Characterisation of actinomycete integrative and conjugative elements

te Poele, Evelien

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2008

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
CHAPTER I

General introduction

Part of the general introduction and concluding remarks is accepted for publication in Antonie van Leeuwenhoek
Ecology of the order *Actinomycetales*

The order *Actinomycetales* consists of high G+C, Gram-positive bacteria, of which many species can form branching mycelium and produce spores. The order encompasses a highly versatile group of bacteria including heterotrophic, autotrophic, methylotrophic, thermophilic, and halophilic members that are present in a wide range of ecosystems and habitats. The actinomycetes comprise important pathogens, such as *Mycobacterium tuberculosis*, the causative agent of tuberculosis, but also symbionts, like *Frankia*, which induce the formation of N$_2$-fixing root nodules on diverse flowering plants.

![Phylogenetic tree](image)

**Fig. 1.** Phylogenetic tree based on 1314 bp of the 16S rRNA gene of several representatives from all currently known suborders of the order *Actinomycetales*. The scale bar represents 0.01 substitutions per nucleotide position.
Actinomycetes are ubiquitous and numerically abundant in soil, in which they play a crucial role in decomposition and nutrient cycling, as they can degrade an exceptionally wide range of natural and man-made polymers. These actinomycetes, including members of the genera Streptomyces, Saccharopolyspora and Amycolatopsis, are adapted to survival in a highly erratic and competitive soil environment. They are equipped with a wide array of enzymes for exploiting nutrients but also produce a broad range of bioactive metabolites of industrial and medical importance, e.g. compounds with antibiotic activity against fungal and bacterial competitors.

Over the past decade, it has become evident that actinomycetes are also widely distributed in marine ecosystems. Besides known actinomycete genera, new genera like Salinispora (Fig. 1) have been discovered in marine environments, which have clearly diverged from their terrestrial counterparts. These obligate marine actinomycetes have adapted to life in the ocean and require high salinity for growth. Their entirely different environmental conditions compared to those of their terrestrial relatives may have influenced the production of secondary metabolites; these marine actinomycetes may therefore represent an important source of novel secondary metabolites. For example, the genus Salinispora, which is widely distributed in tropical and subtropical marine sediments, produces novel secondary metabolites, and also represents a new source of rifamycin-like antibiotics outside the genus Amycolatopsis.

Taxonomically the Actinomycetales are divided in several suborders and families and include the families Streptomycetaceae (suborder Streptomycetaceae) and Pseudonocardiaceae (Fig. 1), the latter comprising the genera Amycolatopsis and Saccharopolyspora, the subjects of research in this PhD thesis. The genus Amycolatopsis, proposed by Lechevalier et al. (1986), currently contains 34 validly described species and can be distinguished from the other genera within the family by using genus-specific oligonucleotide primers and by a combination of chemical and morphological markers.

Given the enormous diversity and genetic versatility within the order of Actinomycetales and its worldwide distribution and occurrence in ecosystems, including those that are difficult to sample (deep subsurface, deep-sea sediments), we are only seeing the tip of the iceberg. Many novel species still await isolation or genetic dissection by for e.g. metagenomic tools and may contain a wealth of novel secondary metabolites of medical and biotechnological importance.
Chapter I

Actinomycete genomes

In recent years, several actinomycete genome sequences have become available. Within the actinomycetes, the genomes may vary largely in size. The genomes of *Streptomyces coelicolor* A3(2) \(^9\), *Streptomyces avermitilis* \(^54\) and *Saccharopolyspora erythraea* \(^106\) are amongst the largest bacterial genomes characterised (8.7, 9.0, and 8.2 Mb, respectively), while the genomes of *Salinispora tropica* \(^142\) and the pathogenic actinomycetes *Corynebacterium diphtheriae* \(^26\) and *M. tuberculosis* H37Rv \(^30\) are much smaller (5.2, 2.5 and 4.4 Mb, respectively). The genome size of three facultative symbiotic *Frankia* species varies greatly, depending on the plant host range and biogeographical history of symbiosis \(^103\). The genome size for the narrow host range strain *Frankia* sp. strain CcI3 is 5.43 Mb, that of the medium host range strain *Frankia alni* strain ACN14a is 7.50 Mb and the broad host range strain *Frankia* sp. strain EAN1pec has a genome size of 9.04 Mb.

