl1614		putative N-acetyltransferase
cg1894	NCgl1615		hypothetical protein
cg1895	NCgl1616		putative secreted protein
cg1896	NCgl1617		putative secreted protein
cg1897	NCgl1618		putative secreted protein
cg1898	NCgl1619		hypothetical protein
cg1899			hypothetical protein
cg1900			hypothetical protein
cg1901			hypothetical protein
cg1902			putative secreted protein
cg1903	NCgl1623		putative ABC-type multidrug transport system, ATPase component
cg1904	NCgl1624		putative membrane protein
cg1905	NCgl1625		hypothetical protein
cg1906			hypothetical protein
cg1907	NCgl1626		putative phosphopantothenoylcysteine synthetase/decarboxylase
cg1908	NCgl1627		hypothetical protein
cg1909			hypothetical protein
cg1910			putative secreted or membrane protein
cg1911	NCgl1629		putative secreted protein

cg1912	NCgl1630		hypothetical protein
cg1913	NCgl1631		hypothetical protein
cg1914			hypothetical protein
cg1915	NCgl1632		hypothetical protein
cg1916	NCgl1633		hypothetical protein
cg1917	NCgl1634		hypothetical protein
cg1918	NCgl1635		putative secreted protein
cg1919	NCgl1636		putative membrane protein
cg1920	NCgl1637		hypothetical protein
cg1921	NCgl1638		hypothetical protein
cg1922	NCgl1639		hypothetical protein
cg1923	NCgl1640		hypothetical protein
cg1924			hypothetical protein
cg1925	NCgl1641		hypothetical protein
cg1926	NCgl1642		hypothetical protein
cg1927	NCgl1643		putative molecular chaperone
cg1928	NCgl1644		hypothetical protein
cg1929	NCgl1645	*res*	resolvase,-family recombinase
cg1930	NCgl1646		putative secreted hydrolase
cg1931	NCgl1647		putative secreted protein
cg1932	NCgl1648	*ppp2*	putative protein phosphatase
cg1934	NCgl1649		hypothetical protein
Intergenic region		*arnA*	small antisense RNA, (Zemanova *et al.*, 2008)
cg1935	NCgl1650	*gntR2*	gluconate-responsive repressor 2, repressor of genes involved in gluconate catabolism and the pentose phosphate pathway, GntR-family, (Frunzke *et al.*, 2008)
cg1936	NCgl1651		putative secreted protein
cg1937	NCgl1652		putative secreted protein
cg1938	NCgl1653		hypothetical protein
cg1940	NCgl1654		putative secreted protein
cg1941	NCgl1655		putative secreted protein
cg1942	NCgl1656		putative secreted protein
cg1943			hypothetical protein
cg1944			hypothetical protein
cg1945	NCgl1658		conserved hypothetical protein
cg1946	NCgl1659		hypothetical protein
cg1947	NCgl1660		hypothetical protein
cg1948			hypothetical protein
cg1949	NCgl1661		hypothetical protein
cg1950	NCgl1662	*tnp14b*	transposase fragment
cg1951	NCgl1664	*tnp14a*	transposase fragment
cg1954	NCgl1665		hypothetical protein
cg1955	NCgl1666		putative secreted protein
cg1956	NCgl1667	*recJ*	single-stranded-DNA-specific exonuclease
cg1957	NCgl1668		hypothetical protein
cg1959	NCgl1669	*priP*	prophage DNA primase
cg1960	NCgl1670		hypothetical protein
cg1961	NCgl1671		hypothetical protein
cg1962	NCgl1672		putative membrane protein
cg1963	NCgl1673		putative superfamily II DNA/RNA helicase

cg1964	NCgl1674		hypothetical protein
cg1965	NCgl1675		putative protein, similarity to gp57-phage N15
cg1966	NCgl1676	*cgpS*	nucleoid-associated protein, Lsr2-type
cg1967	NCgl1677		hypothetical protein
cg1968	NCgl1678		hypothetical protein
cg1969	NCgl1679		hypothetical protein
cg1970	NCgl1680		hypothetical protein
cg1971			hypothetical protein
cg1972			putative translation elongation factor, GTPase
cg1974	NCgl1682	*lys*	putative protein, contains peptidoglycan-binding LysM domain
cg1975	NCgl1683		hypothetical protein, conserved
cg1976	NCgl1684		hypothetical protein
cg1977	NCgl1685		putative secreted protein
cg1978	NCgl1686		hypothetical protein
cg1980	NCgl1687		putative MoxR-like ATPase
cg1981	NCgl1688		hypothetical protein
cg1982	NCgl1689		putative ATPase with chaperone activity, ATP-binding subunit
cg1983	NCgl1690		hypothetical protein
cg1984	NCgl1691		hypothetical protein
cg1985	NCgl1692		putative superfamily I DNA or RNA helicase
cg1986	NCgl1693		hypothetical protein
cg1987	NCgl1694		hypothetical protein
cg1988	NCgl1695		hypothetical protein
cg1989	NCgl1696		hypothetical protein
cg1990	NCgl1697		putative NUDIIX hydrolase
cg1991	NCgl1698		putative protein, similar to gp52-bacteriophage PHIC31
cg1992	NCgl1699		hypothetical protein
cg1993	NCgl1700		hypothetical protein
cg1994	NCgl1701		hypothetical protein
cg1995	NCgl1702		hypothetical protein
cg1996	NCgl1703	*cglIM*	modification methylase, DNA cytosine-5--methyltransferase EC:2.1.1.37, (Schäfer *et al.*, 1994)
cg1997	NCgl1704	*cglIR*	type II restriction endonuclease
cg1998	NCgl1705	*cglIIR*	restriction endonuclease
cg1999	NCgl1706		hypothetical protein
cg2000	NCgl1707		putative membrane protein
cg2001	NCgl1708		conserved hypothetical protein
cg2002	NCgl1709		hypothetical protein
cg2003	NCgl1710		hypothetical protein, conserved // putative terminase predicted by Virfam
cg2004	NCgl1711		conserved putative protein, similar to 232 protein-lactobacillus bacteriophage g1e
cg2005	NCgl1712		conserved putative protein-plasmid encoded
cg2006	NCgl1713		hypothetical protein
cg2007	NCgl1714		putative membrane protein
cg2008	NCgl1715		putative membrane protein
cg2009	NCgl1716		putative CLP-family ATP-binding protease
cg2010	NCgl1717		putative permease of the major facilitator superfamily
cg2011	NCgl1718		putative membrane protein
cg2012	NCgl1719		putative secreted protein

cg2014	NCgl1720		hypothetical protein
cg2015	NCgl1721		hypothetical protein
cg2016	NCgl1722		hypothetical protein
cg2017	NCgl1723		hypothetical protein
cg2018	NCgl1724		putative membrane protein
cg2019	NCgl1725		putative membrane protein
cg2020	NCgl1726		putative membrane protein
cg2021	NCgl1727		putative protein, similar to p18, bacteriophage CP-1 *Streptococcus pneumoniae*
cg2022	NCgl1728		putative secreted protein
cg2023			putative membrane protein
cg2024	NCgl1730		putative nuclease subunit of the excinuclease complex
cg2025			hypothetical protein
cg2026			hypothetical protein
cg2027			hypothetical protein
cg2028	NCgl1734		hypothetical protein
cg2029			hypothetical protein
cg2030	NCgl1735		hypothetical protein
cg2031	NCgl1736		conserved hypothetical protein
cg2032	NCgl1737		putative membrane protein
cg2033	NCgl1738		putative secreted protein
cg2034	NCgl1739		hypothetical protein
cg2035	NCgl1740		putative methyltransferase
cg2036			putative secreted protein
cg2037	NCgl1742		conserved hypothetical protein
cg2038			hypothetical protein
cg2039	NCgl1744		hypothetical protein
cg2040	NCgl1745		putative transcriptional regulator, Cro/CI-family
cg2041	NCgl1746		hypothetical protein
cg2042	NCgl1747		putative secreted protein
cg2043	NCgl1748		hypothetical protein, conserved // putative major capsid protein predicted by Virfam
cg2044	NCgl1749		putative secreted protein
cg2045			hypothetical protein
cg2046			hypothetical protein
cg2047	NCgl1751		putative secreted protein
cg2048	NCgl1752		hypothetical protein
cg2049	NCgl1753		hypothetical protein
cg2050	NCgl1754		hypothetical protein
cg2051			hypothetical protein
cg2052	NCgl1756		putative secreted protein
cg2053			putative membrane protein
cg2054			putative membrane protein
cg2055			putative membrane protein
cg2056	NCgl1760		putative membrane protein
cg2057	NCgl1761		putative secreted protein
cg2058			hypothetical protein
cg2059	NCgl1763		putative secreted protein
cg4007	NCgl1764		hypothetical protein
cg2060	NCgl1765		hypothetical protein
cg2061	NCgl1766	*psp3*	putative secreted protein

cg2062	NCgl1767		putative protein, similar to plasmid-encoded protein PXO2.09
cg2063	NCgl1768		putative membrane protein
cg2064	NCgl1769		putative DNA topoisomerase I omega-protein EC:5.99.1.2
cg2065	NCgl1770		putative superfamily II DNA or RNA helicase
cg2066	NCgl1771		putative low-complexity protein
cg2067	NCgl1772		hypothetical protein
cg2068	NCgl1814		hypothetical protein
cg2069	NCgl1774	*psp1*	putative secreted protein
cg2070		'*int2*	putative phage integrase C-terminal fragment
cg2071	NCgl1816	*int2'*	putative phage integrase N-terminal fragment

11

12

Table S2. Oligonucleotides used in this study.

Overlapping sequences used for Gibson assembly or overlap PCR are marked in bold; restriction sites are marked in italic.

Oligonucleotide	Sequence (5'→3')
PrecA-fw	**TCCTCTTGCTCGTGTCAATT** TTCGCATGAAAATTCGAATT
PrecA-rv-OL-venus	**CTCCTCGCCCTTGCTCACCAT** ATGTATATCTCCTT
Plys-fw	GGAATT *CAATTG* CCTTCTTTGAGGCTTGATGC
Plys-rv	**GCTCACCATATGTATATCTCCTT** ATTTTTCGGCATTGCGC
Venus-fw	ATGGTGAGCAAGGGCGAG
Venus-rv	**GTGTCCATGAGTTCGCTCGA** TTACTTGTACAGCTCGTCCATGC
eYFP-fw	**AAGGAGATATACAT** ATGGTGAGCAAGGGCGA
eYFP-rv	**GGAATT *CTCGAG* TTATCTAGACTTGTACAGCTCGTCCAT**
LacI-fw-pK18	**TCCTCTTGCTCGTGTCAATT** CTCAAGCCTTCGTCACTGG
Ptac-rv-OL-venus	**CTCGCCCTTGCTCACCAT** TGTATATCTCCTTCTGCAGGCATG
Cg1121-Int-fw	AGCACCTTCGGCAAGAAGTA
Cg1122-Int -rv	CATCGAAGGTGTCGCAAAC
LacI-fw-pJC1	**AGCGACGCCGCAGGGGGATC** CTCAAGCCTTCGTCACTGG
Ptac-rv-OL-crimson	**GATGACGTTCTCAGTGCTATCCAT** TGTATATCTCCTTCTGCAGGCATG
Crimson-fw	ATGGATAGCACTGAGAACGTCATC
Crimson-rv-OL-pJC1	**AAAACGACGGCCAGTACTAG** CTACTGGAACAGGTGGTGG
D_recA_1	**GACCATGATTACGCCAAGCT** CCATGCGTTTATCGGAGTTTCG
D_recA_2	**CCCATCCACTAAACTTAAACA** TGTCTTCTTGGGAGCCATTT
D_recA_3	**TGTTTAAGTTTAGTGGATGGG** TCCTCCGTCATTGACTTGGC
D_recA_4	**GGTACCCGGGGATCCTCTAG** CCTGGTCGATTTGCTCAAGC
D_recA_5	GGACATTCCAGCCGGAAATC
D_recA_6	GCGTGAAGTCATGAGTAAGCTC

Supplementary Figures

Fig. S1. Validation of reporter constructs in *C. glutamicum* wild type and Δ*recA* mutant.

A. Cells of the wild type strain *C. glutamicum* ATCC 13032 and the *recA* mutant strain containing either pJC1-P_{recA}-*crimson* or pJC1-P_{lys}-*crimson* were cultivated with and w/o 600 µM MMC for six hours in CGXII minimal media with 2 % glucose. Strains containing the empty plasmid served as reference. Measurements of the fluorescence output were performed using flow cytometry (FACSAria II, BD). The presented values represent the average of three biological replicates including the standard deviation. Both reporter constructs show a significant increase upon treatment with MMC in the wild type background. An increase in reporter signal is still observed in a *recA* mutant. However, this accumulation of reporter protein upon MMC treatment might be a result of the strongly reduced growth rate of a *recA* mutant (data not shown), which is, in addition, also impaired in recombination as a prominent DNA repair strategy. **B.** Flow cytometry analysis of indicated strains (cultivations were performed as described above). Red gates in the scatter plots cover the area in which induced cells are observed. Spontaneously induced cells arise under non-inducing conditions in wild type (0.18 %) as well as in Δ*recA* (0.086 %) populations suggesting RecA-independent SPI in *C. glutamicum*. **Please note:** The promoter for the design of the phage reporter (P_{lys}, cg1974) was chosen on the basis of a previous transcriptome analysis showing the induction of CGP3 promoters upon prophage activation (Donovan et al., 2015). The expression of cg1974 significantly increased after 3 hours upon treatment with MMC in comparison to untreated cells (1 h: 1.7-fold; 3 h: 52-fold, 6 h: 1460-fold).

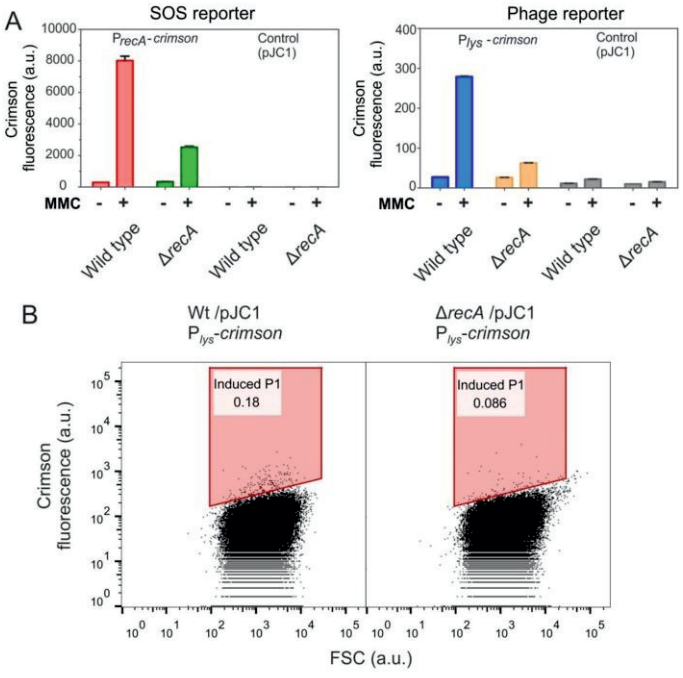


Fig. S2. Time delay of fluorophore maturation.

To test for the difference in induction and maturation time of the fluorescent proteins Venus and Crimson, we have constructed a *C. glutamicum* ATCC 13012 strain in which the expression of both reporter genes is under control of the *tac* promoter. Microfluidic cultivations have been performed in the presence of 100 µg/ml IPTG to induce reporter production. The figure shows the mean values and standard deviations of fluorescence intensities for both reporters. These data indicate, that only a minor delay (ca. 12 min) is observed in the maturation time of the Crimson fluorophore.

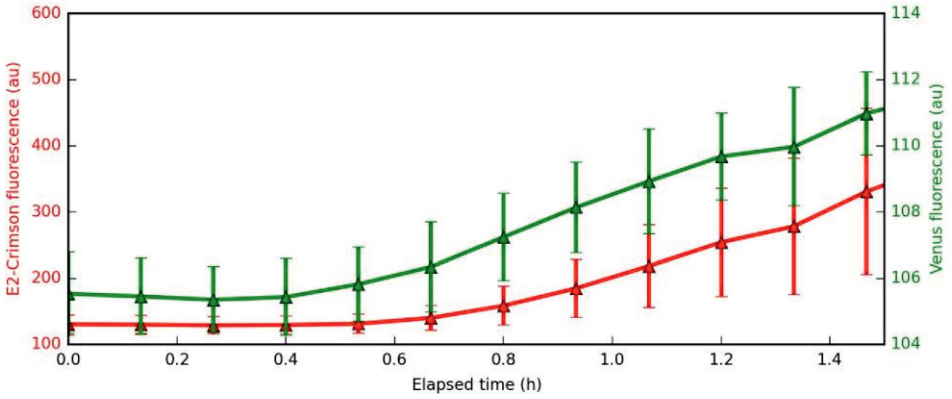


Fig. S3. Growth of the evaluated microcolonies.

Cells were grown in microfluidic chip device in CGXII minimal medium with 2% glucose (for a detailed description, see material and methods). An average growth rate of 0.47 h^{-1} was observed.

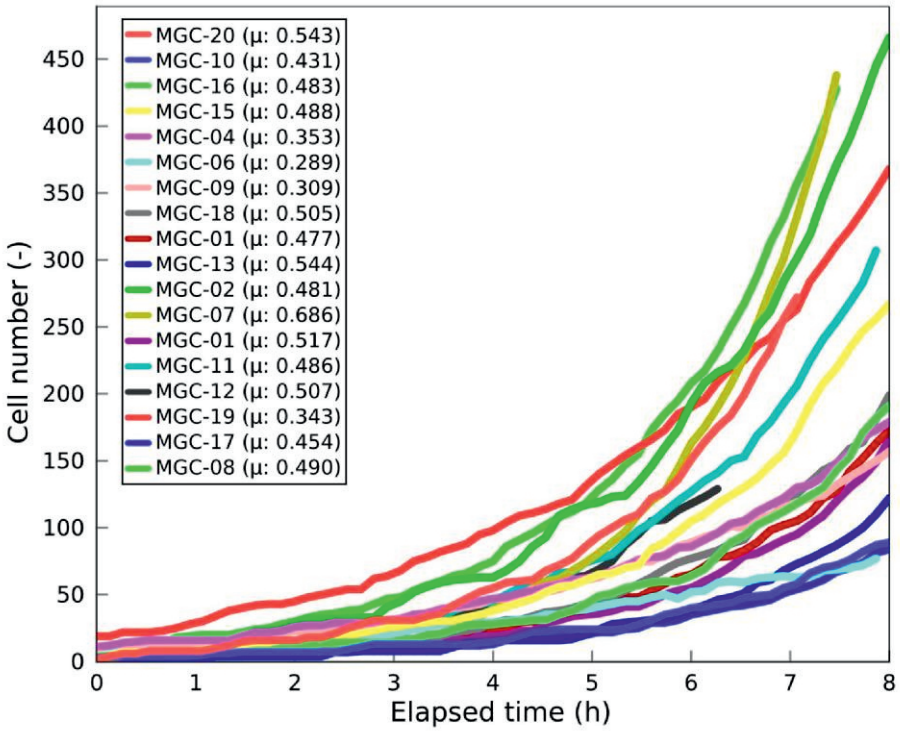


Fig. S4. Setting of thresholds for fluorescence data.

