Sequencing and analysis of the DNA genome of the temperate bacteriophage AaΦ23 of *Actinobacillus actinomycetemcomitans*

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Grégory RESCH

aus Wittelsheim (68310), Frankreich

Basel, 2005

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. J. Meyer, Dr. E.M. Kulik und Prof. T.A. Bickle.

Basel, den 10 Juni 2003

Prof. M. Tanner

Dekan

TABLE OF CONTENTS

REMERCIEMENTS

Tout en écrivant les dernières ligne de cette Thèse, je me surprends à penser aux nombreuses personnes que je ne saurais suffisamment remercier tant leur aide me fut précieuse au cours de ces quatres dernières années.

En premier lieu, je tiens à exprimer ma plus profonde reconnaissance au Pr. Jürg Meyer qui m'a ouvert les portes de son laboratoire pour y effectuer le travail présenté dans cet ouvrage. Ses encouragements permanents ainsi que ses nombreuses inspirations n'ont fait qu'accroitre ma motivation et m'ont permis de mener à bien le principal objet de cette étude. Merci Jürg pour votre inestimable et constant soutien.

Je témoigne ma plus profonde gratitude au Pr. Charles Thompson qui a accepté, sans hésiter, la présidence de mon comité de Thèse.

Je suis également reconnaissant envers le Pr. Thomas Bickle pour m'avoir prodigué de nombreux conseils et m'avoir apporté bon nombre d'informations complémentaires en tant que membre de mon comité de Thèse.

Mes pensées vont vers le Dr. Eva Kulik qui m'a apporté son experience dans le domaine scientifique et qui a su répondre à mes interrogations. Je te suis reconnaissant pour le grand soin que tu as pris pour la correction de ce manuscrit et je te remercie pour la sympathie que tu m'as témoigné durant ces années.

Mes remerciements vont également aux membres actuels et anciens de l'Institut : Irmgard Hauser-Gerspach, Krystyna Lenkeit, Kis Hauck, Elfi Stadelman, Doris Blaser, Caroline Baumann, Tuomas Waltimo, Michael Hänggi, Cathy Weber, Caroline Salathé, Daniel Collet, Teresa Dobr, Karin Hinni, Gilles Kolb pour m'avoir accueilli si gentillement. Merci à Frauke Berres, Manuela Linder, Christian Kunz et Marc Weber pour leur contribution à la bonne entente régnant au sein de cet Institut.

Merci au Pr. Peter Philippsen pour m'avoir permis d'utiliser le séquenceur et au Dr. Frederic Dietrich, au Dr. Sophie Lemire-Brachat, à Sylvia Voegeli et à Anita Lerch pour m'avoir épaulé durant la longue période de séquençage. Merci de m'avoir acceuilli si chaleureusement dans votre laboratoire.

Merci au MTRA Julian Berezowski pour m'avoir permis d'utiliser le dévelopeur automatique de films photographiques.

J'adresse tout particulièrement mes sentiments les plus affectueux à ma mère qui a toujours su faire passer ses enfants au premier plan. Merci maman de m'avoir permis de poursuivre de si longues études.

Une énorme pensée pour mon beau-père. Jean-Luc, tu possèdes des qualités humaines inestimables. Merci pour tout le bonheur que tu apportes à ma mère depuis plus de cinq ans.

Stéphane, mon frère, et Diana, ma belle-sœur, je tenais à vous témoigner ma plus grande admiration. Votre amour rayonne, le simple fait de vous contempler m'a souvent permis de relativiser.

Joël, je profite de cette opportunité pour t'encourager dans tes futures études qui devraient également être longues. Ne baisse jamais les bras, fais ton maximum et surtout ne regrette rien. Tu es de toute façon le beau-fils idéal.

Lisa, je te fais un énorme bisou et je te remercie pour ton insouscience enfantine. Nous aurons dorénavant le temps de faire de longues pauses pour un calin.