In contrast to most bacteria that have circular chromosomes, the *Streptomyces* chromosomes characterised are linear with terminal inverted repeat sequences at both ends that are covalently linked to terminal proteins. It has been suggested that the linear chromosomes resulted from integration of linear plasmids into ancestral circular chromosomes \(^144\). The *S. coelicolor* chromosome contains a central core region containing essential housekeeping genes and is flanked by less conserved and unstable “arms” that were most likely acquired by horizontal gene transfer (HGT) and consist of genes for adaptation to a wide range of conditions in their habitats \(^9\). *Sac. erythraea* has a circular chromosome consisting of a 4.4 Mb core around the origin of replication where most housekeeping genes are located and a non-core region that makes up almost half of the total genome. This non-core region is rich in IS elements, which almost all are associated with transposases. Approximately fifty percent of these IS elements is found in two regions of 0.6 and 0.35 Mb that have substantially lower G+C content than the average and may therefore have been obtained by HGT \(^106\).

The central cores of the *Sac. erythraea* and *Streptomyces* chromosomes and the whole chromosomes of the pathogenic species *M. tuberculosis* and *C. diphtheriae* have a relatively conserved gene organisation which suggests descent from a common actinomycete ancestor \(^9,106\).

Actinomycete bioactive secondary metabolites

Actinomycetes are the most important source of biologically active microbial products, including many medically and commercially important antibiotics \(^11\). Well-known examples are the broad-spectrum antibiotic erythromycin A produced by *Sac. erythraea* \(^90\), vancomycin produced by *Amycolatopsis orientalis* \(^158\), currently the last resort against
methicillin-resistant *Staphylococcus aureus*, and rifamycin which is produced by *Amycolatopsis mediterranei* and used to fight tuberculosis and leprosy.

Exploration of the genome sequences of *Sac. erythraea*, *S. coelicolor* and *S. avermitilis* uncovered a large number of putative antibiotic biosynthesis clusters. The antibiotic biosynthetic potential of *Amycolatopsis* may also be much larger than is presently recognised. Screening of *Amycolatopsis* sp. strain UM16 revealed the potential to produce glycopeptide antibiotics and Type-II polyketide antibiotics. *A. orientalis* harbours at least 10 gene clusters for the synthesis of secondary metabolites other than vancomycin. Using genus-specific primers and novel selective media, Tan *et al.* (2006) recently isolated and characterised several novel *Amycolatopsis* isolates from relatively small sampling sites from several locations in Australia and from a park in the United Kingdom. In doing so, they showed that species diversity within the genus is much higher than previously thought. Dedicated isolation strategies conducted worldwide in terrestrial and marine ecosystems may yield a multitude of novel species that represent a yet unexplored source of new antibiotics and other secondary metabolites.

The widespread use of antibiotics in medicine and agriculture has not only cured many diseases, saving millions of lives, but paradoxically also resulted in an increased antibiotic resistance in pathogenic bacteria making them almost impossible to fight. It has been suggested that many of the resistance genes found in pathogenic bacteria originate from antibiotic producing actinomycetes. They often encode resistance genes to protect themselves against the harmful effects of the antibiotics they produce. These resistance genes might have been acquired by pathogens via HGT. The alarming worldwide resistance of pathogenic organisms to a broad range of unrelated antibiotics urges for the discovery and development of novel antibiotics.