Biological as well as technical noise may result in small fluctuations of fluorescence reporter outputs. Applying a single threshold would result in many false-positive state transitions when the signal fluctuates around the threshold. To prevent such effects, an *UNDEFINED* state around the selected threshold (T) was established (between 95 % and 105 % of T) in which the temporal context of a cell is taken into account to determine the cells' correct state: Cells that enter but leave the *UNDEFINED* state to an *OFF* state are not counted as entering the *ON* state. A cell is assigned the ON state at the time point where the fluorescence reporter intensity exceeds the upper threshold of the *UNDEFINED* state. The same applies for cells entering the *UNDEFINED* state from the *ON* state and for the lower threshold of the UNDEFINED state. Appropriate thresholds for Venus (T_{SOS}) and E2-Crimson (T_{Phage}) intensities were set, for both reporters separately. T_{SOS} is more than 3 standard deviations above the mean of the SOS uninduced population (mean: 117.31 a.u., standard deviation: 3.85 a.u., Fig. 4B on y-axis). T_{Phage} was set to 160 a. u. (mean: 110.44 a.u., standard deviation: 2.44 a.u., Fig. 4D).

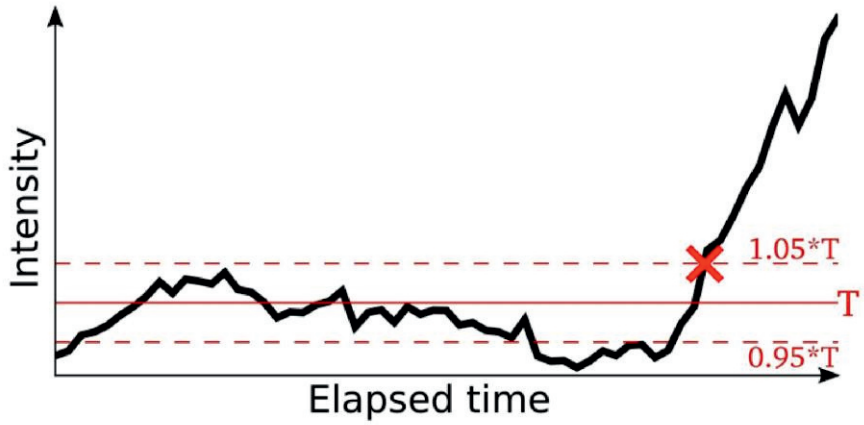


Fig. S5. Influence of T_{SOS} on transitions probabilities.

The graph shows the influence of T_{SOS} definition on transition probabilities regardless of the Offspring state of cells. Changes is transition probabilities are negligible when increasing the threshold to be greater than the value $T_{SOS} = 130$ a.u.

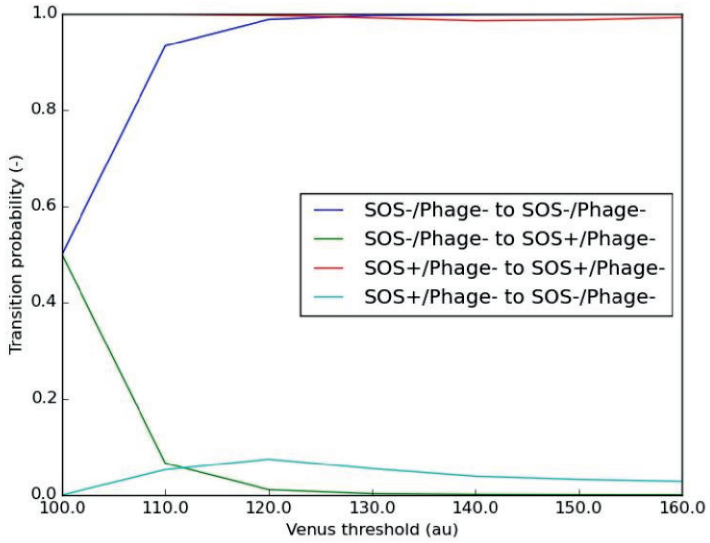
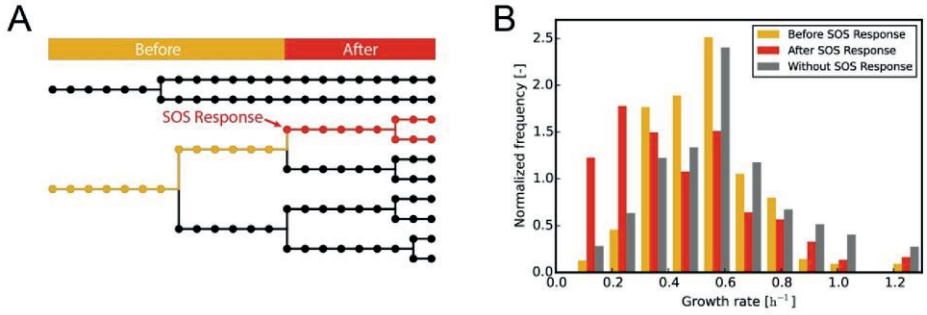


Fig. S6. History and future of SOS+ cells.

Analysis of the recorded lineage data in order to compare the growth of SOS-induced with non-induced cells prior to the induction event. The question we were addressing was: Do cells which will suffer from SOS induction in the future show an altered growth already before the induction of the stress response? For example, are fast growing cells even more prone to spontaneous SOS? Or in contrast, do we even see a reduction in growth rate since the cell is already suffering from e. g. oxidative stress?

A. Schematic lineage tree illustrating the dissection of single traces in a "before" and "after" state upon SOS induction. **B.** Distribution of the growth rate of cells "before" and "after" SOS induction. These values are compared to cells which did not exhibit SOS induction (grey bar) throughout the experiment. Growth rates μ were computed as $\mu = \ln(2)/t_{gen}$. For complete generations, t_{gen} is defined as the time between two cell divisions. For incomplete generations, the end of the experiment is defined as cell division. Consequently, a growth rate >0 is computed even for non-growing cells.



Supplementary Videos

Video S1. Time-lapse video of a selected microcolony (MGC-16).

Cells of the dual reporter strain *C. glutamicum* ATCC 13032::P_{recA}-*venus*/ pJC1-P_{lys}-*e2-crimson* were cultivated in monolayer growth chambers with a continuous flow of CGXII minimal media with 2 % glucose and 25 µg/ ml kanamycin for 20 h at 30 °C. The videos show the first 11 to 13 hours of an experiment. The lineage tree extracted from Video S1 is shown in Fig. 2.

Video S2. Time-lapse video of a selected microcolony (MGC-10).

See caption of Video S1.

Video S3. Time-lapse video of a selected microcolony (MGC-15).

See caption of Video S1.

Video S4. Time-lapse video of a selected microcolony showing RecA-independent SPI in *C. glutamicum*.

Cells of the strain *C. glutamicum* ATCC 13032 Δ*recA*::P_{lys}-*eyfp* were cultivated in microfluidic monolayer growth chambers under continuous flow of CGXII minimal media with 2 % glucose and 25 µg/ml kanamycin for 16 h at 30 °C. The video shows the first 12 hours of an experiment.

Video S5. Time-lapse video of a selected microcolony of the carbon-starvation experiment.

Cells of the dual reporter strain *C. glutamicum* ATCC 13032::P_{recA}-*venus*/ pJC1-P_{lys}-*e2-crimson* were cultivated in monolayer growth chambers. For initiation of the starvation phase the medium was switched to minimal medium CGXII lacking glucose and protocatechuate after 8 h. Cells were exposed to carbon limitation for ~24 h. During the starvation phase, a part of the population is zoomed out to highlight a cell displaying SOS induction during carbon source starvation.

Video S6. Long-term study of SOS induction and SPI in exponentially growing cells.

Cells of the dual reporter strain *C. glutamicum* ATCC 13032::P_{recA}-*venus*/ pJC1-P_{lys}-*e2-crimson* grown in the "mother machine" microfluidic structure. Cells were cultivated under continuous flow of CGXII minimal media with 2 % glucose and 25 µg/ml kanamycin for 24 h at 30 °C. The rate of reporter induction is shown in Fig. 7.

Video S7. Processed time-lapse video of MGC-16.

Video S1 augmented with yellow contours denoting detected cells that have been used for further processing.

123 ## References

124 Donovan, C., Heyer, A., Pfeifer, E., Polen, T., Wittmann, A., Krämer, R., Frunzke, J., and Bramkamp, M.
125 (2015) A prophage-encoded actin-like protein required for efficient viral DNA replication in
126 bacteria. *Nucleic Acids Res* **43**, 5002-5016
127 Frunzke, J., Engels, V., Hasenbein, S., Gatgens, C., and Bott, M. (2008) Co-ordinated regulation of
128 gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two
129 functionally equivalent transcriptional regulators, GntR1 and GntR2. *Mol Mirobiol* **67**, 305-
130 322
131 Schäfer, A., Schwarzer, A., Kalinowski, J., and Pühler, A. (1994) Cloning and characterization of a DNA
132 region encoding a stress-sensitive restriction system from *Corynebacterium glutamicum*
133 ATCC 13032 and analysis of its role in intergeneric conjugation with *Escherichia coli*. *J*
134 *Bacteriol* **176**, 7309-7319
135 Zemanova, M., Kaderabkova, P., Patek, M., Knoppova, M., Silar, R., and Nesvera, J. (2008)
136 Chromosomally encoded small antisense RNA in *Corynebacterium glutamicum*. *FEMS*
137 *Microbiol Lett* **279**, 195-201

138

139

6.4 Supplemental material to: Silencing of cryptic prophages in *Corynebacterium glutamicum*

Supplementary Material to

Silencing of cryptic prophages in *Corynebacterium glutamicum*

Eugen Pfeifer[1], Max Hünnefeld[1], Ovidiu Popa[2], Tino Polen[1], Dietrich Kohlheyer[1], Meike Baumgart[1], and Julia Frunzke[1,*]

Supplementary Tables

Table S1. Strains and plasmids used in this study.

Strains	Relevant characteristics	Reference
E. coli		
DH5α	*supE44* Δ*lacU169* (φ80*lacZ*DM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*, strain used for cloning procedures	Invitrogen
BL21(DE3)	F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm* BL21(DE3), protein production host	(1)
S3974	Derivate of K-12 (CGSC #6300), F^-, λ^-, rph^+ $ilvG^+$	(2)
T221	S3974 $\Delta hns_{FRT,,}$ *E. coli* strain used for complementation studies of Δ*hns* phenotype	(3)
M. tuberculosis		
H37Rv	wild-type laboratory strain, DNA used as PCR template	ATCC 25618
C. diphtheriae		
ATCC 27010	wild-type laboratory strain, DNA used as PCR template	DSM 44123
C. amycolatum		
PAP 272	wild-type, genomic DNA was used for PCR as template	DSM 44737

C. glutamicum		
ATCC 13032	Biotin-auxotrophic wild type	(4)
WT::*cgpS-strep*	Derivative of ATCC 13032 with genomic exchange of the *cgpS* gene to *cgpS-strep*, encoding a C-terminal Strep-tag fusion.	This study
ATCC 13032 ΔCGP3	ATCC 13032 with in-frame deletion of prophage CGP3 (cg1890-cg2071)	(5)
WT::P_{lys} -*eyfp*	Derivative of ATCC 13032 containing the prophage reporter P_{lys}-*eyfp* integrated into the intergenic region of cg1121-cg1122	(6)
Plasmids		
pAN6	*Kan*^{R}*; C. glutamicum*/*E. coli* shuttle vector for gene expression under control of the *tac* promoter; (P_{tac}, *lacI*^{q}, *pBL*$_1$ *ori*$V_{C.g.}$, pUC18 *ori*$V_{E.c.}$)	(7)
pAN6-*cgpS*	Derivative of pAN6 containing the *cgpS* gene	This study
pAN6-*cgpS-Strep*	Derivative of pAN6 containing the *cgpS* gene without stop codon encoding a C-terminal Strep-tag fusion	This study
pAN6-*N-cgpS*	Derivative of pAN6 containing the first 65 amino acids of the *cgpS* gene	This study
pAN6-*N-cgpS-Strep*	Derivative of pAN6 containing the first 65 amino acids of the *cgpS* gene fused C-terminally to a Strep-tag coding region	This study
pAN6-*lsr2-N-M.tub*	Derivative of pAN6 containing the first 58 amino acids of the *lsr2* gene (Rv3597c) of *Mycobacterium tuberculosis* H37Rv	This study
pAN6-*cgpS-N-C.amyc*	Derivative of pAN6 containing the first 66 amino acids of the homologous *cgpS* gene (CORAM0001_2081) of *Corynebacterium amycolatum* DSM 44373	This study
pAN6-*cgpS-N-C.diph*	Derivative of pAN6 containing the first 59 amino acids of the homologous *cgpS* gene (DIP2266) of *Corynebacterium diphtheriae* DSM 44123	This study

pAN6-*alpA-eyfp*	Derivative of pAN6 containing a *alpA-eyfp* fusion	This study
pK19*mobsacB*	Kan^R; plasmid for allelic exchange in *C. glutamicum*; (pK18 *ori*$V_{E.c.}$, *sacB*, *lacZ*α)	(8)
pK19mobsacB-*cgpS-Strep*	Derivative of pK19*mobsacB* containing the *cgpS-Strep* construct for the allelic exchange of the native *cgpS* gene to a C-terminally strep-tagged version in the chromosome of *C. glutamicum*.	This study
pK18*mobsacB*-	Kan^R; plasmid for integration of foreign DNA into the intergenic region between cg1121-cg1122 (*ori*$V_{E.c.}$, *sacB*, *lacZ*α).	(5)
pEC-XC99E	cat_I, $lacI^q$, P_{trc}, *rrnB* (T1 and T2), *ori*$V_{E.c}$, *per* and *repA* (pGA1) $_{c.g.}$. *E. coli – C. glutamicum* shuttle and expression vector conferring chloramphenicol resistance.	(9)
pEC-XC99E-*cgpS-mcherry*	Derivative of pEC-XC99E containing the *cgpS* gene cloned upstream of the *mcherry* gene under control of the *tac* promoter.	This study

Table S2. Oligonucleotides used in this study for cloning, qPCR and affinity chromatography. Bold sequences represent the overlapping sequences needed for Gibson assembly (10). Restriction sites are underlined.

Application	Oligo-nucleotide	Sequence (5´→ 3´) and properties	Comment
pK19*mobsacB-cgpS-strep*	LF_cgpS_pK19_fw	**CCTGCAGGTCGACTCTAGAG** CTGGTCGTCTGTGTAGCTAC	PCR product contains an overlapping sequence to *Bam*HI-digested pK19*mobsacB* plasmid
	LF_cgpS_rv	GTCCATAGTCCTAACCAATCATGTAA	
	cgpS_strep_fw	**GATTGGTTAGGACTATGGAC** ATGGCCATTATTCAGTCGGTC	PCR product contains an overlapping sequence to the left flank of *cgpS* (PCR product above)
	cgpS_strep_rv	TTACTTCTCGAACTGTGGGTG	
	RF_cgpS_fw	**CACCCACAGTTCGAGAAGTAA** GAGCCCTGTGGAGAATTGTTG	PCR product contains overlapping sequences to *cgpS-strep* and to an *Eco*RI-digested pK19*mobsacB* plasmid
	RF_cgpS_pK19_rv	**AAAACGACGGCCAGTGAATT** ACGCGGCGACCTCATC	
	Cgps_indel-fw	GGACATTATCACCCAACCACAC	Oligonucleotides to verify the correct integration of *cgpS-strep*
	CgpS_indel_rv	CAAGGAATCGTTTACCTATATCGAG	
			Restriction enzyme
pAN6 with the coding regions for the N-terminal parts of the CgpS/Lsr2 homologs	C.a.fw	GCGC CATATG ATGGCACGCCGCGAACTAAT	*Nde*I
	C.a.fw	CGCG CCCGGG ATGGCACGCCGCGAACTAAT	*Sma*I
	C.a.N.rv	GCGC GCTAGC CTATACAACCGTGCTGTGATCAATAG	*Nhe*I
	C.a.rv	GCGC GGATCC CTAGTTAGCGCTCTCGTACTTTTC	*Bam*HI
	C.d.fw	CGCG CATATG ATGGCACGTCGTGAAATC	*Nde*I
	C.d.fw	CGCG CCCGGG ATGGCACGTCGTGAAATC	*Sma*I
	C.d.N.rv	GCGC GCTAGC CTAGTGCGCTTTTTCTATGAAGGG	*Nhe*I
	C.d.rv	GCGC GGATCC TTAGCGCTTGGTGGACTTAAG	*Bam*HI
	M.t.fw	GCGC CATATG ATGGCGAAGAAAGTAACCGTC	*Nde*I
	M.t.fw	CGCG CCCGGG ATGGCGAAGAAAGTAACCGTC	*Sma*I
	M.t.N.rv	GCGC GCTAGC CTAGACGCGACGGCCCG	*Nhe*I
	M.t.rv	GCGC TCTAGA TCAGGTCGCCGCGTG	*Xba*I

pAN6 *cgpS*/ *cgpS-strep*/ *cgpS*-N/ *cgpS*-N-*strep*	cgps_fw	CGCGC CATATG ATGGCCATTATTCAGTCGGTCG	*Nde*I
	cgps_strep_rv	CGCGC GCTAGC TTCGAAAGGAATGCCTTCTTTTTC	*Nhe*I
	cgps_rv	CGCGC GAATTC TTA TTCGAAAGGAATGCCTTC	*Eco*RI
	cgpS_n_rv	CGCGC GCTAGC TTA CTGGCGTGCAGATTCCTC	*Nhe*I
	cgpS_n_strep _rv	CGCGC GCTAGC CTGGCGTGCAGATTCCTC	*Nhe*I
pAN6-*alpA-eyfp*	alpA_OL_pAN 6_fw	**TGCAGAAGGAGATATACATA** ATGGCTCAAAAACAGGACACGAC	PCR product contains overlapping sequences to *Nde*I and *Eco*RI-digested pAN6 plasmid
	eYFP-OL_pAN6_rv	**AAAACGACGGCCAGTGAATT** TTATCTAGACTTGTACAGCTCGTCC	
pEC-XC99E-*cgpS-mcherry*	PcgpS-pEC-fw	**GCGGTATTTCACACCGCATATG** CTGGTCGTCTGTGTAGCTAC	PCR product contains overlapping sequences to *Nde*I -digested pEC-XC99E plasmid and to *mcherry*
	cgpS-rv-OL-mcherry	**CTCGCCCTTGCTCACCAT** TTCGAAAGGAATGCCTTCTTTTTCG	
	mcherry_fw	ATGGTGAGCAAGGGCGAG	PCR product contains an overlapping sequence to *PstI* -digested pEC-XC99E plasmid
	mCherry_rv_OL	**AACAGCCAAGCTTGCATGCC** TTACTTGTACAGCTCGTCCATGC	
Application	**Oligo-nucleotide**	**Sequence (5´→ 3´)**	**Comments**
qPCR (circular phage DNA and reference gene)	Phage-LC-for	CCCACGTTCACCCCACAAACG	
	Phage-LC-rev	CTAAAATGAAGCCATCGCGACC	
	ddh-LC-for	ACGTGCTGTTCCTGTGCATGG	
	ddh-LC-rev	GCTCGGCTAAGACTGCCGCT	
Affinity chromatography with P_{alpAC}	PalpAC-Biotin-Tag fw	*GAGGAGTCGTCGATGTGGAGACC* TCGCACTCAATAATGCGGTGG	Asterisks highlight the biotin labelled sequences
	Biotin-oligo	*GAGGAGTCGTCGATGTGGAGACC*	
	PalpAC rv	GCGCATACGCACATTACGC	

Table S3. Oligonucleotides used for the generation of DNA fragments for EMSA experiments.