Finalement, j'aimerais terminer en témoignant ma plus profonde gratitude à ma femme. Martine, en entrant dans ma vie, tu m'as apporté l'amour et l'équilibre qui m'ont été indispensables pour réussir. Tu représentes à mes yeux une personne irremplaçable, je t'aime.

SUMMARY

The entire genomic sequence of the AaΦ23 bacteriophage is presented. Its size is 43,033 bp and it has an overall molar G+C content of 41mol%. 57 potential open reading frames (ORFs) were identified. A putative function could be assigned to 20 of the 57 ORFs, i.e. to 35% of them. While another 25 ORFs share homologies with hypothetical proteins present in several bacteria or bacteriophages, 12 seem to be specific for the phage AaΦ23. Based on the genetic organization of its genome, AaΦ23 shares extensive similarities with lambdoid bacteriophages. Most functions described for lambda are also found on the AaΦ23 genome. We identified potential ORFs coding for the integrase mediating recombination at the attachment sites; the generalized recombination proteins ninB and ninG; both C1 and Cro regulators of lysogeny; the replication proteins O and P; the antitermination protein Q; two components of the lytic system; structural components of the virion head, tail, baseplate and tail fibers and both subunits of the terminase enzyme involved in DNA packaging.

The sequence of the phage attachment site *(attP*) is located 19 bp upstream of the gene coding for the integrase. The *att*L and *att*R of two lysogenic *A. actinomycetemcomitans* strains and the *att*B sites in two non-lysogenic strains were also identified. The common *att* core is 49 bp in length.

No gene coding for a known virulence factor was detected on the AaΦ23 genome. However, a DNA adenine methylase was characterized that may be functionally expressed from the prophage. This protein is of particular interest because DNA adenine methylase enzymes are involved in the virulence of several bacterial pathogens such as *Salmonella typhimurium*, *Yersinia pseudotuberculosis,* and *Vibrio cholerae*. Still, presence of ORFs coding for proteins modulating the *A. actinomycetemcomitans* pathogenicity cannot be exclude as no putative function was assigned to 37 ORFs.

Preliminary data on the putative lytic enzyme of AaΦ23 suggest that expression, even at low level, of the cloned *lys* gene is lethal for *Escherichia coli* cells.

Location on the genetic map, amino acid sequence homologies and conservation of structural key residues strongly suggest that the integrase is coded by *orf*1. Preliminary data on the functional characterization were inconclusive. Delineation of the integration elements of AaΦ23 will allow the construction of an integration vector for *A. actinomycetemcomitans*.

INTRODUCTION

Pathogenesis of periodontal diseases

Periodontal disease is a general term used to describe inflammatory processes in the periodontium caused by bacterial plaque accumulation. Microorganisms contained in the bacterial plaque are the primary etiologic agents of the different forms of periodontitis, namely localized juvenile periodontitis (LJP), adult periodontitis (AP) and rapidly progressive periodontitis (RPP). LJP is an early-onset periodontitis principally involving the permanent first molar and incisor teeth. LJP starts during the circumpubertal period of life and often results in tooth loss prior to age 20. The prevalence and severity of periodontal disease increase over time with inadequate oral hygiene (Löe et al. 1978). A positive correlation between the amount of microbial plaque and the degree of gingival inflammation has been demonstrated and this inflammation was reversible by removing the plaque from the teeth (Löe et al. 1965). However, it is not clear how much of the destruction of periodontal tissue is caused by direct effects of the bacteria and how much is due to the immune response of the host against the bacterial invasion. Periodontitis evolves in four distinct steps : the initial, the early, the established and the advanced lesion that is generally viewed as the periodontitis status (Page and Schroeder 1976). In response to bacterial plaque proliferation in the sulcus during the initial and early stages, leukocytes (predominantly neutrophils) migrate in the junctional epithelium and underlying tissue. Neutrophils, within the crevices, can phagocytose and digest bacteria but if they become overloaded they degranulate. Bacterial toxins secretion and release of enzymes following the neutrophils degranulation causes tissue damage and leads to collagen and fibroblast alteration in the marginal gingiva and to proliferation of basal cells of the junctional epithelium. In the established lesion the inflammatory infiltrate is characterized by the predominance of plasmocytes and by several