**Mobile genetic elements**

Mobile genetic elements (MGEs) mediate the exchange of DNA within a genome (intracellular) or between individuals (intercellular) and are therefore key players in maintaining genome plasticity and HGT. MGEs are traditionally classified as transposable elements, plasmids or bacteriophages, but this classification has become redundant due to the existence of many intermediate forms. For example, integrative and conjugative elements (ICEs) have both plasmid- and phage-like features. MGEs can be regarded as modular structures consisting of different functional modules, each including the genes and sequences that are required for a specific biological function. MGEs evolve by exchange or acquisition of modules between similar and different types of MGEs, resulting in various combinations of functional modules. A functional module can perform a local function, which is defined by its biochemical interactions with other molecules, and an integrated function, which is the role it has in larger cellular
processes of the MGE or host cell. All the local and integrated functions can be grouped in three global categories, i.e. intracellular mobility, intercellular mobility, and stability. Several functional modules can perform more than one integrated function and may therefore be part of more than one category. For example, rolling circle replication (RCR) can be used for autonomous replication of a plasmid (intracellular mobility), but can also be an essential step in conjugal transfer (intercellular mobility). Moreover, a module can also be used by different elements to fulfil the same integrated function. For example, RCR can be used by plasmids and phages.

Transposable elements (TEs), i.e. insertion sequences (IS) and transposons, are numerous and widely distributed among bacteria. They play important roles in genome diversity causing insertional mutations, genomic rearrangements and genome expansion, as they move from one chromosomal location to another within a cell. TEs can mobilise themselves to other cells by hitchhiking on conjugative plasmids or bacteriophages. TEs contain a recombination module that consists of a recombinase catalysing the translocation of DNA from a donor site to a target site (either within a genome or exchange between vector and genome), and a conserved short DNA sequence, the recombination site that is specifically recognised and cleaved by the recombinase. The simplest TE, the IS element, only consists of a recombinase required for transposition, and mostly has two short terminal inverted repeats. Transposons often carry additional accessory genes, such as antibiotic- or heavy metal resistance genes, or catabolic genes.

Three unrelated families of recombinases can be distinguished, i.e. the DD-E transposases, containing the conserved amino acid motif DD-E as part of the catalytic site, tyrosine (Tyr) recombinases and the serine (Ser) recombinases. DD-E transposases translocate DNA to more or less random sites and generate small duplications on insertion. Transposition mediated by Tyr and Ser recombinases differs from DD-E transposition in that recombination is mostly restricted to specific sites, restores the original site after excision and integration, and leaves no small duplications. Site-specific recombinases covalently bind a tyrosine (or serine) residue to the DNA substrate. They can perform several integrated functions, such as reduction of plasmid dimers to monomers, co-integrate resolution, and promote the integration and excision of a wide range of integrative and conjugative elements, pathogenicity islands and temperate bacteriophages. Bacteriophage λ encodes a Tyr recombinase/integrase that mediates site-specific integration into the 3’ end of a tRNA gene leaving the gene intact after integration.
Fig. 2. Overview of the modular structure of mobile genetic elements. Adapted from Toussaint & Merlin (2002). The three global categories of the integrated functions intracellular mobility, intercellular mobility, and stability, are depicted in the middle of the figure. The global categories are surrounded by the functional modules. Arrows pointing from a functional module towards a global category indicate that this functional module can perform an integrated function belonging to that category. Several MGEs are shown in the periphery and are connected by arrows with the functional modules from which they can be built. AICE, actinomycete integrative and conjugative element (see below).

**Integrative and Conjugative Elements**

The name integrative and conjugative elements (ICEs) was initially proposed by Burrus *et al.* (2002) for a diverse group of MGEs, which have both plasmid-and bacteriophage-like features. ICEs are present in all major divisions of bacteria and include mobile genomic islands and conjugative transposons. They are characterised by their prophage-like mode of maintenance, i.e. replication along with the host chromosome, and their ability to excise, conjugate to a new host and integrate in the host chromosome by site-specific recombination, irrespective of the specificity and mechanism of integration and conjugation.²²
The ICE backbone is composed of three modules, which are involved in maintenance, conjugation and regulation. ICEs often contain genetic elements, such as transposons and insertion sequences as well as genes encoding specific recombinases. These genetic elements and recombinases mediate the acquisition of additional modules encoding functions, such as resistance and metabolic traits, which confer a selective advantage to the host under certain environmental conditions. For example, the 100 kb ICE SXT of *Vibrio cholerae*, the agent of cholera, encodes resistance to multiple antibiotics.