Oligonucleotide	Sequence (5´→ 3´) and properties	Product length (bp)	GC content of product (%)
gntK-Prom-fw	ATGGTGGCGTCATGCTCGGCCG	560	49.3
gntK-Prom-rv	GGATTTGCCGCAGCCAGAAACGC		
cg0150fw	GGGGTAATAAGACAAAACAGTGGG	500	39.6
cg0150rv	TAGAAATCAGCGACAACCATGCTTC		
cg0421fw	GGATACTTTCTGTTTTGGTTGGTC	500	41.5
cg0421rv	GAAATTACCAAGATGCACCACCTC		
cg0432fw	CCTTTTCTAGACAAGACCTGATC	500	42.0
cg0432rv	ACCAACGACGTCGGATTAGG		
cg0718fw	ATAAGTCATGGTTCAACCTCGG	500	44.0
cg0718rv	CCTAAAACGACACCATCTCAAAAG		
cg0726fw	TACCACTTGCCTTTGTAGCGTTC	500	46.0
cg0726rv	ACTTGGAAACCGGCAGCAAG		
cg1028fw	TGGTCAGCGCAGCGAC	500	50.3
cg1028rv	AAGTTGAGTCTTGGGCCGG		
cg1517fw	GTATGACCAAATGGGACGAAGG	500	42.0
cg1517rv	GATAAGCCACTCAACCACCAAAC		
cg2782fw	GACGCTGAGAAGGACTACG	500	49.5
cg2782rv	TTGAAGGTATCTCCGACAGCAAC		
cg2805fw	AAGAAGGCTGAGTTTAGTGGGG	500	44.8
cg2805rv	AGAAGACGTCCAAAATCCCGTC		
cg3060fw	CAAAATCAATGCGAGAGCGAAG	500	44.0
cg3060rv	CTGCAGAGCTGAAATTATCGAC		
cg3304fw	GGATAACTTCCCCACAATTGAC	500	47.7
cg3304rv	AAGCGTGCCATTGTTCTCCC		
cg1951fw	CTCTATTGGCTCTTAATGGTCAATTAC	500	33.4
cg1951rv	GCCTCTTAAAGCACAGTTATTGCG		
cg1966fw	GCTCAGTATCAATGTCGTCACC	500	36.3
cg1966rv	GTCGAAGTGGTGTCGTTATTTAGG		
cg2023fw	GCACCACCAACAAGTGCC	500	40.7
cg2023rv	TGGGAGCATTTCACTGCACG		
cg1977fw	GTTCTAAACATAAGGAACGCGC	500	39.1
cg1977rv	CGATGGTGCAGTGACCATG		
cg1936fw	CATCGCTCATTGTTACTTAATTACCC	500	36.0
cg1936rv	CCTGAAGAATTTGCTCAGCCG		
cg1940fw	CCATAGTCAAGATTCCCAATCAAC	500	39.5
cg1940rv	GATTCAGGTGATGTAGCGCTG		
cg1917fw	CCTGTAGCCTGCGACGTTAA	500	42.2
cg1917rv	GTGCACCGGTAGCCATAATAG		
cg1895fw	TCACGGGTGGAATCGGAG	500	38.3
cg1895rv	GCTTGGATCATCTGAACAGAGTG		
cg2014fw	AGCGTCAATCGGAATCTGCG	500	40.7
cg2014rv	CAGTTGCGCTAGATAAGCGAG		
cg1890fw	GCGACAAACAAATAGATCAGCTG	500	41.8
cg1890rv	GGGGTTTATTACCTGCCTGC		

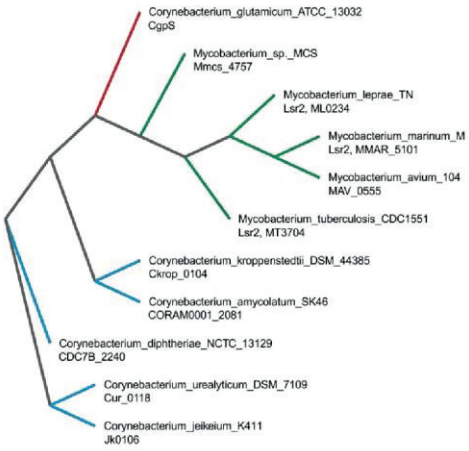
Table S4. Results of the ChAP-Seq experiment. The 90 identified regions are evaluated regarding their peak width, peak maxima and area. Furthermore, the regions are classified into three categories as described in Figure S3. Genes within the CGP3 region are highlighted in green.

Table S5. Impact of CgpS countersilencing on the *C. glutamicum* transcriptome. CGP3 prophage genes are highlighted in green. ORFs exhibiting are more than two-fold altered mRNA ratio (of >2 or < 0.5, *p*-value <0.05) are shown.

Table S6. PSI-BLAST results of CgpS. e-value was set <= 0.005 across several orders of the phylum Actinobacteria and phages as annotated in the NCBI database (http://www.ncbi.nlm.nih.gov/).

Supplementary Figures

A



B

	Corynebacterium glutamicum ATCC 13032	100	27.52	25.66	25.96	27.27	25.45	21.55	23.42	24.76	24.55	21.1
1	*Mycobacterium spec.* MCS	27.52	100	87.93	90.09	87.83	86.09	28.45	27.35	30.77	36.61	34.82
2	*Mycobacterium marinum* M	25.66	87.93	100	93.69	94.87	94.02	26.89	29.91	28.57	40.18	35.71
3	*Mycobacterium avium* 104	25.96	90.09	93.69	100	91.82	91.82	27.27	27.93	30.3	36.45	35.19
4	*Mycobacterium tuberculosis* CDC1551	27.27	87.83	94.87	91.82	100	92.31	28.45	31.03	28.85	40.54	34.23
5	*Mycobacterium leprae* TN	25.45	86.09	94.02	91.82	92.31	100	27.59	28.45	28.85	37.84	35.14
6	*Corynebacterium kroppenstedtii* DSM 44385	21.55	28.45	26.89	27.27	28.45	27.59	100	41.32	39.81	40.35	38.6
7	*Corynebacterium amycolatum* SK46	23.42	27.35	29.91	27.93	31.03	28.45	41.32	100	40.74	41.74	43.48
8	*Corynebacterium diphtheria* NCTC 13129	24.76	30.77	28.57	30.3	28.85	28.85	39.81	40.74	100	44.86	49.06
9	*Corynebacterium urealyticum* DSM 7109	24.55	36.61	40.18	36.45	40.54	37.84	40.35	41.74	44.86	100	53.51
10	*Corynebacterium jeikeium* K411	21.1	34.82	35.71	35.19	34.23	35.14	38.6	43.48	49.06	53.51	100

Figure S1: CgpS orthologs. A. Phylogenetic tree based on the multiple sequence alignments of CgpS/Lsr2 homologs of selected *Corynebacteria* (*C. kroppenstedtii*, *C. amycolatum*, *C. diphteriae*, *C. urealyticum*, *C. jeikeium*), and *Mycobacteria* (*M. tuberculosis*, *M. spec.*, *M. leprae*, *M. marinum*, *M. avium*). Alignments were performed using Clustal Omega (11) with standard configurations. Data for phylogenetic tree were derived from alignments and visualized using tree vector (12). Analysis indicates that CgpS displays a higher sequence identity to mycobacterial Lsr2 proteins than to the corynebacterial orthologs.

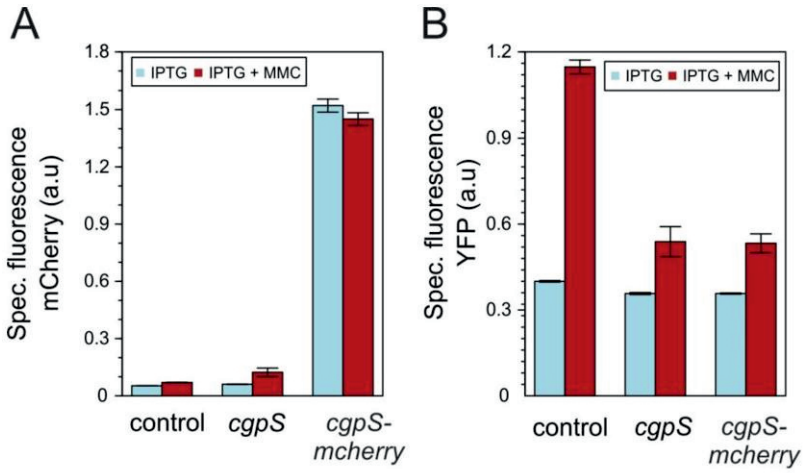


Figure S2: Silencing of CGP3 prophage induction. A and B. Phage reporter cells (WT::P_{lys}-*eyfp*) were transformed with pAN6, pAN6-*cgpS* and pAN6-*cgpS-mcherry* and were cultivated in CGXII with 50 µM IPTG and in the presence or absence of 0.6 µM MMC. The mCherry (A) and eYFP (B) fluorescence as well as backscattered light were measured in the BioLector® microcultivation system and were used to calculate the specific fluorescence. The specific fluorescence after 20 h of cultivation is shown. The data represent average values from three biological replicates including the standard deviation.

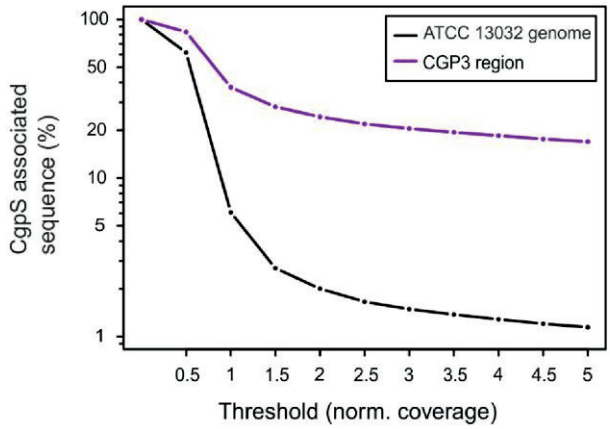


Figure S3: Threshold variation of the CgpS ChAP-Seq data. Based on mean normalized coverage values which were obtained by ChAP-sequencing experiments, thresholds were varied to validate its impact on the estimated binding of CgpS to the CGP3 region and to the entire genome of ATCC 13032. Based on this analysis, bound regions showing a threshold T >3 were considered as CgpS targets in this study (20.46% of CGP3 and 1.49% of the genome).

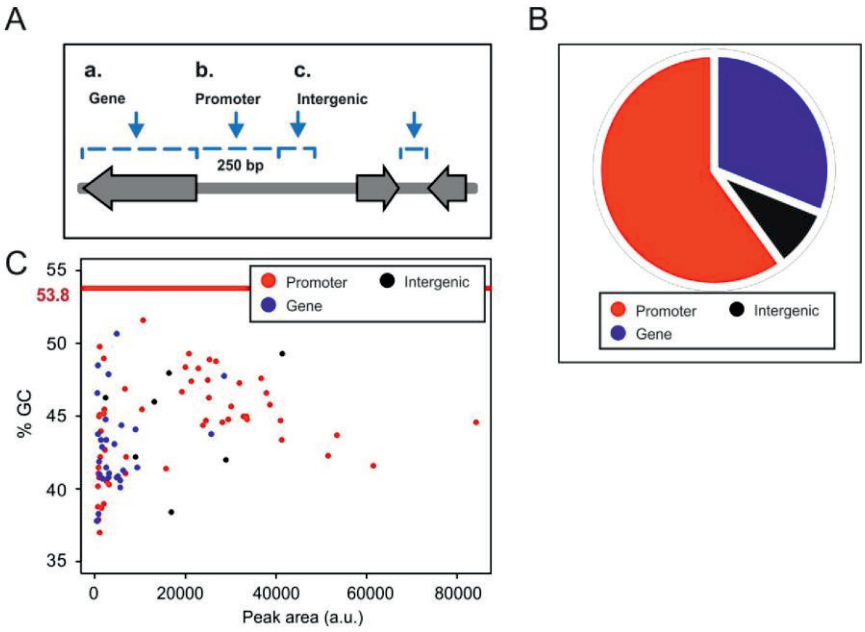


Figure S4: **Genomic distribution of CgpS binding sites within genes, promoters or intergenic regions. A.** The 90 regions bound to CgpS were classified into three categories: i. Binding sites within open reading frames (genes), ii. 250 bp upstream of translational start or according to published transcription start sites (promoter regions), and, iii. intergenic regions. **B**. Distribution of the 90 CgpS-bound genomic regions. Overall, 60% of the peaks are located in promoter regions and 31% within genes. Only 9% are assigned to intergenic regions. **C**. The %GC content of the regions were plotted against peak areas. Red line illustrates average GC content of *C. glutamicum* ATCC 13032, which is about 53.8% (13). Interestingly, a trend to higher peak areas was observed for promoter regions in comparison to intergenic regions or ORFs.

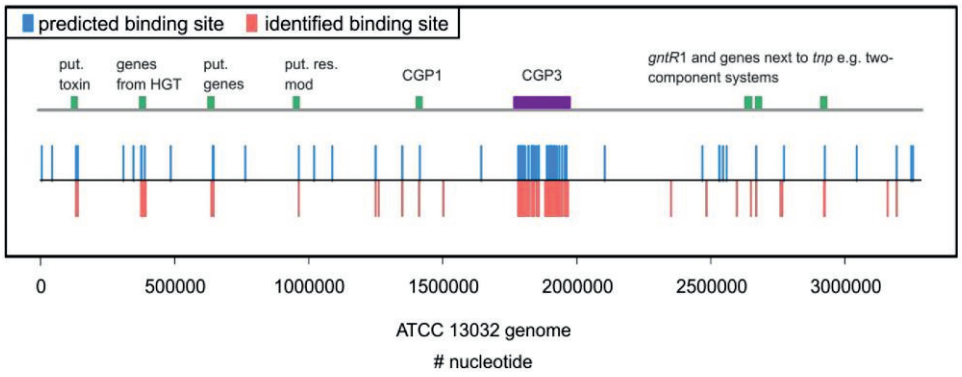
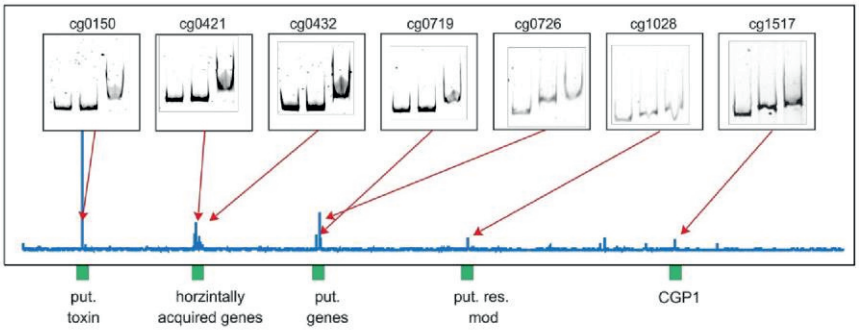
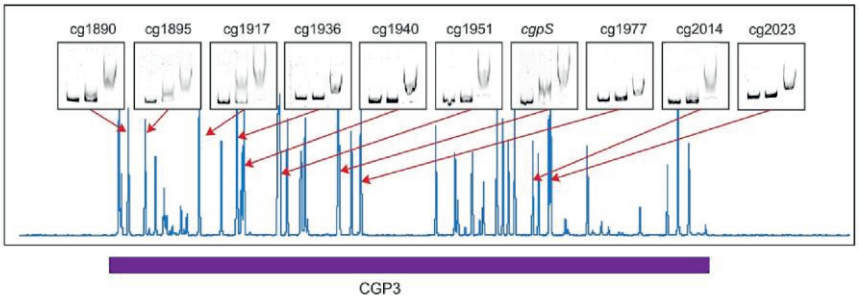


Figure S5: Comparison of predicted and experimentally identified CgpS binding sites. The DNA binding motif derived from ChAP-Seq results (Fig. 3C) was checked for further hits in the genome of ATCC 13032 using FIMO (14). Here, 90 positions exhibiting highest probability (p-Values: $2.7 \cdot 10^{-10}$ – $2.3 \cdot 10^{-6}$) (in blue) were compared with the 90 experimentally identified binding sites acquired by ChAP-Seq binding studies (in red). Potential CgpS site within the CGP3 region (purple boxes) and outside (green boxes) are highlighted. Correlation between experimentally identified and predicted CgpS binding sites ~75 %.