modifications of the milieu : secretion of antibodies, activation of the complement by opsonized antigens, production of several interleukines (notably IL1) and increase of the production of inflammatory prostaglandine E_2. The junctional epithelium extends apically and inflammatory cells migrate deeper in the sulcus thus forming a periodontal pocket. Later in the advanced stage, anaerobic microorganisms accumulate deeper in the pocket, some of which produce tissue-destructing enzymes like collagenases and/or toxins. Collagen destruction and junctional epithelium migration continues. Exacerbation of the inflammatory response leads to bone loss by activation of osteoclast cells.

Certain genetic and environmental factors such as compromised immunity, hormonal factors and smoking are thought to be risk factors for periodontal diseases (Page and Kornman 1997).

Bacteriology of periodontal diseases

The oral cavity represents an ecological niche for a large number of bacteria. In fact, to date, more than 500 cultivable bacterial species have been isolated from the gingival crevices (Moore and Moore 1994) and this number may be doubled if we take under consideration the non-cultivable and non-identified bacteria (Wilson et al. 1997). Facultatively anaerobic Gram–positive strains are predominant in the supragingival dental plaque of healthy subjects. In gingivitis, the proportion of Gram-negative anaerobic species increase. Of the many species present in the mouth, only 10 to 30 are putative periodonto-pathogens (Socransky and Haffajee 1994), but are all also found in periodontally healthy individuals. Only three species, *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis* and *Bacteroides forsythus* are at present considered to play an important etiological role in the initiation of periodontal diseases (World Workshop in Periodontics 1996). Additional putative periodontal pathogens include *Fusobacterium nucleatum*, *Prevotella*

intermedia, *Prevotella nigrescens* and several spirochetes (Haffajee and Socransky 1994, World Workshop in Periodontics 1996).

ACTINOBACILLUS ACTINOMYCETEMCOMITANS

General characteristics

A. actinomycetemcomitans is a Gram-negative coccobacillus, member of the family *Pasteurellaceae* which has been describe for the first time in 1912 by Klinger as *Bacterium actinomycetem comitans* (Klinger 1912). It is non-motile, facultatively anaerobic and capnophilic. It is cultivated on agar plates, supplemented with human blood. It grows into small colonies within 48 to 72 h. On agar plates, circular colonies are rough, translucent with a star-shaped morphology. Bacteria adhere together and to the agar plate with their frimbriae (Rosan et al. 1988). After repeated subculture, the rough phenotype is lost and a smooth phenotype appears with a decrease of adherence (Inouye et al. 1990). This bacterium is most closely related to *Haemophilus aphrophilus* and *Haemophilus paraphrophilus* species (for review see Zambon 1985) but does not require X (hemin) or V (nicotinamide adenine dinucleotide) growth factors (Slots 1982).

A. actinomycetemcomitans isolates were further characterized by several typing methods. Serotyping revealed six different serotypes, named a to e (Zambon et al. 1983) and more recently f (Kaplan et al. 2001). Three to nine % of the *A. actinomycetemcomitans* strains characterized did not react with any of the serotype-specific antisera tested and were therefore classified as non-serotypeable strains (Paju et al. 1998). Generally only one serotype is isolated from one patient (Zambon et al. 1983). Serotypes b (Asikainen et al. 1991, Zambon et al. 1990, Zambon et al. 1983) or c (Tinoco et al. 1997) are most frequently associated with LJP or AP. Serotype c is also found in more than 70% among isolates of

healthy persons (Asikainen et al. 1991) and in non-oral infections (Zambon et al. 1988). Other typing methods like analysis of restriction enzyme digestions and ribotyping confirmed the heterogeneity among the species *A. actinomyctemcomitans* (van Steenbergen et al. 1994).