Integration of ICEs is mediated by site-specific recombinases, mainly tyrosine recombinases, which function in many different hosts and have a minimal requirement for specific host factors. These features may contribute to the extremely broad host range of many ICEs and hence their important role in the spreading of antibiotic resistance genes.

Prior to conjugative transfer, ICEs excise from the host chromosome to form a circular, mostly non-replicative molecule that is nicked at the origin of transfer and is transferred to recipient cells as a single-stranded DNA (ssDNA) intermediate. Excision and transfer of ICEs can be triggered by different factors, such as sub-inhibitory concentrations of tetracycline in the case of Tn916, the conjugative transposon from *Enterococcus faecalis*, and the SOS response triggering transfer of SXT. In the latter case, the SOS response causes release from repression by a phage lambda repressor orthologue SetR encoded by SXT and subsequent derepression of the excision and transfer functions.

**Actinomycete Integrative and Conjugative Elements**

The focus of this thesis is on a special class of ICEs found in actinomycetes. These actinomycete ICEs (AICEs) are maintained integrated in a specific tRNA gene in the host chromosome. Unlike other ICEs, they have the ability to replicate autonomously like a plasmid. The majority are self-transmissible and a number of them can mobilise chromosomal markers. However, genes encoding clear beneficial functions for the physiology of the hosts are not, or very rarely, found on these elements. Several AICEs have been described in literature (Table 1), but only a few of these have been characterised in detail. So far, little is known about their distribution in actinomycetes and their role in the evolution of their host. In this study, we show that AICEs are present in several actinomycete genera, and provide more insight into their evolutionary relationships (Chapters III and IV). Furthermore, we provide data that suggest that the AICEs play an important role in host genome plasticity and HGT (Chapter V).
Table 1. Overview of AICEs and several element-specific functions

<table>
<thead>
<tr>
<th>AICE</th>
<th>strain</th>
<th>reference</th>
<th>size (kb)</th>
<th>tRNA</th>
<th>replication</th>
<th>conjugation</th>
<th>pock formation</th>
<th>CMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMEA100</td>
<td>A. mediterranei</td>
<td>(Moretti et al., 1985)</td>
<td>23.3</td>
<td>Phe</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pMEA300</td>
<td>A. methanolica</td>
<td>(Vrijbloed et al., 1994)</td>
<td>13.3</td>
<td>Ile</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pSE211</td>
<td>Sac. erythraea</td>
<td>(Brown et al., 1988a)</td>
<td>17.2</td>
<td>Phe</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pSE101</td>
<td>Sac. erythraea</td>
<td>(Brown et al., 1988b)</td>
<td>10.9</td>
<td>Thr</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pMR2</td>
<td>M. rosaria</td>
<td>(Hosted Jr et al., 2005)</td>
<td>11.2</td>
<td>Phe</td>
<td>+</td>
<td>nd d</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>pSAM2</td>
<td>S. ambofaciens</td>
<td>(Pernodet et al., 1984)</td>
<td>10.9</td>
<td>Pro</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SLP1</td>
<td>S. coelicolor</td>
<td>(Bibb et al., 1981)</td>
<td>17.3</td>
<td>Tyr</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pSG1</td>
<td>S. griseus</td>
<td>(Cohen et al., 1985)</td>
<td>16.9</td>
<td>Ser</td>
<td>+</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>pJ110</td>
<td>S. parvulus</td>
<td>(Hopwood et al., 1984)</td>
<td>13.6</td>
<td>nd</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pD408</td>
<td>S. glaucescens</td>
<td>(Hopwood et al., 1984)</td>
<td>15.05</td>
<td>nd</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

a reference to paper first describing the element; b see text for references to element-specific functions; c CMA: mobilisation of chromosomal markers; d nd: not determined

Modular organisation of AICEs

As shown in chapter IV, AICEs have a highly conserved modular structure, with functional modules for excision/integration, replication, conjugative transfer, and regulation. These modules are described in more detail below.