A Upstream of CGP3

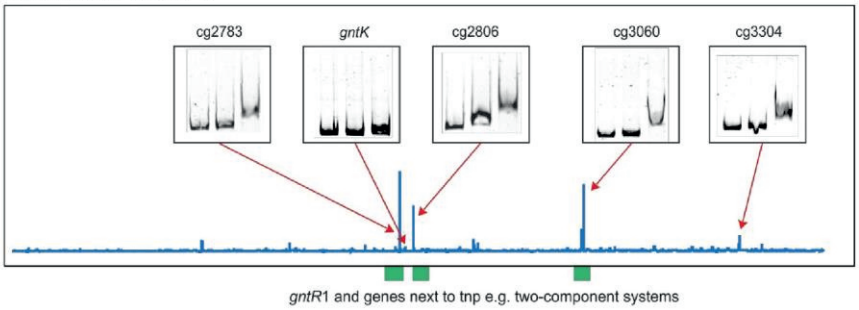


B



Downstream of CGP3

C



D

cg0150	putative transcriptional regulator, Fic/Doc- family	**cg1936**	putative secreted protein, CGP3 region
cg0421	*wzx*, putative translocase involved in export of a cell surface polysaccaride	**cg1940**	putative secreted protein, CGP3 region
cg0432	putative lipopolysaccharide modification acyltransferase	**cg1951**	*tnp*14a, transposase fragment, CGP3 region
cg0719	*crtYe*, C50 carotenoid cyclase, terpenoid synthesis	**cg1966**	*cgpS*, Lsr2-like protein, CGP3 region
cg0726	putative secreted lipoprotein	**cg1977**	putative secreted protein, CGP3 region
cg1028	putative restriction modification system: methylase (EC:2.1.1.72)	**cg2014**	hypothetical protein, CGP3 region
cg1517	putative secreted protein, CGP1 region	**cg2023**	putative membrane protein, CGP3 region
cg1890	*alpC*, actin like protein, CGP3 region	**cg2783**	*gntR1*, gluconate-responsive repressor
cg1895	putative secreted protein, CGP3 region	**cg2806**	putative membrane protein
cg1917	hypothetical protein, CGP3 region	**cg3060**	*cgtS6*, two-component sensor kinase
		cg3304	*dnaB*, replicative DNA helicase

Figure S6: *In vitro* binding studies of CgpS to its putative target sites. Electrophoretic mobility shift assays (EMSAs) were performed with purified CgpS-Strep protein and 21 putative target DNA regions derived from ChAP-Seq data (Fig. 3). Green boxes indicate regions outside of CGP3 and the purple box sites within the CGP3 region. All tested DNA fragments had a size of about 500 bp and were chosen 250 bp up and downstream of the peak maxima, which were detected by the ChAP-Seq analysis. Overall, eleven candidate regions were chosen outside of CGP3 ((**A**) seven upstream and (**C**) four downstream of CGP3) and ten sites within the CGP3 region (**B**). In all lanes 90 ng DNA (12-14 pM) were incubated without (lane 1) or with increasing amounts of CgpS protein (lane 2: 1 µM and lane 3: 2 µM). The promoter region of *gntK* (560 bp) was used as a negative control. Annotations und potential functions of the bound regions are listed in **D**.

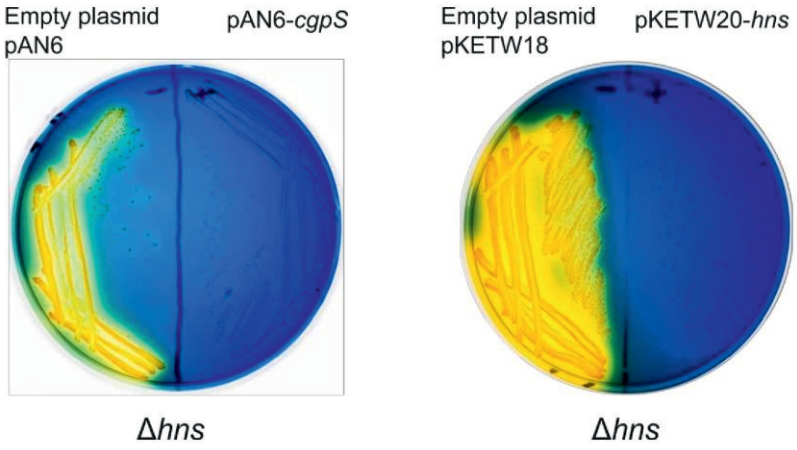


Figure S7: Complementation studies of a *E. coli* K-12 Δ*hns* strain with *cgpS* cloned into the overexpression plasmid pAN6. Cells were grown on bromothymol blue salicin indicator plates as described in Dole et al., 2002 (15). *E. coli* cells lacking *hns* were transformed with the empty plasmid pAN6, pAN6-cgpS or with the empty plasmid pKETW18 or pKETW20 carrying *hns*. Plates were incubated at 37°C overnight. Complementation is based on the utilization of salicin. Salicin can be used as carbon source if the *bgl* operon is expressed. This operon is repressed by H-NS in the wild type situation. Thus, in the absence of H-NS, salicin is metabolized leading to a decrease of the pH resulting in a colour shift from blue to yellow. Complementation of the Δ*hns* phenotype was achieved by expressing either *hns* or *cgpS* suggesting a similar function of both proteins.

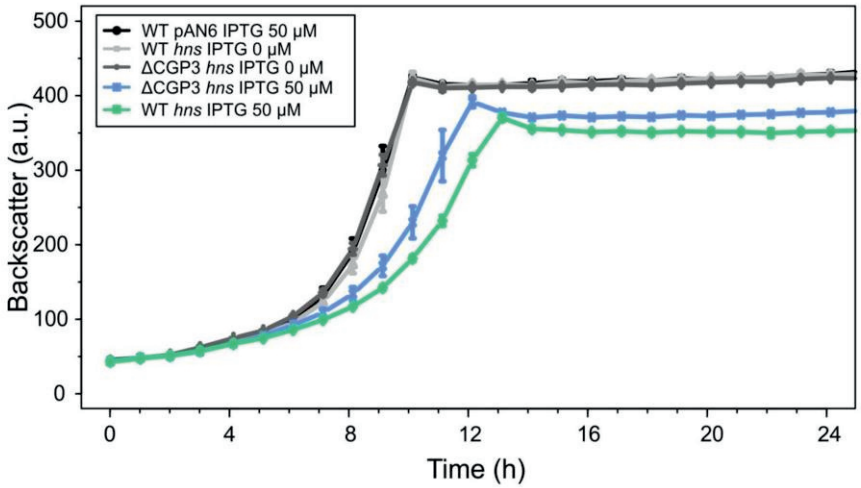


Figure S8: Overexpression of *hns* in *C. glutamicum* strains. H-NS encoding gene located on the overexpression plasmid pAN6 was overexpressed in the prophage reporter strain WT::P_{lys}-*eyfp* and in the ΔCGP3 strain. Cells were cultivated in CGXII minimal medium and *hns* expression was induced with 50 µM IPTG. The data represent average values of three biological replicates including the standard deviation.

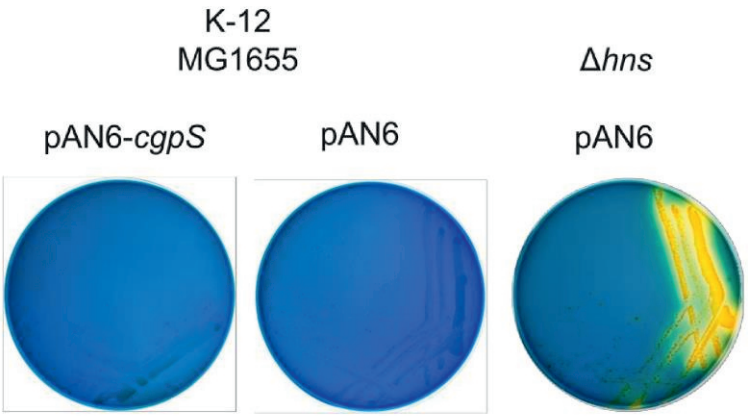


Figure S9: *cgpS* overexpression in *E. coli* wild type cells. To verify whether CgpS is interfering with the function of H-NS in its native host, *E. coli* K-12 MG1655 wild type cells were transformed with the pAN6-*cgpS* plasmid. Cells were streaked on bromothymol blue salicin indicator plates (15) supplemented with 100 µM IPTG. As control, the wild type strain and a Δ*hns* mutant were transformed with the empty plasmid pAN6. The obtained results suggest that heterologous *cgpS* expression is not able to counteract H-NS silencing at the *bgl* promoter when compared to a mutant lacking the *hns* gene. However, it needs to be highlighted that the resulting *E. coli* strain expressing the *cgpS* gene (left plate) showed a significant growth defect in comparison to the empty vector controls (middle and right).

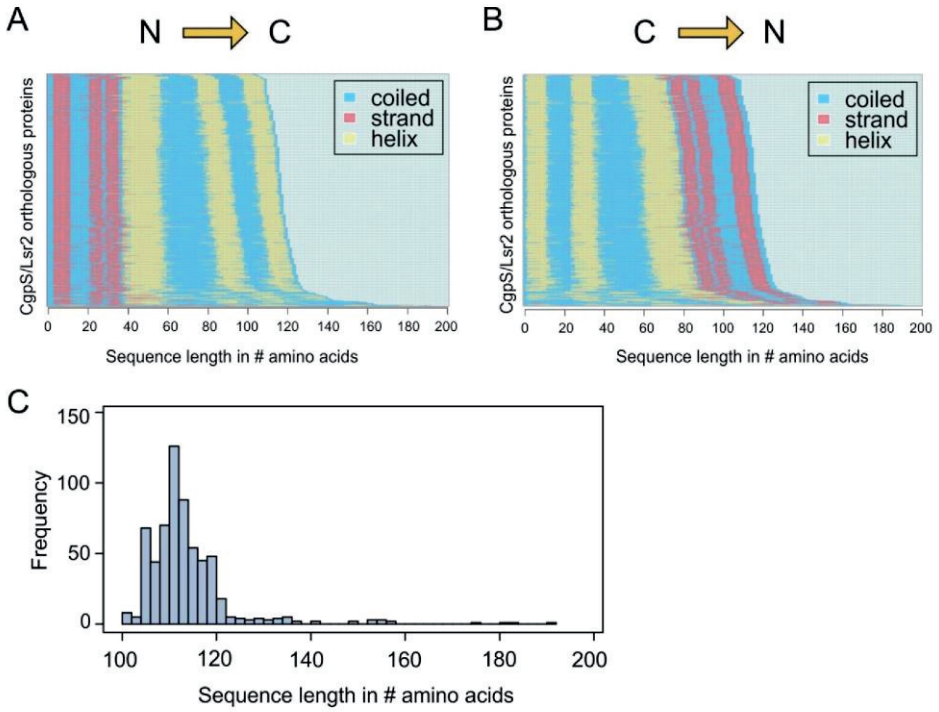


Figure S10. Bioinformatic analysis of CgpS related proteins. A PSI-BLAST search on CgpS homologs with an *e*-value of 0.005 was conducted and achieved 5230 hits (Table S6). 1920 sequence are individual and can be assigned to 863 taxonomical units; 618 of these can be allocated to bacteria or phages. Secondary structure predictions of the 618 sequences are shown in direct comparison in N->C (**A**) and C->N (**B**) orientation. The increasing length of the amino acid sequences entails distortet matches in secondary structure prediction and hence for a better overview the two possibilites are shown. **C.** Histogramm of the 618 sequences ordered according to their amino acid sequence length. The maximum of this distribution is located around 110 amino acids.

Supplementary Videos

Video S1: Time lapse video of a *C. glutamicum* microcolony under standard conditions (without IPTG, control). Cells of the prophage reporter strain ATCC 13032::P_{lys}-*eyfp* carrying the countersilencing plasmid pAN6-N-*cgpS* were cultivated in microfluidic chambers (16) in standard minimal medium (CGXII with 2% (w/v) glucose, 25 µg·ml kanamycin for 20 h without IPTG). The video shows the first 12 h of the cultivation.

Video S2: Time lapse video of the effect of CgpS countersilencing (150 µM IPTG) on prophage activation. The same reporter strain (Video S1) was grown in the presence of 150 µM IPTG inducing the expression of the truncated CgpS protein (aa 1-65) covering its oligomerization domain. The video shows the first 16.5 h of the experiment.

References

1. Studier, F.W. and Moffatt, B.A. (1986) Use of Bacteriophage-T7 Rna-Polymerase to Direct Selective High-Level Expression of Cloned Genes. *J Mol Biol*, **189**, 113-130.
2. Venkatesh, G.R., Koungni, F.C.K., Paukner, A., Stratmann, T., Blissenbach, B. and Schnetz, K. (2010) BglJ-RcsB Heterodimers Relieve Repression of the *Escherichia coli bgl* Operon by H-NS. *J Bacteriol*, **192**, 6456-6464.
3. Stratmann, T., Pul, U., Wurm, R., Wagner, R. and Schnetz, K. (2012) RcsB-BglJ activates the *Escherichia coli* leuO gene, encoding an H-NS antagonist and pleiotropic regulator of virulence determinants. *Mol Microbiol*, **83**, 1109-1123.
4. Kinoshita, S., Udaka, S. and Shimono, M. (1957) Studies on the amino acid fermentation - Part I. Production of L-glutamic acid by various microorganisms. *J Gen Appl Microbiol*, **50**, 331-343.
5. Baumgart, M., Unthan, S., Rückert, C., Sivalingam, J., Grünberger, A., Kalinowski, J., Bott, M., Noack, S. and Frunzke, J. (2013) Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Appl Environ Microbiol*, **79**, 6006-6015.
6. Helfrich, S., Pfeifer, E., Krämer, C., Sachs, C.C., Wiechert, W., Kohlheyer, D., Nöh, K. and Frunzke, J. (2015) Live cell imaging of SOS and prophage dynamics in isogenic bacterial populations. *Mol Microbiol*, **98**, 636-650.
7. Frunzke, J., Engels, V., Hasenbein, S., Gätgens, C. and Bott, M. (2008) Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2. *Mol Microbiol*, **67**, 305-322.
8. Schäfer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G. and Pühler, A. (1994) Small Mobilizable Multipurpose Cloning Vectors Derived from the *Escherichia-Coli* Plasmids Pk18 and Pk19 - Selection of Defined Deletions in the Chromosome of *Corynebacterium-Glutamicum*. *Gene*, **145**, 69-73.
9. Kirchner, O. and Tauch, A. (2003) Tools for genetic engineering in the amino acid-producing bacterium Corynebacterium glutamicum. *J Biotechnol*, **104**, 287-299.
10. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A. and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, **6**, 343-U341.
11. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*, **7**, 539.
12. Pethica, R., Barker, G., Kovacs, T. and Gough, J. (2010) TreeVector: Scalable, Interactive, Phylogenetic Trees for the Web. *Plos One*, **5**.
13. Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B.J., Gaigalat, L. *et al.* (2003) The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J Biotechnol*, **104**, 5-25.
14. Grant, C.E., Bailey, T.L. and Noble, W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017-1018.
15. Dole, S., Kühn, S. and Schnetz, K. (2002) Post-transcriptional enhancement of *Escherichia coli bgl* operon silencing by limitation of BglG-mediated antitermination at low transcription rates. *Mol Microbiol*, **43**, 217-226.
16. Grünberger, A., Probst, C., Helfrich, S., Nanda, A., Stute, B., Wiechert, W., von Lieres, E., Nöh, K., Frunzke, J. and Kohlheyer, D. (2015) Spatiotemporal Microbial Single-Cell Analysis Using a High-Throughput Microfluidics Cultivation Platform. *Cytometry*, **87A**, 1101-1115.

6.5 Supplemental material to: Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium.

Supplementary material to

Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium.

Eugen Pfeifer, Cornelia Gätgens, Tino Polen*, and Julia Frunzke*

Institute of Bio- und Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH,

52425 Jülich, Germany

*Corresponding author:

Tino Polen; Email: t.polen@fz.-juelich.de; Phone +49 2461 61 6205

Julia Frunzke; Email: j.frunzke@fz-juelich.de; Phone: +49 2461 61 5430

Supplementary Tables

Table S1: Bacterial strains used in this study.

Strains	Relevant characteristics	Reference
E. coli		
DH5α	*supE44 ΔlacU169* (φ80*lacZ*DM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1.*	Invitrogen
C. glutamicum		
ATCC 13032	Biotin-auxotrophic wild type.	(1)
MB001	ATCC 13032 with in-frame deletion of prophage CGP1, CGP2 and CGP3	(2)
Wildtype_bc	Wild type derivative; To barcode the strains the sequence 5'GCGCTTCATATGGTCAGTTATGAAG-3' was integrated into the locus between cg1121 and cg1122	This study
MB001_bc	MB001 derivative; the barcode 5'CATTCC AAGCTT GATCCGTACATAG-3' was integrated into the intergenic region between cg1121 and cg1122	This study
ATCC 13032::Ptac-eyfp	Wild type derivative; contains an *eyfp* gene which is under the control of *tac* promoter and which were integrated into the region between cg1121-cg1122	(2)
ATCC 13032::Ptac-crimson	Wild type derivative; contains an *crimson* gene which is under the control of *tac* promoter and which were integrated into the region between cg1121-cg1122	(2)
MB001::Ptac-eyfp	MB001 derivative; contains an *eyfp* gene which is under the control of *tac* promoter and which were integrated into the region between cg1121-cg1122	(2)
MB001::Ptac-crimson	MB001 derivative; contains an *crimson* gene which is under the control of *tac* promoter and which were integrated into the region between cg1121-cg1122	(2)
ATCC 13032 pyk_T12A	Wild type derivative; in the *pyk* gene (cg2291) adenine at the position 34 was changed to guanine	This study
ATCC 13032 pyk_A20V	Wild type derivative; in the *pyk* gene (cg2291) cytosine at the position 34 was changed to adenine	This study
ATCC 13032 pyk_A271T	Wild type derivative; in the *pyk* gene (cg2291) guanine at the position 811 was changed to adenine	This study
ATCC 13032 fruK_T6I	Wild type derivative; in the *fruK* gene (cg2119) cytosine at the position 17 was changed to adenine	This study
ATCC 13032 fruK_R71L	Wild type derivative; in the *fruK* gene (cg2119) guanine at the position 212 was changed to adenine	This study
ATCC 13032 corA_del	Wild type derivative; in gene *corA* the 12 bp CGTCGACGATGG position 593 to 604 in cg0080	This study
ATCC 13032 corA _Q307*	Wild type derivative; in gene *corA* cytosine at the position 919 was changed to thymine	This study

ATCC 13032 pyk_A271T_fruK_R71L	Wild type derivative; in *pyk* gene (cg2291) guanine at the position 811 was changed to adenine and in *fruK* guanine at the position 212 was changed to adenine	This study

Table S2: Plasmids used in this study.