A. actinomycetemcomitans as a periodontal pathogen

A. actinomyctemcomitans is tought to play a role as primary etiologic agent in LJP because it fulfills Socransky's postulates (1979) which identify periodontal pathogens:

- *A. actinomycetemcomitans* is present in increased numbers or proportions in diseased sites and in lower numbers or even below detection level in healthy sites (Zambon et al. 1983, Slots et al. 1980).
- Its elimination or suppression by mechanical treatment or chemotherapy brings the disease progression to a halt, whereas recurrence of the disease is directly correlated with the reappearence of the bacterium in periodontal sites (van Winkelhoff and de Graaff 1991, Slots and Rosling 1983).
- A detectable specific immune response to *A. actinomycetemcomitans* is found in diseased patients (Ebersole et al. 1982, Ranney et al. 1982, Genco et al. 1980).
- It can induce a periodontal disease in animals upon experimental infection (Beighton et al. 1989). However, animal models differ in several aspects from the human disease.
- Several putative virulence factors, such as a leukotoxin (Mangan et al.1991), a hemolysin (Kimizuka et al. 1996), a lipopolysaccharide (Kiley and Holt 1980), a bacteriocin (Hammond et al. 1987), cytolethal distending toxins (Mayer et al. 1999) and several others (for review see Fives-Taylor et al. 1999), have been identified for this bacterium.

A. actinomycetemcomitans in non-oral infections

In the literature several studies reported the recovery of *A. actinomycetemcomitans* from infections of other body sites and organs than the oral cavity, for examples, brain abscesses, heart infections and endocarditis (van Winkelhoff and Slots 1999). Thus, *A. actinomycetemcomitans* is able to mediate focal infections.

Genome sequencing project of the *A. actinomycetemcomitans* strain HK1651.

The genome of the *A. actinomycetemcomitans* strain HK1651, a clinical isolate obtained from Mogens Kilian (University of Aarhus, Denmark), has been sequenced at the Advanced Center for Genome Technology of the Oklahoma's University (unpublished results). This sequence consisting of 2,105,329 bp can be obtained by file transfer protocol on the following internet site (http://www.genome.ou.edu/act.html).

BACTERIOPHAGES

Definition of bacteriophages

First observed by Frederick Twort (1915) and, independently (Duckworth 1976), by Félix d'Hérelle (1917), bacteriophages (bacteria-eater) are simple viruses consisting of a nucleic acid molecule (single or double-stranded RNA or DNA) enveloped in a protein shell, the capsid (Schlesinger 1936). The main distinction with other viruses is their host range. In fact, bacteriophages are obligate intracellular parasites that propagate only within a bacteria host cell. Two types of bacteriophages are distinguished. The lytic phages (like *E. coli* phage T4) on the one hand, kill and lyse their hosts after replication, and the temperate

bacteriophages (like *E. coli* phage lambda) on the other, which in addition can lysogenize their hosts. After circularisation, the double-stranded DNA is integrated, by a site-specific recombination process, within the host chromosome (Campbell 1962). This integrated bacteriophage genome represents a prophage and is replicated along with the host cell chromosome as the lysogenized cell grows and divides. Repression of the viral genes required for lytic replication is maintained in progeny cells. However, these prophages can initiate a lytic cycle either spontaneously, or after induction by mitomycin C treatment, UV-light irradiation or X-rays.