Site-specific integration

The AICE integration system resembles that of temperate bacteriophages in that AICEs integrate into a specific tRNA gene in the host chromosome 15. The site-specific recombination is mediated by a tyrosine integrase and occurs between two identical short sequences, the \( att \) identity segments, in the attachment site on the element (\( attP \)) and on the chromosome (\( attB \)). The \( attB \) overlaps the 3’ end of the tRNA, keeping the gene functional after integration and excision. In most cases, an integrase and excisionase are both required for excision, whereas the integrase alone can mediate integration of the element. The SLP1-encoded integrase carries out both integration and excision 17.

Some elements are able to integrate site-specifically in tRNA genes of genomes other than their own host. Integration of pSAM2 was observed in many other Streptomyces species 16 71 besides its original host Streptomyces ambofaciens and vectors based on the pSAM2 integration system were shown to integrate in a Pro-tRNA gene in Mycobacterium smegmatis 88 122. The pSE101 element integrates at multiple chromosomal locations in Streptomyces lividans, whereas it is only integrated site-specifically in a Thr-tRNA of its original host Sac. erythraea 21.
Chapter I

Replication

AICEs can replicate like genuine plasmids, but replication is not required for maintenance. The elements are maintained integrated in the host chromosome, in which they are propagated along with the host during cell division. Mutations in the replication gene or origin of replication (ori) of pSAM2 lead to loss of the transfer function, showing that autonomous replication is required for conjugal transfer 44,127.

Replication of plasmids is initiated by a chromosomal or plasmid-encoded replication initiator protein (Rep) that binds to a specific DNA sequence at the ori and guides the assembly of the replication initiation complex. This complex largely consists of proteins of the host replication machinery and in the subsequent steps of elongation and termination of replication, the plasmid often relies extensively on host proteins as well.

After binding to the ori, replication proceeds by one of two basic mechanisms. The first mechanism involves opening of the parental strands at an AT-rich region followed by RNA-priming and is used in theta replication and in the related strand displacement. The second mechanism, which is used by RCR plasmids, introduces a nick to relax the DNA and to generate a 3′-OH that is used as a primer for initiation of replication.

The ori of theta-replicating plasmids typically consist of several short DNA repeats, called iterons, to which Rep specially binds. Opening of the double-stranded plasmid DNA occurs at an AT-rich region adjacent to these iterons. Unwinding of the strands is performed by a helicase and an RNA primer is synthesised by either RNA polymerase or by a host- or plasmid-encoded primase. DNA synthesis is performed by DNA polymerase III holoenzyme, and is continuous on the leading strand and discontinuous on the lagging strand 33. For more details on the mechanism of replication and proteins involved therein see the review of del Solar et al. (1998).

The ori of RCR plasmids is called the double strand origin (DSO) and consists of a binding region and a nicking site. The DSO regions of many RCR plasmids have secondary structures, such as hairpins and cruciforms. Also binding of Rep may cause bending of the DNA and the generation of a hairpin, in which the nicking site is located in the single-stranded loop 65. This nicking site is highly conserved among plasmids belonging to the same family, whereas the binding site is less well conserved. Hence, the replication specificity of RCR plasmids is determined by the specific binding of Rep to the DSO 65.

Rep nicks the DNA and becomes covalently attached to the 5′ phosphate end, leaving a free 3′ OH end which is used as primer for leading strand synthesis. Rep recruits other proteins, such as a DNA helicase which unwinds the DNA, single-stranded DNA-binding proteins (SSBs) which coat the displaced ssDNA, and DNA polymerase III which initiates leading strand replication 65. After the leading strand has been fully displaced, Rep introduces a second single-stranded break at the DSO and ligates the ssDNA ends,
general introduction

...generating a double-stranded plasmid and a circular single-stranded plasmid. Generation of ssDNA replicative intermediates is the hallmark for RCR plasmids. The single-stranded plasmid is converted into dsDNA using the single strand origin (SSO) and the host replication machinery.