Plasmid	Relevant characteristics	Reference
pK19mobsacB	plasmid that is used for allelic exchange in *C. glutamicum*; $oriV_{E.c.}$, *sacB lacZα KanR*, *Eco*RI and *Hin*dIII restrictions sites were for Gibson assembly	(3)
pK18mobsacB-int1	KanR; plasmid is used for integration of DNA into the intergenic region between cg1121 and cg1122 ($oriV_{E.c}$ *sacB lacZ*)	(2)
pK18_BC_WT	derivative of pK18 that was used to barcode (5'GCGCTTCATATGGTCAGTTATGAAG-3') the wild type strain	This study
pK18_BC_MB	derivative of pK18 that was used to barcode (5'CATTCCAAGCTTGATCCGTACATAG-3') the MB001 strain	This study
pK19_pyk_T12A	derivative of pk19 to replace native *pyk* gene with *pyk*_T12A	This study
pK19_pyk_A20V	derivative of pk19 to replace native *pyk* gene with *pyk*_A20V	This study
pK19_pyk_A271T	derivative of pk19 to replace native *pyk* gene with *pyk*_A271T	This study
pK19_fruK_T6I	derivative of pk19 that was used to insert the mutation T6I into the *fruK* gene	This study
pK19_fruK_R71L	derivative of pk19 that was used to insert the mutation R271L into the *fruK* gene	This study
pK19_ corA _del	derivative of pk19 that was used to delete 18 bp of the *corA* gene	This study
pK19_ corA _Q307*	derivative of pk19 that was used to insert the mutation Q307* in *corA*	This study
pJC1-	Kan^R, Amp^R, *C. glutamicum* and *E. coli* shuttle vector	(4)
pJC1-*ptsG-eyfp*	pJC1 derivative containing promoter of *ptsG* fused to *eyfp*. 657 bp upstream of cg1537 (*ptsG*) were considered as promoter region.	This study

Table S3. Oligonucleotides used in this study. Bold sequences point out overlapping parts for Gibson Assembly.

Oligonucleotide	Sequence (5' -> 3')	Comment
WT_OL_fw	**TCCTCTTGCTCGTGTCAATT** GCGCTT CATATG GTCAGTTATGAAG	Label of wild type strain
WT_OL_rv	**GTGTCCATGAGTTCGCTCGA** CTTCATAACTGAC CATATG AAGCGC	
WT_BC_fw	GCGCTT CATATG GTCAGTTATGAAG	
WT_BC_rv	CTTCATAACTGAC CATATG AAGCGC	
MB001_OL_fw	**TCCTCTTGCTCGTGTCAATT** CATTCC AAGCTT GATCCGTACATAG	Label of MB001 strain
MB001_OL_rv	**GTGTCCATGAGTTCGCTCGA** CTATGTACGGATC AAGCTT GGAATG	
MB001_BC_fw	CATTCC AAGCTT GATCCGTACATAG	
MB001_BC_rv	CTATGTACGGATC AAGCTT GGAATG	
		Construction of
pyk_T12A-P1	**GACCATGATTACGCCAAGCT** ACCTTGGGCCGTTTCTGG	pK19_pyk_T12A
pyk_T12A-P2	**GCCACCGCTGGGCCTAGGGC** ACATACAATCTTAGTTCGTCTATCCAC	
pyk_T12A-P3	**GAACTAAGATTGTATGTGCC** CTAGGCCCAGCGGTG	
pyk_T12A-P4	**AAAACGACGGCCAGTGAATT** AGGTCACGGATATCCTTTTCGG	
pyk_T12A-seq_fw	ACCTTGGGCCGTTTCTGG	
pyk_T12A-seq_rv	AGGTCACGGATATCCTTTTCGG	
pyk_A20V-P2	**AAACGCAGAATTCCATC TAC** ACTAGCCACCGCTGGG	pK19_pyk_A20V
pyk_A20V-P3	**GCCCAGCGGTGGCTAGTGTA** GATGGAATTCTGCGTTTGGTAGAAG	
pyk_A271T-P1	**GACCATGATTACGCCAAGCT** CACGATCGTGTGTCCACC	pK19_pyk_A271T
pyk_A271T-P2	**GCCACGATAACTGGCTT TGT** GTTCTCACGGGCAATCTGG	
pyk_A271T-P3	**AGATTGCCCGTGAGAACACA** AAGCCAGTTATCGTGGCAACC	
pyk_A271T-P4	**AAAACGACGGCCAGTGAATT** GGCATTGCTAAAAGAGCACGG	
pyk_A271T-seq_fw	ACGATCGTGTGTCCACC	
pyk_A271T-seq_rv	GGCATTGCTAAAAGAGCACGG	
corA_del_P1	**GACCATGATTACGCCAAGCT** CTGTACCTGAAGAGGTTTCTCC	pK19_ corA _del
corA_del_P2	**CTAAGAACGGGGGAGAAGC** CGTCGACAAGCCTATAGGC	
corA_del_P3	**GCCTATAGGCTTGTCGACG** GCTTCTCCCCCGTTCTTAG	
corA_del_P4	**AAAACGACGGCCAGTGAATT** ACTGCGTTTGAAGATCCAGTAC	
corA_del_seq_fw	CGTCGGTACCTTCCTGATTCGGC	
corA_del_seq_rv	CCTGGAATTTCCGAGGCC	
corA_Q307*_P1	**GACCATGATTACGCCAAGCT** CATGAAGCCTCAGGCCATAG	pK19_ corA _Q307*
corA_Q307*_P2	**ATGTCTTCGTTTTGGCGTTA** GGCTACAAGGGTGGCG	
corA_Q307*_P3	**ACGCCACCCTTGTAGCCTAA** CGCCAAAACGAAGACATGAAG	
corA_Q307*_P4	**AAAACGACGGCCAGTGAATT** CAGCTCAAGGAACGTTTCCTC	

fruK_T6I_P1	GACCATGATTACGCCAAGCT CCTGGATTCAGGTTTAGCGTG	pK19_fruK_T6I
fruK_T6I_P2	**TCAATACTCGGGTTTGGGAT** GAATGTGATGATCATGGGGTTACC	
fruK_T6I_P3	**CCATGATCATCACATTCATC** CCAAACCCGAGTATTGATTCC	
fruK_T6I_P4	**AAAACGACGGCCAGTGAATT** CCATCAGTGGCTTGTCTGAG	
fruK_T6I_seq_fw	CTCATCACACCTCTAGCACG	
fruK_T6I_seq_rv	AGCTGGCCCAGTTCCAG	
fruK_R71L_P2	**CGGGCAAGCCGATGTCGAG** GACCAGTGGGACGAAGG	pK19_fruK_R71L
fruK_R71L_P3	**CCTTCGTCCCACTGGTC CTC** GACATCGGCTTGCCCG	

Table S4: Frequencies (in %) of mutations in *fruK*, *pyk*, *corA*, *soxA* and *hmuV*, which were identified in the evolved strains. WT_A and WT_B: wild type clones A and B; MB_A, MB_B: MB001 clones A and B.

Gene	Mutation	Strain	#100	#140	#540	#630	Top1	Top2	Top3
fruK	exchange T6I	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	-	20	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	96	97.73	-	-	-
	exchange R71L	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	-	-	100	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
	Insertion of *tnp13b*	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	73	79	-	-	100
		MB_B	-	-	-	-	-	-	-
pyk	exchange A271T	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	91.84	96.43	100	100	99.53
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
	exchange P257L	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	33.08	37.72	-	-	-
		MB_B	-	-	-	-	-	-	-
	exchange D175G	WT_A	-	14	-	-	-	-	-
		MB_A	-	12.5	-	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	11.05	-	-	-	-	-
	exchange A20V	WT_A	-	15.38	47.69	28.26	100	100	-
		MB_A	10.67	27.03	-	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	12	17.74	-	-	-	-	-
	exchange T12A	WT_A	-	-	53.42	73.81	-	-	-
		MB_A	-	11.43	-	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
corA	Stop W365*	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	96.4	97.8	-	-	-
	Stop Q307*	WT_A	-	-	54.47	25	-	99.49	-
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
	exchange A302P	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	29.13	25.81	-	99.29	-
		WT_B	-	-	-	-	-	-	-

		MB_B	-	-	-	-	-	-	-
	exchange R291H	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	18.63	11.54	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
	Deletion 12 bp, CGTCGACGATGG	WT_A	-	-	-	-	-	-	-
	62905..62916	MB_A	-	-	13.58	30.77	98.55	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
	Insertion of *tnp13b*	WT_A	-	-	36	48	72	-	-
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	76	72	-	-	89
		MB_B	-	-	-	-	-	-	-
	Insertion of *tnp1b*	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	22	-	-	-	73
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
soxA	L26L	WT_A	-	-	23.26	25.51	40.1	-	36.03
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	27.88	24.22	-	-	-
		MB_B	-	-	-	-	-	-	-
	exchange H29N	WT_A	-	-	-	-	-	-	33.33
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	25.95	26.71	-	-	-
		MB_B	-	-	-	-	-	-	-
	Deletion,C	WT_A	-	-	-	-	-	-	33.56
	1673063	MB_A	-	-	-	-	-	-	-
		WT_B	-	-	25.95	27.44	-	-	-
		MB_B	-	-	-	-	-	-	-
	exchange L271F	WT_A	-	-	-	-	-	-	-
		MB_A	-	-	-	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	27.57	21.55	-	-	-

hmuV	Insertion, G	WT_A	-	-	-	-	-	-	-
	414917^414918	MB_A	-	13.64	-	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
	Insertion, A	WT_A	11.54	-	-	-	-	-	-
	414923^414924	MB_A	-	15	12.5	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-
	exchange S73Y	WT_A	16.67	15.15	-	-	10.87	-	-
		MB_A	-	19.23	15.15	11.76	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	11.54	-	-	10.77
	exchange D74E	WT_A	14.29	14.29	-	-	-	-	11.11
		MB_A	-	15.15	10.81	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	11.11	-	-	-
	exchange A70E	WT_A	13.04	-	-	-	-	-	-
		MB_A	-	14.29	11.43	-	-	-	-
		WT_B	-	-	-	-	-	-	-
		MB_B	-	-	-	-	-	-	-

Table S5: Comparison of mutations found in samples of the 2nd experiment taken after eight (#50 generations) and twelve (#80 generations) cultivation cycles.

	WT		MB001	
SNV and Annotation	**#50 Freq. (%)**	**#80 Freq. (%)**	**#50 Freq. (%)**	**#80 Freq. (%)**
exchange A70E in cg0469, *hmuV*, hemin transport system, ATP-binding protein				10.71
Insertion of transposon *tnp13a* in cg2119, *fruK*		43		
exchange D334Y in cg3213, putative secreted protein			10.87	10.87
exchange M92I in cg2458,*pgp2*, predicted phosphatase, HAD-family (EC:3.1.3.18)				12.5
exchange S73Y in cg0469, *hmuV*, hemin transport system, ATP-binding protein				13.64
exchange V123L in cg2069, *psp1*, putative secreted protein CGP3 region	100	100		
414917^414918_Insertion_G, *hmuV*, hemin transport system, ATP-binding protein				11.54
414923^414924_Insertion_A, *hmuV*, hemin transport system, ATP-binding protein				12.5
MNV, 3078739..3078740_TA_AT, in cg3213(+), putative secreted protein			10.64	10.64
IGR between cg1121 (putative permease of the major facilitator superfamily) and cg1122 (putative secreted protein)	47.95	43.29	47.19	49.48
S270S in cg3197(+), *psp5*, putative secreted protein	100	100		
exchange F26L in cg2807, *tnp11a*, transposase, putative pseudogene		19.75		
exchange Q12H in cg3342, putative secreted protein			10.53	10.53

Supplementary figures

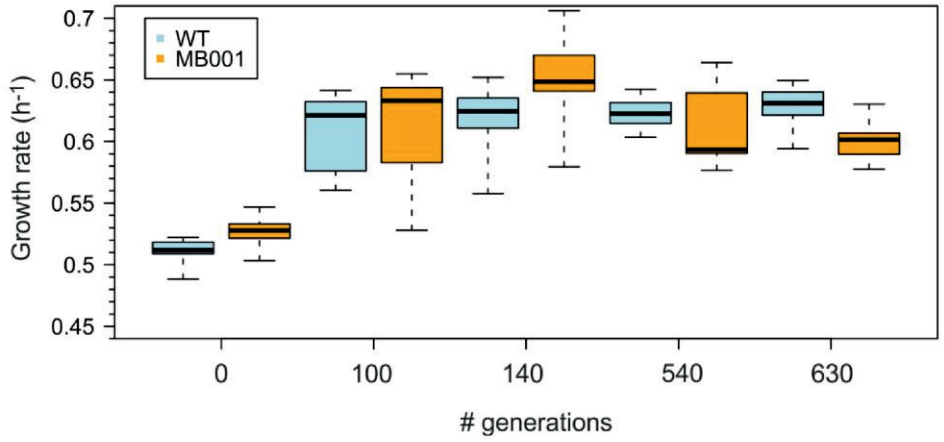


Figure S1: Growth rate distribution of each strain from a second cell line of the first ALE experiment. Growth studies were conducted as described in material and methods part. Single colonies were used for the inoculation of the 1st preculture of the 2nd initial cell line. Growth rates of 24 colonies are compared for each time point and strain. Whiskers represent the maximum and minimum values. Data of the first cell line are given in Figure 3.

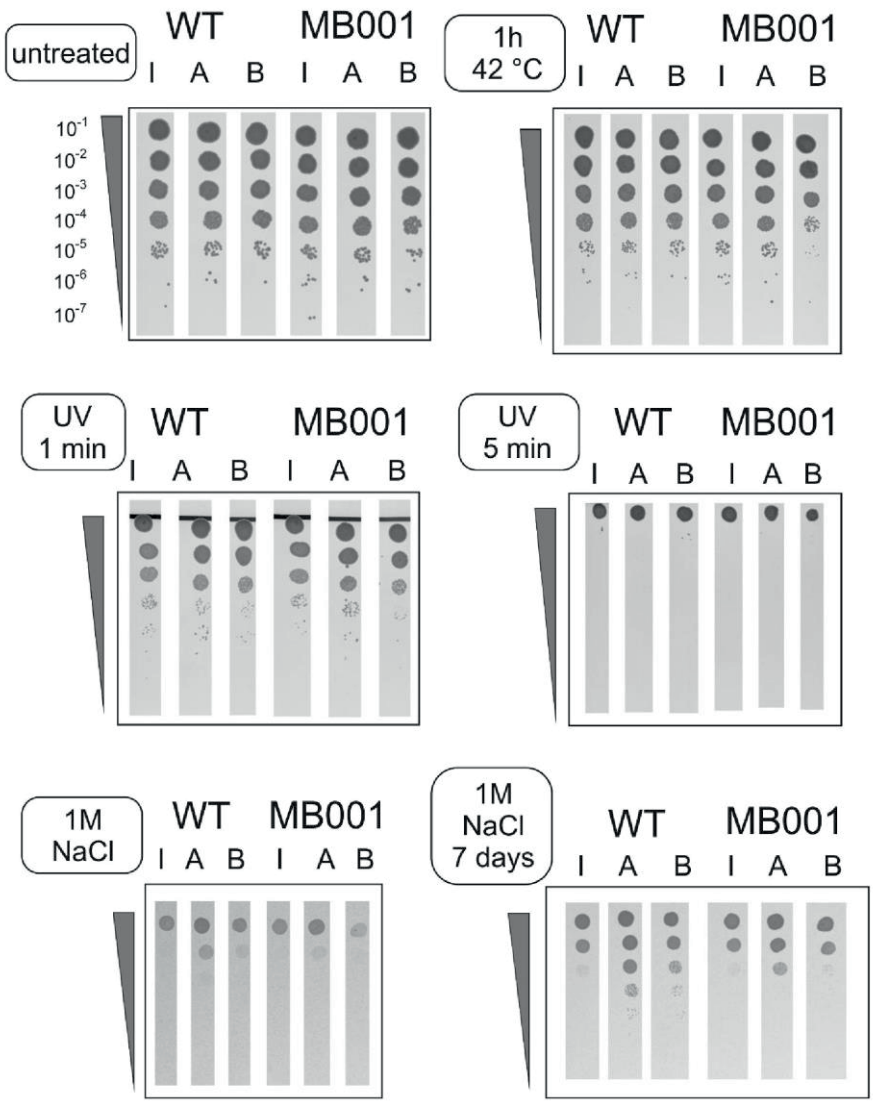


Figure S2: Stress tests of the evolved strains in comparison to the wild type and MB001 strain. Sensitivity towards stress conditions e.g. heat shock, UV stress and osmotic stress of evolved populations were investigated (630 generations, A = mixture of clone A, B = mixture of B) and initial strains (I). The pre-cultivations were conducted according to routine protocols as described in the material and methods part. 3 µl of diluted cell suspensions were spotted on CGXII agar plates with 2% (w/v) glucose. If not stated otherwise plates were incubated for 48 h at 30 °C.