Classification of bacteriophages

Phages are classified with respect to their morphology and properties of their nucleic acid genomes. They are indexed in the Universal Virus Database of the International Committee on Taxonomy of Viruses (http://www.ictvdb.rothamsted.ac.uk/index.htm). The last phage survey has been realised in the year 2000 at the Félix D'Hérelle Reference Center for Bacterial Viruses, Quebec, Canada (Ackermann 2001). More than 5000 bacteriophages have been observed under the electron microscope since 1959. 96% are tailed phages and only 186 (3,6%) have a cubic, filamentous or pleïomorphic morphology. Phages have been assigned to 13 families; they have been isolated from about 140 bacterial and several archaeal genera. The most frequently represented family is the *Siphoviridae* that comprizes 61% of the 4950 tailed phages.

Importance of bacteriophages

Ecological importance of bacteriophages

Bacteriophages are considered as being more numerous than any other organism in the biosphere, prokaryotes included. They are found in all ecosystems. The major quantity is found in the coastal oceanic waters at a density of 10^6 particles *per* ml but significant quantities are also found in soil and in the gut of mammals. The population of only the tailed phage on earth is estimated to be 10^{30} virion particles (Wommack and Colwell 2000).

Economical importance of bacteriophages

In the dairy industries, many different *lactobacilli* are used in the production of dairy products such as cheese and yoghurt. Infections of *Lactobacilllus* species by lytic bacteriophages could lead to the arrest of fermentations and thereby loss of production (Brüssow 2001). Phages have also been shown to interfere with the industrial production of amino acids from coryneform bacteria (Hongo et al. 1972) and bacitracin from *Bacillus licheniformis* (Tran et al. 1999).

Medical importance of bacteriophages

Role of bacteriophages in the horizontal transfer of bacterial virulence genes

Bacteriophages are mediators of horizontal gene transfer, called transduction, whereby a genetic trait is carried by phage particles from a bacterial donor to a recipient cell. Specialized and generalized transduction processes were recently summarized (Birge 2000).

In specialized transducing phages (Lambda, ϕ80) bacterial chromosomal sequences are covalently linked with the phage genome and are thus packaged into bacteriophage heads. Most frequently the genes flanking the integration site(s) are excised from the host chromosome along with the prophage. Many toxin genes are transduced, e.g. the staphylokinase and enterotoxin A from *Staphylococcus aureus* (Coleman et al. 1989), the erythrogenic toxins from *Streptococcus pyogenes* (Johnson et al. 1986) the diphtheria toxin produced by *Corynebacterium diphtheriae* (Laird and Groman 1976) and the cytotoxin encoded by phi CTX of *Pseudomonas aeruginosa* (Hayashi et al. 1993). The genes are located close to the bacteriophage attachment site, suggesting that these phages may have evolved as specialized transducing phages.

In generalized transducing phages (P22, P1) phage heads usually contain only bacterial DNA packaged from any chromosomal location. These phages are non-infectious, i.e., they can inject DNA but it is not replicated. If this DNA is not integrated into the host chromosome, it is lost but may be expressed transiently (abortive transduction). Finally, plasmid transduction can also occur when plasmid DNA is packaged into transducing phages.

The Vibrio pathogenicity island (VPI), which contains the gene coding for the receptor of the cholera toxin encoding filamentous phage CTXphi, can be transferred between *Vibrio cholerae* strains of the O1 serogroup via the generalized transducing phage CT-P1 (O'Shea and Boyd 2002).

Bacteriophages as therapeutical agents

Soon after the observation of the "bacteriophage phenomenon" and the description of plaque assays (d'Hérelle 1917), bacteriophages were tested as antibacterial agents in several infectious diseases as for example, typhoid fever (Davison 1922), bacillary dysentery (Beckerich and Haudroy 1922), cholera (Morrison 1932) and in *Staphylococcus* infections