**Fig. 3.** Evolutionary relationship between (putative) replication proteins of previously characterised AICEs (red), newly found AICEs (green), actinomycete plasmids (blue) and other chromosome-encoded homologues with more than 35% amino acid identity to AICE-encoded Rep proteins (black). Abbreviations: SACE, *Sac. erythraea* NRRL23338; SCO/Sco, *S. coelicolor* A3(2); SAV/Sav, *S. avermitilis* MA-4680; Sare, *Sal. arenicola* CNS205; Strop, *Sal. tropica* CNB-440; FRAAL, *Frankia alni* strain ACN14a; Franean, *Frankia* sp. strain EAN1pec; Francci3, *Frankia* sp. strain CcI3; Mflv, *Myc. gilvum*; Mmcs, *Mycobacterium* sp. MCS; φRv2, prophage of *Mycobacterium tuberculosis* H37Rv; Mvan, *Mycobacterium vanbaalenii* PYR-1; MAV, *Mycobacterium avium* 104; nfa, *Nocardia farcinica* IFM 10152; PTH, *Pelotomaculum thermopropionicum* SI. Plasmids: pSLS of *Streptomyces laurentii*; pSG5 of *S. ghanensis*; pSVH1 of *Streptomyces venezuelae*; pSA1.1, *Streptomyces cyaneus*. The phylogenetic tree was constructed using the neighbour-joining algorithm of Mega version 4.0. The scale bar represents 0.2 substitutions per site.
Chapter I

pSAM2 of *Streptomyces ambofaciens*, the only AICE for which the replication mechanism has been elucidated thus far, replicates via the RCR mechanism \(^45\) and homologues of its replication initiator protein RepSA were present on newly found AICEs of *Streptomyces*, *Mycobacterium*, *Salinispora*, and *Frankia* (Fig. 3). MGEs replicating by RCR are divided into five families: pT181, pC194, pMV158 and pSN2 \(^33\), and a family based on *Corynebacterium* replicons \(^108\). Based on the presence of conserved motifs and a conserved nicking site, pSAM2 belongs to the pC194 family together with most of the characterised actinomycete RCR plasmids, like pSG5 of *Streptomyces ghanaensis* and pSVH1 of *Streptomyces venezuelae* \(^100\) \(^116\). However, the amino acid sequence of RepSA of pSAM2 appears not to be related to those of the Rep proteins of pSG5 and pSVH1 (Fig. 3).

On SLP1, two regions are required for autonomous replication \(^107\). The first region contains the transcriptional regulator impA, and the second region consists of two genes, at least one of which most likely encodes a replication protein; the replication mechanism of SLP1 is still unknown.

pMEA300 requires two regions for replication as well. Similar to SLP1, one region encodes a transcriptional regulator, KorA, related to ImpA, and the second region contains two genes, orfA and orfB, which are transcriptionally coupled. The orfB gene encodes the replication protein of pMEA300. Since orfA and orfB are transcriptionally coupled, a role for OrfA in replication is anticipated. However, its disruption did not significantly affect replication, therefore its function remains unknown \(^146\); Chapter II). Closely related homologues of OrfB can be found on pMEA100 and pSE211. The proteins do not resemble any previously known replication proteins, but are related to hypothetical proteins of *Mycobacterium* sp. MCS and *Mycobacterium vanbaalenii* PYR-1, and to a possible φRv2 prophage protein of *M. tuberculosis* (Fig. 3).

We have shown that OrfB, now designated RepAM, and its homologues of pMEA100 and pSE211 form a novel class of replication initiator proteins and that RepAM has unique DNA-binding properties (Chapter II). Purified pMEA300 RepAM protein binds specifically to multiple identical 8 bp repeats within its own coding sequence. The repeat sequences within this putative ori can form a stable secondary hairpin structure. Similar hairpin structures with multiple identical 8 bp inverted repeats are also present on pMEA100 and pSE211.

Conjugation

The conjugation process of AICEs appears to be similar to that of *Streptomyces* plasmids, but differs greatly from that of other bacteria. Only one protein is essential for intermycelial transfer from donor to recipient. Streptomycetes grow as mycelia and it has been suggested that the hyphal tips of a plasmid-carrying donor and a recipient...
General introduction

mycelium can grow together when they intertwine, making special aggregation systems for cell-to-cell contact, as seen for Gram-negative and unicellular Gram-positive bacteria, unnecessary 41 157. The transfer protein TraB of pSG5 was found to be localised at the hyphal tips of S. lividans, indicating that conjugation takes place at the tips of the mating mycelium 115. The main transfer proteins of AICEs and Streptomyces plasmids all contain a FtsK-SpoIIIE domain that is also present in proteins involved in chromosome partitioning during sporulation and cell division, translocating double-stranded chromosomal DNA 8 157. Mating experiments with pSAM2 indicated that it was transferred as dsDNA 113. In fact, the TraB proteins of pSG5 and pSVH1 were shown to be DNA translocators mediating the translocation of unprocessed dsDNA molecules to recipient strains 115.