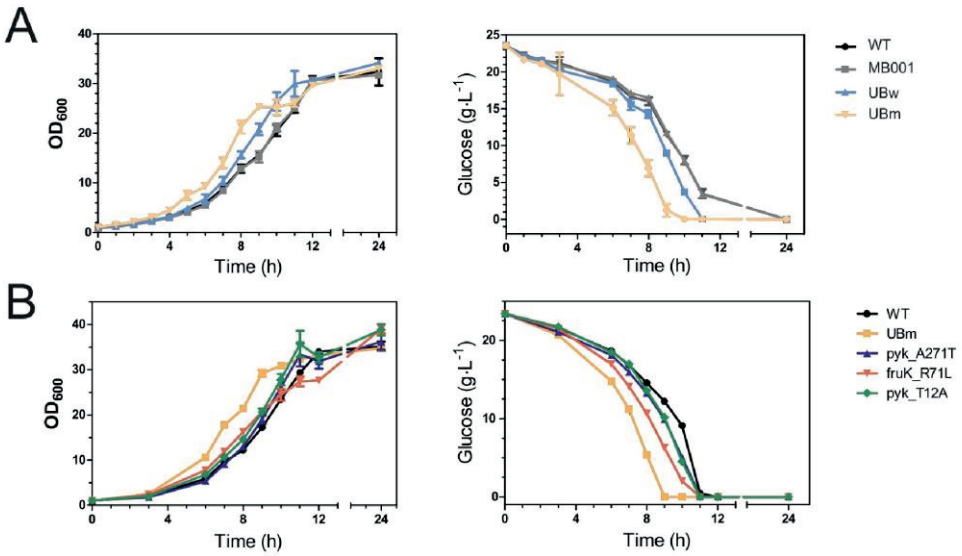


Figure S3: Growth and glucose consumption of initial, evolved and single-mutant strains. **A**. Growth rates of evolved strains UBm and UBw were determined by shaking flask experiments. Glucose concentrations of supernatants were measured as described in the material and methods section and glucose consumption rates are listed in Table 3. Cells were cultivated in CGXII minimal media with 2% (w/v) glucose. Standard deviations are based on three biological replicates. **B** Glucose uptake rates of ATCC 13032 derived strains containing the mutations *pyk* A271T, *pyk* T12A and *fruK* R71L were determined and are also given in Table 3.

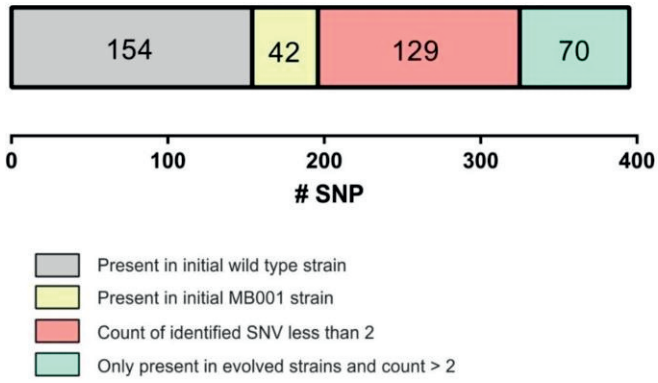


Figure S4: Evaluation of SNPs acquired for the ALE samples.
In sum 395 SNPs were identified by referring all evolved samples to the *C. glutamicum* ATCC 13032 genome sequence published in 2003 (5). 154 of them (grey) were also present within the initial wild type cells and are therefore excluded. Further 42 SNPs (yellow) were also found in the sequence of MB001 (and for this reason they are also excluded). 129 mutations (red) were counted less than two times and, thus are likely not responsible for the observed growth phenotypes. 70 SNPs (green) were only identified in evolved strains and were futher studied in detail.

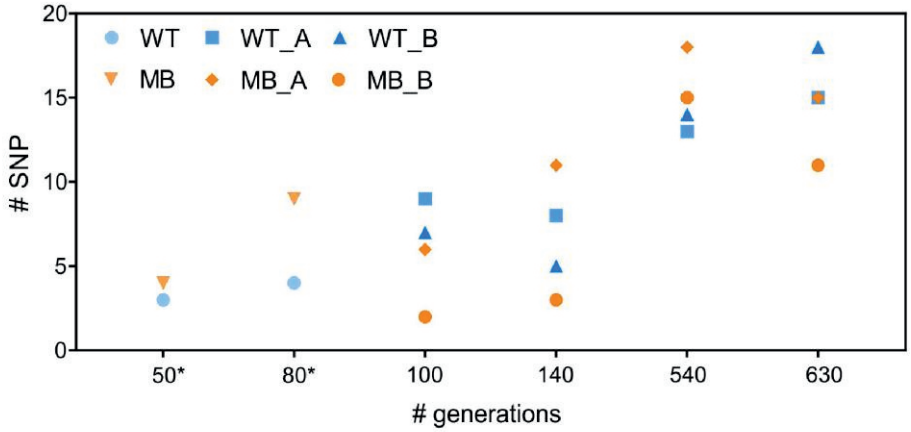


Figure S5: Distribution of the 70 relevant SNPs (presented in green in Figure S4). Number of SNPs identified in ALE samples plotted for each time point. The amount of SNPs found in wild type strains (blue) and prophage-free strains (orange) increased over time. A and B represent two independent cell lines of the main experiment. Samples were taken after 100, 140, 540 and 630 generations. Samples marked with an asterisk represent samples from the second ALE experiment, which was conducted to verify the fitness leap after roughly 100 generations.

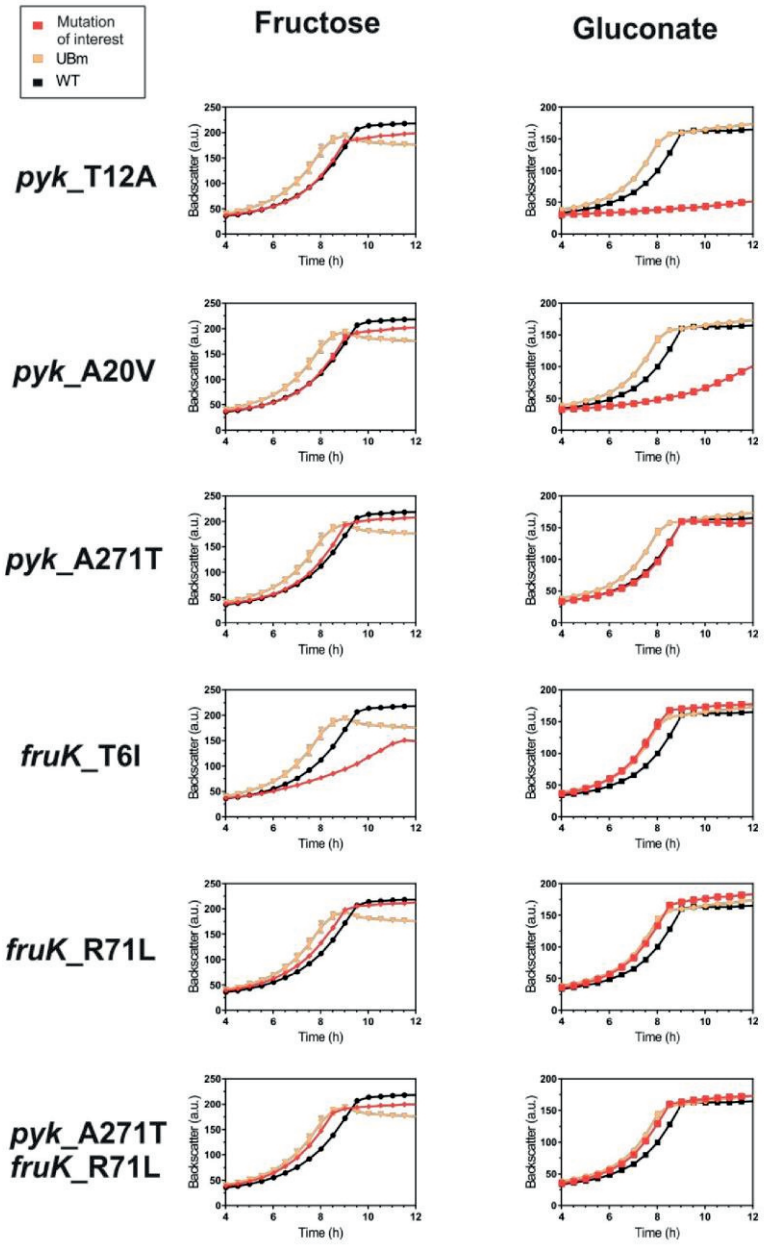


Figure S6: Growth experiments of single-mutant strains on fructose and gluconate. To investigate impacts of mutations on the utilization of fructose and gluconate single-mutant strains were cultivated in CGXII with 2% (w/v) fructose (left) or 2% (w/v) gluconate (right) in a microtiter plates. Wild type and UBm were used as reference strains. The data represent average values of three biological replicates including standard deviation.

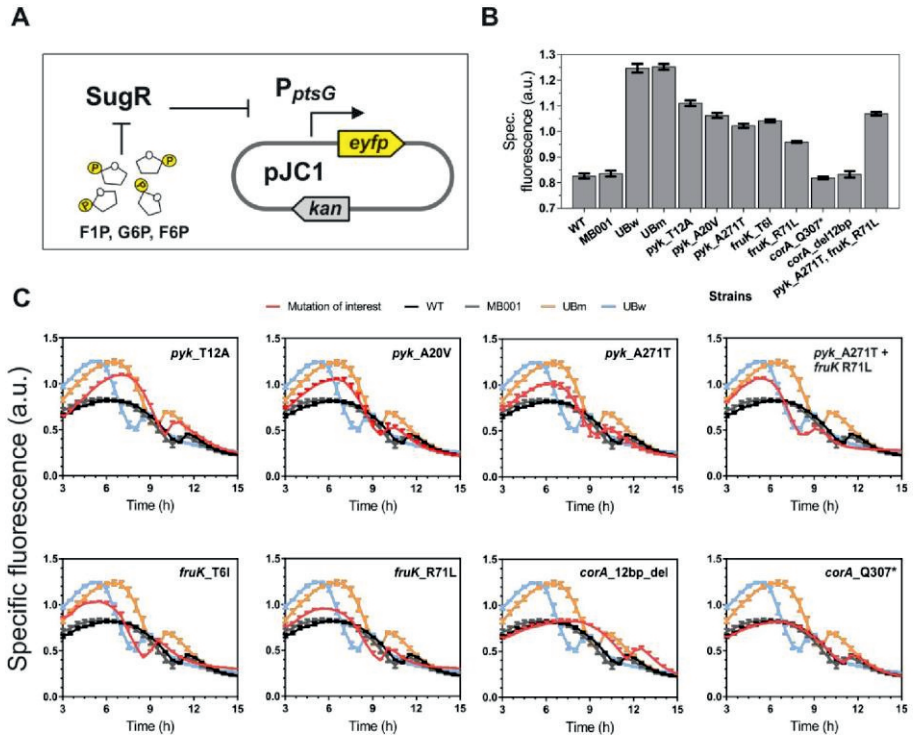


Figure S7: Impact of identified key mutations on *ptsG* expression.

A. A plasmid-based reporter was constructed to visualize expression of *ptsG* encoding enzyme 2 of the PTS system. The promoter region of *ptsG* (500 bp upstream) was fused to *eyfp* encoding the yellow fluorescent protein. The DeoR-type regulator SugR represses expression of *ptsG*, which can be relieved by hexose phosphates. Here, different effector molecules were reported in recent studies (6, 7). **B.** Fluorescence output of the *ptsG* reporter was measured in parental, evolved and single-mutant strains during cultivation in CGXII with 2% (w/v) glucose in microtiter plates. The maximum fluorescence (F) output is shown for each strain. Values were normalized to corresponding backscatter values (B) (F/ B =spec. fluorescence) **C.** Time course data of *ptsG* expression represented by the specific fluorescence of *ptsG* reporter. Corresponding growth curves are shown in Figure 5. The kink observed for the specific YFP fluorescence is a result of delayed chromophore maturation due to oxygen limitation in exponentially growing cells.

References

1. **Kinoshita S, Udaka S, Shimono M.** 1957. Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms. J Gen Appl Microbiol **50:**331-343.

2. **Baumgart M, Unthan S, Rückert C, Sivalingam J, Grünberger A, Kalinowski J, Bott M, Noack S, Frunzke J.** 2013. Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. Appl Environ Microbiol **79:**6006-6015.

3. **Schäfer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Pühler A.** 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene **145:**69-73.

4. **Cremer J, Eggeling L, Sahm H.** 1990. Cloning the Dapa Dapb Cluster of the Lysine-Secreting Bacterium *Corynebacterium-Glutamicum*. Mol Gen Genet **220:**478-480.

5. **Kalinowski J, Bathe B, Bartels D, Bischoff N, Bott M, Burkovski A, Dusch N, Eggeling L, Eikmanns BJ, Gaigalat L, Goesmann A, Hartmann M, Huthmacher K, Krämer R, Linke B, McHardy AC, Meyer F, Mockel B, Pfefferle W, Pühler A, Rey DA, Rückert C, Rupp O, Sahm H, Wendisch VF, Wiegrabe I, Tauch A.** 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. J Biotechnol **104:**5-25.

6. **Engels V, Wendisch VF.** 2007. The DeoR-type regulator SugR represses expression of *ptsG* in *Corynebacterium glutamicum*. J Bacteriol **189:**2955-2966.

7. **Gaigalat L, Schluter JP, Hartmann M, Mormann S, Tauch A, Pühler A, Kalinowski J.** 2007. The DeoR-type transcriptional regulator SugR acts as a repressor for genes encoding the phosphoenolpyruvate:sugar phosphotransferase system (PTS) in *Corynebacterium glutamicum*. BMC Mol Biol **8:**104.

6.6 Supplemental material to: A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria

Supplemental Tables, Figures and Movies

A prophage encoded actin-like protein required for efficient viral DNA replication in bacteria

Catriona Donovan[1,3,*], Antonia Heyer[2,*], Eugen Pfeifer, Tino Polen[2], Anja Wittmann[1], Reinhard Krämer[1], Julia Frunzke[2,‡], & Marc Bramkamp[1,3‡]

[1] Department of Biology I, Ludwig-Maximilians-University Munich, Großhaderner Str. 2-4, 82152 Planegg-Martinsried, Germany.

[2]Institut für Bio- und Geowissenschaften, IBG-1: Biotechnologie, Forschungszentrum Jülich, D-52425 Jülich, Germany;

[3] Institute for Biochemistry, University of Cologne, Zülpicherstr. 47, 50674 Cologne, Germany.

Strain construction

The in-frame deletion mutant of the *alpC* gene (cg1890) was constructed *via* the two-step homologous recombination procedure as described previously (1). The *alpC* up- and downstream regions were amplified using the oligonucleotide pairs DalpC-1/DalpC-2 and DalpC-3/DalpC-4. The resulting PCR products were used as a template for an overlap extension PCR with the oligonucleotides DalpC-1/DalpC-4. The purified PCR product of approximately 1 kb was digested with EcoRI and BamHI and cloned into the pK19*mobsacB* vector. The resulting plasmid pK19*mobsacB*-Δ*alpC* was used for performing an allelic exchange by homologous recombination (1) in the chromosome of *C. glutamicum* ATCC 13032 resulting in the mutant strain *C. glutamicum* Δ*alpC*.

The strain expressing *ecfp-alpC* (CDC020) at physiological concentration was generated by integrating *ecfp* upstream of and in-frame with *alpC*. As describe above, a two-step homologous recombination procedure was used (1). The *ecfp* gene was PCR amplified using the primer pair eCFP-SalI-F/eCFP-XbaI-R. The resulting PCR product was restriction digested with SalI and XbaI and ligated into an identically digested pK19*mobsacB* vector, placing *ecfp* in the middle of the multiple cloning site. The region upstream of the *alpC* start codon and the first 0.5 kp of *alpC* were amplified using primer pairs Alp-up-Hind-F/Alp-up-Sal-R and Alp-D-Xba-F/Alp-D-Bam-R, respectively. The resulting PCR products were digested with HindIII/SalI and XbaI/BamHI, respectively and sequentially ligated into a pK19*mobsacB-ecpf* vector. This vector was used for allelic replacement at the *alpC* locus, integrating *ecfp* at the 5' end of *alpC*. Chromosomal integration was confirmed by PCR.

To generate a strain overproducing *alpC-cfp* (CDC021), *cfp*, including the stop codon, was amplified using primer pair CFP-SacI-F/cfp EX2 Eco R ms. The resulting PCR product was restriction digested with SacI and EcoRI and ligated into an identically treated pEKEx2

vector. Subsequently, *alpC*, lacking the stop codon, was amplified using primer pair AlpC-SalI- F/ AlpC-BamHI-os-R, restriction digested with SalI and BamHI and ligated into the pEKEx2-*cfp* vector. The resulting vector (pCD129) was transferred into *C. glutamicum* resulting in extrachromosomal, IPTG-inducible expression of *alpC-cfp*.

To generate a strain expressing *alpC*^{D301A}*-cfp* (CDC022), pCD129 was used as a template and subjected to site directed mutagenesis using primer pair 1890-D301A-F/1890-D301A-R, giving rise to vector pCD130. Plasmids were verified by sequencing. The resulting vector was transferred into *C. glutamicum*, giving rise to strain.

For plasmid-encoded expression of *alpC-mCherry*, the *alpC* orf was fused to the gene encoding the mCherry fluorescent protein and cloned into the vector pEC-XC99E under control of the native P_{alpAC} promoter. The fragments were amplified using the primer pairs PalpA-F/PalpA-R, AlpC_F/AlpC_R and mCherry_F/mCherry_R and the resulting fragments were assembled using Gibson Assembly (2).

For an in-frame deletion of *alpA* (cg1891) the plasmid pK19*mobsacB-ΔalpA* was constructed using the oligonucleotides DalpA-1/DalpC2 and DalpA-3/DalpA-4. For deletion of *alpA* in the strain CDC020 (ATCC 13032 Δ*alpA alpC*::*ecfp-alpC*), in which *alpC* is genomically replaced by *ecfp-alpC*, the flanking downstream region of *alpA* was amplified with the oligonucleotides DalpA-3 and DalpA-cfp-alpC-4. As a template DNA for this PCR served genomic DNA of strain CDC020. In an overlap extension PCR, the PCR products were combined by amplification with the oligonucleotides DalpA-1 and DalpA-4 or DalpA-cfp-alpA-4, respectively. After restriction with EcoRI and BamHI and ligation with pK19*mobsacB,* the plasmids pK19*mobsacB-ΔalpA* and pK19*mobsacB-*Δ*alpA-cfp-alp* were obtained. The construction of an in-frame deletion mutant of *alpA* and the deletion of *alpA* in

the strain CDC020 were perform as described for Δ*alpC* resulting in the mutant strains Δ*alpA* and *ΔalpA alpC::ecfp-alpC*.

To localize AlpA in single cells, an AlpA-eYFP protein fusion was used. For this purpose, the natural promoter P*alpA* and *alpA*, excluding the stop codon, were amplified by PCR with the oligonucleotides pairs PalpA-BamHI-fw and alpA-link-rv. To amplify *eyfp* the oligonucleotides eyfp-link-fw and eyfp-SalI-rv were used, thereby adding the sequence 5'-GGCGCTGCTGGC-3' in front of *eyfp* as a linker sequence. Both PCR products were combined in an overlap-extension PCR with the oligonucleotides PalpA-BamHI-fw and eyfp-SalI-rv and digested with BamHI and SalI. Subsequently, the product was ligated into pJC1, resulting in the plasmid pJC1-*PalpA-alpA-eyfp*. For co-localization studies the strain ATCC 13032 Δ*alpA alpC*::e*cfp-alpC* was transformed with the plasmid resulting in strain ATCC 13032 *ΔalpA alpC::ecfp-alpC* P_{alpA}-*alpA-eyfp*.

For co-visualization of eCFP-AlpC filaments and CGP3 prophages (CDC023), strain CDC020 was used as the background strain and the CGP3 prophage was tagged as described previously (3).

For co-visualization of AlpC-mCherry and AlpA-eYFP, wild type *C. glutamicum* was co-transformed with plasmids pJC1-*PalpA-alpA-eyfp* and pEXC-XC99E-P*alpC-alpC-mCherry* resulting in strainATCC 13032 *alpC-mCHerry alpA-eyfp*.

For heterologous overexpression of His_{10}-AlpC the *alpC* gene (cg1890) was PCR amplified using primers AlpC-et-F and AlpC-et-R, restriction digested with XhoI and BamHI and ligated into pET16b (Novagen). The resulting plasmid (pCD115) was transformed into BL21(DE3) pLysS for heterologous protein production. Site directed mutagenesis was

employed to generate pCD116 (His_{10}-$AlpC^{D301A}$). For heterologous overexpression of AlpA the gene cg1891 (*alpA*) was amplified with the oligonucleotides alpA-fw and alpA-rv. The PCR product and the plasmid pET-TEV were digested with NdeI and EcoRI and ligated. Thereby, the sequence of a His_{10}-tag was fused to the 5'- end of *alpA*, resulting in the plasmid pET-TEV-*alpA*. Subsequently, *E. coli* BL21(DE3) was transformed with the plasmid for heterologous protein production.

Table S1: Bacterial strains, plasmids and oligonucleotides.

Strains	Relevant characteristics	Source or reference
***C. glutamicum* ATCC 13032**	Biotin-auxotrophic wild type	(4)
ATCC 13032 Δ*alpC*	In-frame deletion of cg1890 (*alpC*)	This study
CDC020	ATCC 13032 with an in-frame allelic replacement of *alpC* to *ecfp-alpC*	This study
CDC021	ATCC 13032, IPTG-inducible extra-chromosomal copy of *alpC-cfp*, Kan^r	This study
CDC022	ATCC 13032, IPTG-inducible extra-chromosomal copy of $alpC^{D301A}$-*cfp*, Kan^r	This study
ATCC 13032::pLAU44-CGP3-Spec	ATCC 13032 derivative containing plasmid pLAU44-CGP3-Spec integrated into the cg1905-cg1906 intergenic region	(3)
CDC023	ATCC 13032 derivative containing plasmid pLAU44-CGP3-Spec integrated into the cg1905-cg1906 intergenic region in the strain CDC020	This study
CDC024	ATCC 13032 derivative containing plasmid pLAU44-CGP3-Spec integrated into the cg1905-cg1906 intergenic region in the strain ATCC 13032 Δ*alpC*	This study
ATCC 13032 Δ*alpA*	ATCC 13032 with in-frame deletion of cg1891 (*alpA)*	This study

ATCC 13032 *ΔalpA alpC::ecfp-alpC*	ATCC 13032 with in-frame deletion of cg1891 (*alpA*) in the strain CDC020 with an allelic replacement of *alpC* to *ecfp-alpC*	This study
ATCC 13032 *ΔalpA alpC::ecfp-alpC* P_{alpA}-*alpA-eyfp*	ATCC 13032 with in-frame deletion of *alpA* (cg1891) in the strain CDC020 with an allelic replacement of *alpC* to *ecfp-alpC* and plasmid-bound expression of *alpA-eyfp* (pJC1-P_{alpA}-*alpA-eyfp*)	This study
ATCC 13032 *alpC-mCherry*	ATCC 13032 derivative containing plasmid pEXC-XC99E-P*alpC-alpC-mCherry* Cm^r	This study
ATCC 13032 *alpC-mCherry alpA-eYFP*	ATCC 13032 derivative containing plasmid pEXC-XC99E-P*alpC-alpC-mCherry* Cm^r and plasmid-bound expression of *alpA-eyfp* (pJC1-P_{alpA}-*alpA-eyfp*) Kan^r	This study
***E. coli* DH5α**	*supE44* Δ*lacU169* (ϕ80*lacZ*DM15) *hsdR17* *recA1* *endA1 gyrA96 thi-1 relA1*	Invitrogen
***E. coli* BL21 (DE3) pLysS**	F^- *ompT hsdS*(r_B^- m_B^-) *gal dcm* λ(DE3) pLysS (Cam^r) (λ(DE3): *lac*I, *lac*UV5-T7 gene 1, *ind*1, *sam*7, *nin*5)	Promega
Plasmids	**Relevant characteristics**	**Source or reference**
pK19*mobsacB*	Kan^r; vector for allelic exchange in *C. glutamicum*; (pK18 $oriV_{E.c.}$, *sacB*, *lacZ*α)	(5)
pK19*mobsacB-ΔalpA*	Kan^r; pK19*mobsacB* derivative containing a crossover PCR product covering the up- and downstream regions of *alpA* (cg1891)	This study

pK19*mobsacB*-*ΔalpA-ecfp-alpC*	Kan^r; pK19*mobsacB* derivative containing a crossover PCR product covering the flanking regions of *alpA* (cg1891) in strain CDC020 (*alpC* replaced by *ecfp*-	This study
pK19*mobsacB*-*ΔalpC*	Kan^r; pK19*mobsacB* derivative containing a crossover PCR product covering the up- and downstream regions of *alpC* (cg1890)	This study
pCD127	Integration vector, *ori pUC, Kan^r, mob sacB ecfp-alpC*	This study
pEKEx2	Kan^r; *C. glutamicum*/*E. coli* shuttle vector for regulated gene expression (P_{tac}, *lacI*^{q}, pBL1 *ori*V_{Cg}, pUC18	(6)
pCD129	Kan^r, P_{tac} *lacI*^{q} pBL1 *ori*$V_{C.g.}$ pUC18 *ori*$V_{E.c}$, AlpC$^+$-CFP	This study
pCD130	Kan^r, P_{tac} *lacI*^{q} pBL1 *ori*$V_{C.g.}$ pUC18 *ori*$V_{E.c}$, AlpC^{D301A+}-CFP	This study
pEKEx2-*yfp-tetR*	Kan^r, pEKEx2 derivative containing *yfp-tetR* under control of the P_{tac} promoter	(3)
pET16(b)	*bla PT7lac-10his lacI*	Novagen
pCD115	*bla PT7lac-10his-alpC lacI*	This study
pCD116	*bla PT7lac-10his-alpC^{D301A} lacI*	This study
pET-TEV	Kan^r; pET28b derivative for overexpression of genes in *E. coli* (pBR322 *ori*$V_{E.c.}$, *PT7, lacI*)	(7)
pET-TEV-*alpA*	Kan^r, pET-TEV derivative to overproduce AlpA with an N-terminal His_{10} tag	This study
pJC1	Kan^r, Amp^r, *E. coli – C. glutamicum* shuttle vector	(8)

pJC1-P$_{alpA}$-*alpA-eyfp*	Kan^r, pJC1 derivative containing *alpA-eyfp*, encoding an AlpA-eYFP protein fusion under the control of the native promoter P$_{alpA}$	This study
pEC-XC99E	*cat$_I$*, *lacI^q*, P$_{trc}$, *rrnB* (T1 and T2), *oriV$_{E.c}$*, *per* and *repA* (pGA1)$_{C.g.}$, - *E. coli* – *C. glutamicum* shuttle and expression vector with chloramphenicol resistance	[(9)]
pEC-XC99E-P$_{alpA}$-*alpC-mCherry*	Cm$^+$, pEC-XC99E containing *alpC-mcherry*, encoding an AlpC-mcherry protein fusion under the control of the native promoter P$_{alpA}$ (used restriction site: *Pst*I,	This study

Oligonucleotide	**Sequence (5´→ 3´) and properties[a]**
DalpC-1	ATATATGAATTCTTGTTTGTCGCTGAATACGGTG (EcoRI)
DalpC-2	*CCCATCCACTAAACTTAAACA*CACATTCACAGCGCTGGTCAT AATC
DalpC-3	*TGTTTAAGTTTAGTGGATGGG*CGCTCGATTGCAGCGAAAGC ACG
DalpC-4	TATATAGGATCCAGCGCGCCAAAGAAAACACAGAG (BamHI)
DalpA-1	TATATAGAATTCCCATTTTCGGGGTGATGGTTAC (EcoRI)
DalpA-2	*CCCATCCACTAAACTTAAACA*CGTGTCCTGTTTTTGAGCCAT GTG
DalpA-3	*TGTTTAAGTTTAGTGGATGGG*GATGTCATGGGCCAAGCGTT CG
DalpA-4	TATATAGGATCCTCGTGCAGACAAGCTGCGCGTGCC (BamHI)
DalpA-cfp-alpC-4	TATATAGGATCCGTCCTCCTTGAAGTCGATGCC (BamHI)
alpA-NdeI-fw	CGCCATATGGCTCAAAAACAGGACACGAC (NdeI)

alpA-EcoRI-rv	CCGGAATTCCTAGCGACCGAACGCTTGG (EcoRI)
PalpA-BamHI-fw	CGCGGATCCGTCATGGTGGGGCTCCATTAG (BamHI)
alpA-link-rv	*GCCAGCAGCGCC*GCGACCGAACGCTTGGCC
eyfp-link-fw	*GGCGCTGCTGGC*ATGGTGAGCAAGGGCGAGG
eyfp-SalI-rv	CGCGTCGACTTATCTAGACTTGTACAGCTCGTC (SalI)
alpA-up-fw	TGTCATGGTGGGGCTCCATTA
alpA-up-rv	TGTGTTTCCAAAAACCGCTCAAC
alpC-down-fw	AAGGACTTATGATTGCGGCAC
alpC-down-rv	CTGGGGGCTAGAGCGCGCCAA
Cg2036-up-fw	CACTTCCAAGGAAGATACACGC
Cg2036-up-rv	ATGTTAAGAGCGTAACGGATATCA
Phage-LC-for	CCCACGTTCACCCCACAAACG
Phage-LC-rev	CTAAAATGAAGCCATCGCGACC
ddh-LC-for	ACGTGCTGTTCCTGTGCATGG
ddh-LC-rev	GCTCGGCTAAGACTGCCGCT
Alp-up-Hind-F	CAGAAGCTTTGTGGGTGAAGGTACT (HindIII)

Alp-up-Sal-R	CAGGTCGACCTTGCGTGCTTTCGC (SalI)
Alp-D-Xba-F	CAGTCTAGATAATTAATACCTAGTT (XbaI)
Alp-D-Bam-R	CATGGATCCCCACTCATTACCGCC (BamHI)
eCFP-SalI-F	CATGTCGACATGGTGAGCAAGGGC (SalI)
eCFP-XbaI-R	CATCTCGAGCTTGTACAGCTCGTC (XbaI)
AlpC-SalI-F	CAGGTCGACATGACCAGCGCTGTGAAT (SalI)
AlpC-BamHI-os-R	CAGGGATCCCTTGCGTGCTTTCGCTGC (BamHI)
CFP-SacI-F	CAGGAGCTCATGGTGAGCAAGGGCGAG (SacI)
cfp EX2 Eco R mS	GCGGAATTCTTACTTGTACAGCTCGTC (EcoRI)
1890-D301A-F	CGCCACACAATCGGTGTG**GCC**GTGGGTGAAGGTACTG
1890-D301A-R	CAGTACCTTCACCCAC**GGC**CACACCGATTGTGTGGCG
AlpC-et-F	CAGCTCGAGATGACCAGCGCTGTG (XhoI)
AlpC-et-R	CAGGGATCCTTACTTGCGTGCTTTCGC (BamHI)
PalpA-F	**GCGGTATTTCACACCGCATATG**GGTGATGGTTACAGCAT CGC
PalpA-R	**ATATCTCCTTCTTAAAGTCTA**TACATGTGTCGTGTCCTGT TTTTG

AlpC-F	**TAGACTTTAAGAAGGAGATAT**ATGACCAGCGCTGTGAATGTG
AlpC-R	**TCCTCGCCCTTGCTCACCAT**CTTGCGTGCTTTCGCTGCAAT
mCherry-F	ATGGTGAGCAAGGGCGAG
mCherry-R	**AACAGCCAAGCTTGCATGCC**TTACTTGTACAGCTCGTCCATGC

[a] In some cases oligonucleotides were designed to introduce recognition sites for restriction endonucleases (recognition sites underlined, restriction endonucleases indicated in parentheses) or complementary 21mer sequences for generating overlap PCR products (italics). Amino acid mutations on mutagenesis primers are indicated in grey boldfacing. In bold, overlaps for Gibson assembly are indicated (2).

Supplementary Figures

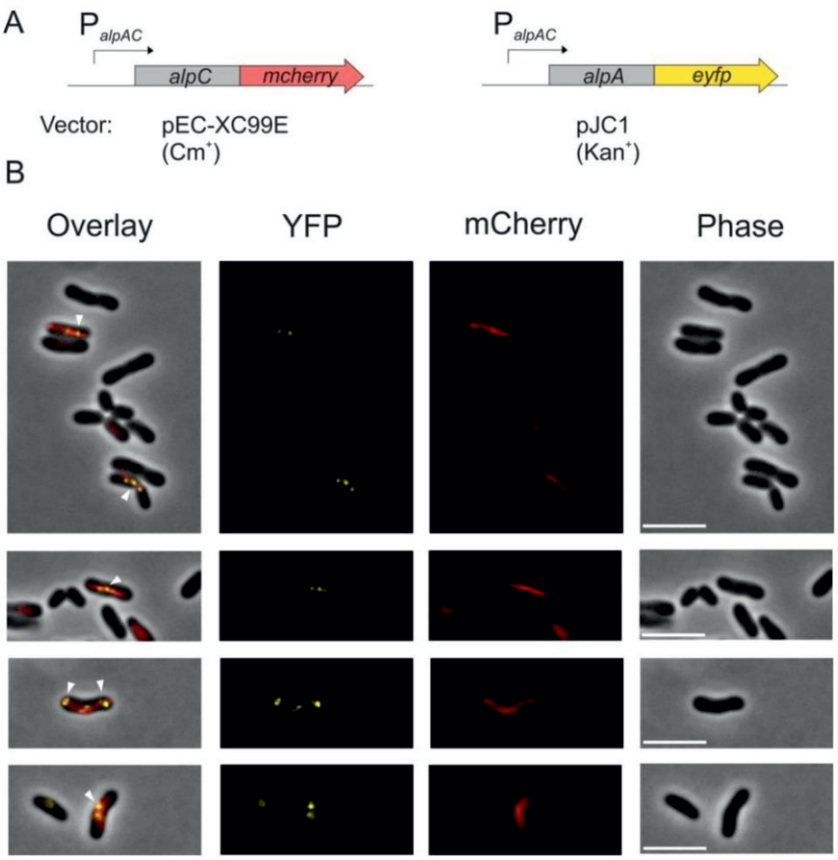


Figure S1: Co-localization studies of AlpA and AlpC filaments after mitomycin C treatment. A: Reporter constructs which were used to monitor localization of AlpA-eYFP and AlpC-mCherry in *C. glutamicum* ATCC 13032. B: Wild type cells of the strain ATCC 13032 containing both plasmids (pEC-XC99E-alpA-yfp and pJC1-alpC-mcherry) were grown in CGXII minimal media supplemented with 2% glucose, 10 µg/ml chloramphenicol and 25 µg/ml kanamycin to an OD_{600} of approximately three. Activation of the CGP3 prophage was induced by adding mitomycin C (0.6 µM). Three to five hours after induction pictures were taken by using fluorescence microscopy (length of the bar scale is 5 µm). Overall, co-localization of AlpA foci and AlpC filaments was observed in 91% of the cells (n=116).

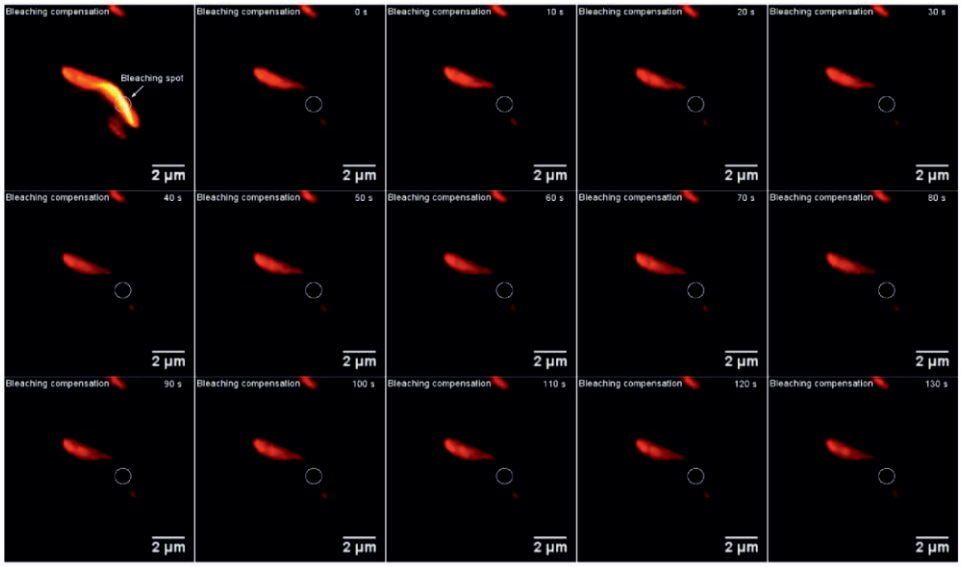


Figure S2: FRAP experiment to control for dark-state-reversal of mCherry fluorophore. To control for dark-state-reversal of the mCherry fluorophore, cells expressing AlpC-mCherry were fixing with 1% formaldehyde for 30 minutes at room temperature prior to analysis. FRAP analysis was carried out and evaluated as described in materials and methods. After bleaching a region in the cell, no recovery of mCherry was detected.