A few small and often hydrophobic spread proteins (Spd) are involved in spread of the element through the septal crosswalls of the compartments of the recipient mycelium. This intramyCelial spread is accompanied by so-called pock formation, a phenotype that appears to be restricted to Streptomyces plasmids and AICEs 21 97 127 149. Pocks are macroscopically visible inhibition zones that reflect temporary growth delay of plasmid-acquiring recipient cells when these are grown in a confluent lawn of plasmid-free recipients 13. The proposed function is to delay growth until the copy number of the element is sufficiently high for efficient spread in the recipient mycelium 41 46.

Regulation

The AICEs have a prophage-like mode of maintenance as they are mainly integrated in the host chromosome. The copy number of the freely replicating form is normally low, in case of pMEA300 less than 1 free replicating copy per 5 to 10 chromosomes 145. For pSAM2 it has been observed that only under conditions favourable for conjugation, i.e. the presence of potential recipient cells, the element is excised and freely replicating molecules appear prior to transfer to recipients 113. Once transferred, the element replicates in the recipient cell, spreads in its mycelium and integrates into its chromosome.

Studies on pSAM2 showed that two element-encoded proteins, the transcriptional regulator KorSA and the hydrolase Pif, keep pSAM2 integrated in the absence of potential recipient cells 112 125. KorSA belongs to the GntR family of repressor proteins and has homologues on all the other AICEs. KorSA represses pra, the activator of the rep, xis, int operon 124 125. Initiation of transfer is established by temporary inactivation of KorSA. This leads to derepression of Pra and subsequent replication and conjugal transfer.

Pif, the pSAM2 immunity factor, contains a Nudix hydrolase motif that is required for the immunity activity 112. Nudix proteins hydrolyse the pyrophosphate bond in a Nucleoside diphosphate linked to some other moiety, X. The Nudix protein MutT of E. coli
degrades potentially mutagenic nucleotides; other Nudix proteins control the levels of metabolic intermediates and signalling compounds. Pif of pSAM2 is the first Nudix protein shown to be involved in bacterial conjugation. It prevents redundant transfer between pSAM2 harbouring cells by rendering its host unable to induce DNA transfer into neighbouring donor cells. It has been suggested that Pif prevents recognition between donor cells by modifying a host component in donor cells. However, the mechanism triggering conjugal transfer is unknown.

When *A. methanolica* is grown in medium containing autoclaved sucrose or fructose, a drastic increase in free replicating pMEA300 molecules is observed. Conceivably, a specific degradation product of the sugars somehow triggers the onset of conjugation thus leading to excision and autonomous replication of pMEA300. pMEA300 also encodes a Nudix hydrolase protein, but it is not known whether it is involved in immunity as well.

SLP1 is one of the few elements that misses a Nudix hydrolase homologue. For this element, another mechanism for preventing redundant exchange was proposed, in which the *imp* (inhibition of plasmid maintenance) locus works as a master regulator of the replication, integration and transfer functions. It encodes two proteins, ImpA and ImpC that both contain a helix-turn-helix (HTH)-motif, suggesting that they function as DNA-binding proteins. ImpA is related to the Kor transcriptional regulators found on other AICEs. Cell-to-cell contact between donor and recipient may cause derepression of the *imp* controlled genes resulting in replication and conjugation. Derepression may be caused by dilution of the Imp proteins or by signal transduction mechanisms altering *imp* expression or activity.