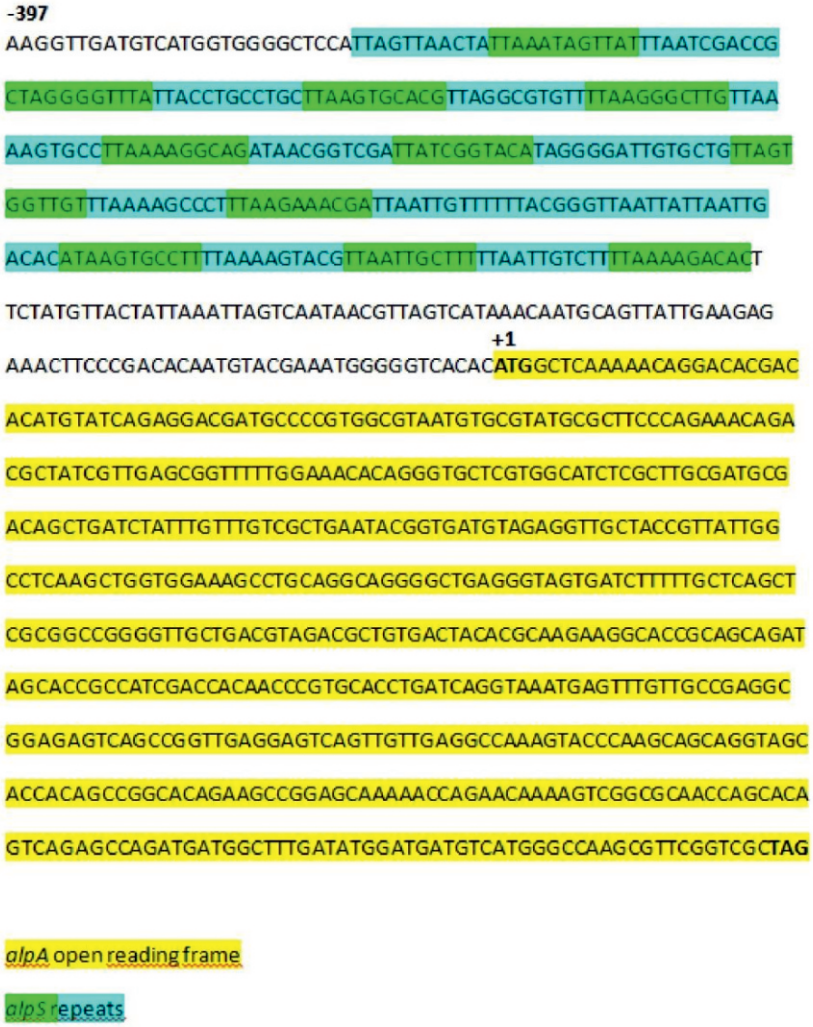


Figure S3: Conserved *alpS* repeats in the *alpAC* promoter region. Shown are 397 bases of the *alpAC* operon followed by the *alpA* orf (yellow). The *alpS* repeats, which were shown to be bound by AlpA *in vitro*, are highlighted in green and blue.

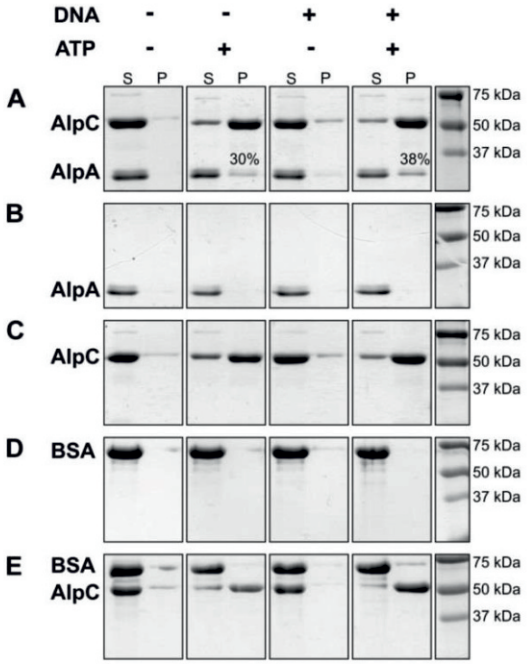


Figure S4: Co-sedimentation analysis of AlpA / AlpC interaction *in vitro*. (**A**) Equimolar concentrations of purified AlpA and AlpC were incubated with ATP, DNA or both ATP and DNA. Sedimentated protein was separated from non-sedimentated protein by centrifugation. The supernatant (S) and pellet (P) fractions were analysed by SDS-PAGE and visualized by coomassie staining. In the absence of ATP, AlpC does not polymerize and remains in the supernatant fraction, similar to AlpA. In the presence of ATP or ATP and DNA, AlpC polymerises and is found in the pellet fraction. Under these conditions, AlpA interacts and co-sediments with AlpC. (**B**) In the absence of AlpC, AlpA does not sediment. (**C**) Sedimentation of AlpC is dependent of nucleotide. (**D-E**) As a control, BSA was incubated with AlpC. Unlike AlpA, BSA does not co-sediment with AlpC.

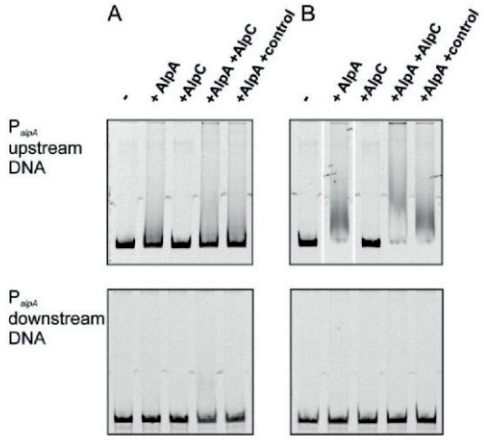


Figure S5: Evidence for an interaction of AlpA and AlpC in the presence of the upstream region of *alpAC*. 100 ng DNA of the upstream region of *alpAC* (500bp) containing the binding site of AlpA, *alpS,* were incubated with AlpA (200-fold molar excess) as indicated without ATP (A) or with ATP (B). AlpC (30-fold molar excess) or the control protein HrrA (2-fold molar excess) were added as indicated. (B) To induce polymerization of AlpC filaments, ATP was added to the reaction mixture. The downstream region of the *alpAC* operon (500 bp) was used as control fragment in both experiments. After20 min of incubation the samples were separated on a 10% non-denaturing polyacrylamide gel and stained with SYBR green I. No difference was observed without ATP, but the addition of ATP resulted in a super shift of AlpA-bound DNA. This provides additional evidence for a direct interaction of AlpA with AlpC filaments in the presence of *alpS* DNA.

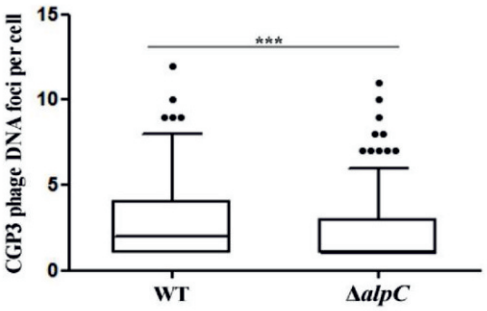
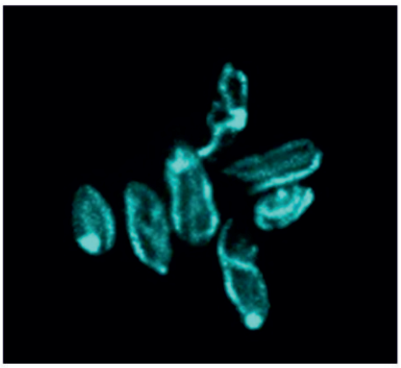
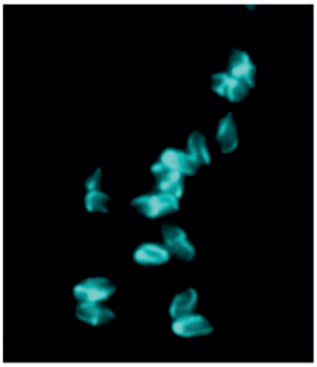


Figure S6: Induction of CGP3 phage is reduced in the absence of AlpC, *in vivo*. The number of CGP3 phage DNA particles was measured in both WT (strain ATCC 13032::pLAU44-CGP3-Spec)and Δ*alpC* (strain CDC024) cells (n≥210). Cells were treated with mitomycin C to induce the CGP3 phage. On average WT cells contained 2.9 CGP3 foci per cell, while Δ*alpC* contained 1.9. ***T-test assessment shows that there is a significant difference between the mean values (p value <0.0001).

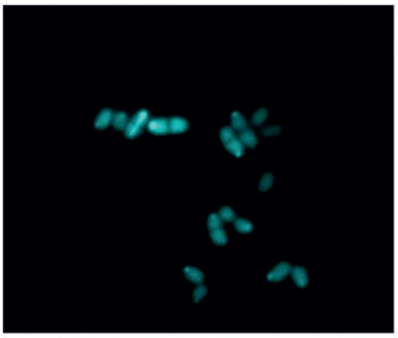
Legend to supplemental movies



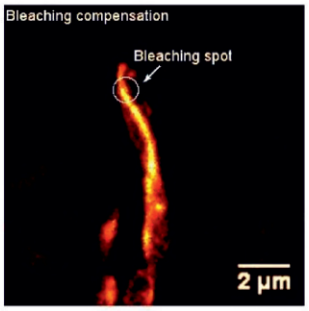
Movie S1: AlpC assembles into filaments. Z-stack through *C. glutamicum* cells expressing AlpC-CFP. Cells were grown to logarithmic growth phase and imaged using a Zeiss Axioimager M1 as described in material and methods. A still image is shown above.



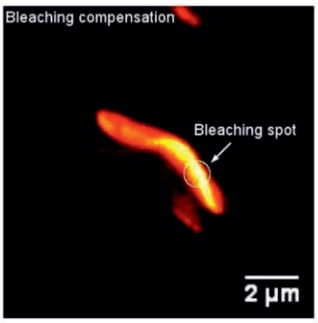
Movie S2: AlpC filaments are dynamic. Time lapse image series of *C. glutamicum* cells expressing AlpC-CFP. Cells were grown to logarithmic growth phase and imaged using a Zeiss Axioimager M1 as described in material and methods. A still image is shown above.



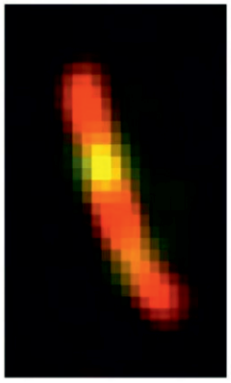
Movie S3: AlpC filament dynamics require nucleotide hydrolysis. Time lapse image series of *C. glutamicum* cells expressing AlpC^{D301A}-CFP. Cells were grown to logarithmic growth phase and imaged using a Zeiss Axioimager M1 as described in material and methods. Many cells only show foci or patches of AlpC-CFP, but no filaments. Indeed, only 1.6% of all cells expressing AlpC^{D301A}-CFP show clear filaments. Time lapse analysis shows that these filaments are not as dynamic as wild type AlpC-CFP filaments. A still image is shown above.



Movie S4: Analysis of AlpC-mCherry filament dynamics by FRAP. Part of AlpC-mCherry filaments were bleached using a 561 nm laser. The recovery of AlpC-mCherry in the bleached area was monitored by time lapse analysis, where images were acquired every 10 seconds prior to bleaching.



Movie S5: Control for dark-state-reversal of mCherry fluorophore. AlpC-mCherry expressing cells were fixing with 1% formaldehyde for 30 minutes at room temperature prior to analysis. FRAP analysis was carried out and evaluated as described in materials and methods. A region within the cell was bleached with the 561 nm laser and recovery of AlpC-mCherry was monitored over time, with image acquisition every 10 seconds. There was no recovery of mCherry in the bleached region.



Movie S6: Time lapse analysis of AlpA-eYFP and AlpC-mCherry. Under CGP3 phage induction conditions, the dynamics of AlpA-eYFP and AlpC-mCherry was analyzed. Images were acquired every 5 seconds for a total of 65 seconds. The AlpA-eYFP focus moves from a centrally located position to the cell pole.

References

1. Niebisch, A. and Bott, M. (2001) Molecular analysis of the cytochrome bc1-aa3 branch of the *Corynebacterium glutamicum* respiratory chain containing an unusual diheme cytochrome c1. *Archives of microbiology*, **175**, 282-294.
2. Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., Stockwell, T.B., Brownley, A., Thomas, D.W., Algire, M.A. *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, **319**, 1215-1220.
3. Frunzke, J., Bramkamp, M., Schweitzer, J.E. and Bott, M. (2008) Population Heterogeneity in *Corynebacterium glutamicum* ATCC 13032 caused by prophage CGP3. *Journal of bacteriology*, **190**, 5111-5119.
4. Kinoshita, S., Udaka, S. and Shimono, M. (1957) Studies on the amino acid fermentation: I. Production of L-glutamic acid by various microorganisms. *J. Gen. Appl. Microbiol.*, **3**, 193-205.
5. Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G. and Pühler, A. (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene*, **145**, 69-73.
6. Eikmanns, B.J., Kleinertz, E., Liebl, W. and Sahm, H. (1991) A family of *Corynebacterium glutamicum/Escherichia coli* shuttle vectors for cloning, controlled gene expression, and promoter probing. *Gene*, **102**, 93-98.
7. Bussmann, M., Baumgart, M. and Bott, M. (2010) RosR (Cg1324), a hydrogen peroxide-sensitive MarR-type transcriptional regulator of *Corynebacterium glutamicum*. *The Journal of biological chemistry*, **285**, 29305-29318.
8. Cremer, J., Eggeling, L. and Sahm, H. (1990) Cloning the *dapA dapB* cluster of the lysine-secreting bacterium *Corynebacterium glutamicum*. *Mol. Gen. Genet.*, **220**, 478-480.
9. Kirchner, O. and Tauch, A. (2002) Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. *Journal of biotechnology*, **104**, 287-299.

Danksagung

Ganz besonders möchte ich mich bei Juniorprof. Dr. Julia Frunzke für die vorbildliche Betreuung, das Überlassen des interessanten Themas sowie über zahlreiche spannende Diskussionen Bedanken. Besonders in schwierigen Momenten während der Promotion, wo man unter anderem den roten Faden verloren hat, schaffte Sie es immer durch Ihren Enthusiasmus, Ihre zuverlässige und beständige Erreichbarkeit die Motivation wieder aufzubauen, einen auf die richtige Bahn wieder zu lenken um somit das Beste aus einem rauszuholen.

Auch ein ganz besonderer Dank gilt Prof. Dr. Michael Feldbrügge für das Übernehmen der Zweitgutachtens und für die interessanten Diskussionen während einiger Konferenzen.

Des Weiteren bedanke ich mich bei Prof. Dr. Michael Bott, Dr. Iris Eggeling, Iris Heitzer und Marion Dreßen-Combach für das Ermöglichen der Promotion am Forschungszentrum Jülich und für die Hilfe mit dem dazu verbundenen bürokratischen Aufwand.

Bei den ehemaligen und jetzigen Mitgliedern der AG Frunzke, Antonia Heyer, Nurije Mustafi, Eva Davoudi (ehem. Hentschel), Meike Baumgart, Cornelia Gätgens, Isabel Huber, Max Hünnefeld, Regina Mahr, Raphael Freiherr von Boeselager, Marc Keppel, Johanna Wiechert, Dominik Wolf, Kira Ludwig, Nguyen van Nhi Le, Sabrina Faßbender, Simon Lemm, möchte ich mich natürlich für die besonders gute Unterstützung, die lustigen Mittagsgespräche, die zahlreichen Grillabende, die lustigen und spannenden Ausflüge/Retreats bedanken. Nicht zu vergessen, möchte ich auch für tolle Hilfe im Labor oder das besonders motivierende Anfeuern bei den BSG-Läufen bedanken!

Allen Mitarbeiten des IBG-1 Instituts danke ich auch für die super angenehme Arbeits-atmosphäre, die große Hilfsbereitschaft und die sehr netten Gespräche sowohl während der Sommer- als auch Weihnachtsfeste.

Besonders bedanken möchte ich mich bei Kim J. Kraxner, Maike Kortmann, Julia Schulte, Lucas Horst, Andreas Küberl, Cedric Davoudi, Alina Spielmann, Natalie Wolf, Michael Voigt, Tina Mack, Nicolai Kallscheuer, Lion Flachbart, Alexander Grünberger, Iska Steffens, Loreen Andrick, allen Mitglieder des DFG-SPP1617 Schwerpunktes und allen weiteren Institutsmitgliedern, die ich hier aus Platzgründen nicht mehr auflisten konnte, für die lustigen Grillfeiern, die anspruchsvollen Boulder-Exkursionen, die leckeren Kuchenschlachten, die glorreichen Pub-Abende und für alle weitere unvergessbaren Momente!

Nicht zu vergessen, möchte ich zum Schluss meinen Eltern, Großeltern, meinem Bruder und seiner Frau, meinen Freunden und dem Rest meiner Verwandtschaft für die immer vorhandene und stetige Unterstützung während der Doktorarbeit und für alles was danach kommt, danken!

Erklärung

Hiermit versichere ich an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine Universität Düsseldorf“ erstellt worden ist. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Eugen Pfeifer