Several AICE-encoded *kor* genes are also part of a *kil-kor* system. The expression of certain *kil* genes is lethal in the absence of a *kor* (*kil override*) gene that controls expression of the Kil-phenotype. The *kil-kor* systems of AICEs, like the *kil-kor* systems found on conjugative and pock forming *Streptomyces* plasmids, are associated with conjugation, in which Kor transcriptionally represses transfer genes responsible for the Kil-phenotype. The proposed function is to retard growth of recipient cells, resulting in pock formation.

On pMEA300, *traA* and *traB* are most likely involved in the Kil phenotype, since disruption of the two genes strongly affected pock size. *KorA* binds to the *korA–traA* intergenic region, and may therefore regulate its own expression and that of the transfer genes. The binding region contains a 14 bp inverted repeat, that is also present directly upstream of *orfA*, one of the genes involved in replication. Furthermore, *korA* cannot be deleted from autonomous replicating derivatives of pMEA300, suggesting a role for KorA in replication of pMEA300 as well.
ImpA, the Kor-like protein encoded by the *imp* locus of SLP1, also regulates its own expression by binding to a 16 bp inverted repeat that is highly similar to the 14 bp repeat found on pMEA300, and regulates transfer by binding to a second promoter upstream of the transfer genes. However, in contrast to the situation on pMEA300, the *imp* locus represses replication of SLP1.

KorSA of the pSAM2 encoded *kil-kor* system does not bind to the transfer genes. Instead, it negatively regulates transfer indirectly by repressing *pra*, which encodes the activator of replication and transfer, to maintain pSAM2 integrated in the chromosome.

Deletion of the pMEA300-encoded *stf* gene was shown to result in reduced transformation frequencies with pMEA300 DNA in *A. methanolica*. The Stf protein contains a partial C-terminal DUF921 domain that is found in several putative regulatory proteins in *Streptomyces*, one of which also has a putative N-terminal HTH-motif and is thought to be involved in sporulation regulation. Whether Stf is the activator protein of replication and transfer of pMEA300, similar as the function of Pra on pSAM2, is not known.
Chapter I

Scope of this thesis

The aim of this thesis was to unravel the ecophysiological role of pMEA300 and related AICEs and to increase our knowledge about their putative role in modulating host genome diversity, HGT, and dissemination of antibiotic resistance. Therefore, the functional properties, origin and evolution of pMEA300 and related elements were studied. In addition, characterisation of these elements is important for the construction of genetic tools that allow the transfer, manipulation and expression of genes encoding the biosynthetic enzymes for interesting secondary metabolites. pMEA300 of *A. methanolica* has already been successfully used for this purpose. *E. coli - A. methanolica* shuttle vectors have been constructed from pMEA300 and used to clone genes involved in glucose and methanol metabolism and in aromatic amino acid biosynthesis of *A. methanolica*.

Chapter I reviews our current knowledge about the functional properties of AICEs and will be part of a review in which the distribution and evolution of these elements in actinomycetes will be discussed as well as their role in host genome plasticity and the acquisition of secondary metabolite clusters and foreign DNA via HGT.

Chapter II describes the unique DNA-binding properties of RepAM, the replication initiator protein of pMEA300 of *A. methanolica*, and shows that RepAM and its homologues of pMEA100 of *A. mediterranei* and pSE211 of *Sac. erythraea* form a novel class of replication initiator proteins.

Chapter III describes the distribution of pMEA-sequences in the genus *Amycolatopsis*. More than 100 recently isolated *Amycolatopsis* strains were examined for the presence of two unique pMEA-sequences, *repAM* and *traJ*. The molecular systematics dataset generated for a large number of *Amycolatopsis* strains provides an interesting biogeographical profile allowing us to analyse the evolution and distribution of pMEA-like elements.

Chapter IV describes the complete nucleotide sequence of pMEA100 of *A. mediterranei* and presents a structural and evolutionary comparison to the other pMEA-like elements and other AICEs. In addition, we report and discuss the sequence of two newly discovered elements in the genome sequence of *Sac. erythraea* NRRL23338. We show that AICEs are present in diverse genera of the actinomycetes and that they have a highly conserved structural organisation consisting of four functional modules. Furthermore, data are presented that suggest that they play an important role in HGT and spread of antibiotic resistance.

In Chapter V the data obtained in this thesis are summarised and discussed.