(Brunoghe and Maisin 1921). The hopes were great and a first extensive review dealing with phage therapy was published in 1934 (Eaton and Stanhope 1934), when the exact nature of the so-called "bacteriophage" was not known. While d'Hérelle suspected a living organism he described as an "invisible microbe", Twort and several other eminent scientists postulated an enzyme produced by bacterial activity. Answers to the question of the bacteriophage nature came in 1936 when Schlesinger reported the composition of bacteriophage particles being 50% proteins and 50% nucleic acid (Schlesinger 1936) and later with the first electron microscopy observation of bacteriophages (Ruska 1940). In the following years, the "miracle antibacterial drugs", called antibiotics, became available. Phage therapy research was stopped in Western countries but, under d'Hérelle's influence, was continued in Eastern European countries, especially in the Ex-Soviet Union. In this country, more than 700 publications reporting the successful use of phages as therapeutic agents have been published and several phage preparations against the most common human pathogenic bacterial strains are available on the market. The increasing prevalence of multidrug-resistant bacterial pathogens led the Western scientific community to seriously re-thinking the potential applications of phages and phage-derived products in the treatment of certain infectious diseases. Promising results about the antibacterial properties of bacteriophages were recently published. A single injection of $3x10^8$ PFU of a lytic phage strain, administred 45 h after injection of 10^9 CFU of vancomycin-resistant *Enterococcus faecium* was shown to rescue 100% (n=5) of the bacteremic mice (Biswas et al. 2002). Moreover, potential antibacterial properties of purified bacteriophage derived lytic enzymes were also tested. 5 ng of the purified bacteriophage C_1 lysin were able to eliminate group A *streptococci* from 9 heavily colonized mice (Nelson et al. 2001). 100 U and 1400 U of purified bacteriophage DP-1 lytic enzyme (Pal) were respectively able to kill, in *vitro*, 15 common serotypes of *pneumococci* and to reduce to undetectable level *pneumococcal* titers in the nasal wash of 18 previously intranasally

colonized mice (Loeffler et al. 2001). More recently, an injection of 50 U of the PlyG lysin isolated from the γ phage of *Bacillus anthracis* 15 minutes after infection showed a pronounced therapeutic effect in mice as 13 of the 19 animals previously inoculated with *Bacillus cereus* (10^6 cells) recovered fully (Schuch et al. 2002).

Bacteriophages of *Actinobacillus actinomycetemcomitans*

The first *A. actinomycetemcomitans* phage (ΦAa17) was demonstrated by mitomycin C induction of a leukotoxic *A. actinomycetemcomitans* strain, which was isolated from a patient with LJP (Stevens et al. 1982). The phage was classified in Bradley's group A and its genome was found, by restriction fragments analysis, to be a double-stranded linear DNA molecule of 47kb (Stevens et al. 1993). No clear correlation between ΦAa17 infection susceptibility and the leukotoxin production by strains of *A. actinomycetemcomitans* was found in this study (n=12). Occurrence of phages in this bacterial species is widespread. *A. actinomycetemcomitans* bacteriophages have been isolated either from healthy (Willi et al. 1997a, Haubek et al. 1997) or periodontally diseased patients (Barretto-Tinoco et al. 1997, Iff et al. 1997, Haubek et al. 1997, Sandmeier et al. 1995, Willi et al. 1993, Olsen et al. 1993, Preus et al. 1987a, Preus et al. 1987b). Whereas several authors suspected that bacteriophages might modulate the virulence of *A. actinomycetemcomitans*, no clear correlation has been established between lysogeny in *A. actinomycetemcomitans* and severity of the disease. *A. actinomycetemcomitans* phages have been classified in Bradley's groups A (Iff et al. 1997, Loftus and Delisle 1995, Sandmeier et al. 1995, Olsen et al. 1993, Preus et al. 1987a), B (Loftus and Delisle 1995, Olsen et al. 1993, Preus et al. 1987a) and C (Preus et al. 1987a). The majority of phages induced from *A. actinomycetemcomitans* belong to the *Myoviridae* family (Sandmeier et al. 1995, Willi et al. 1993). Five bacteriophages of *A. actinomycetemcomitans* isolated from periodontal pockets in three different European

countries, were genetically highly related and morphologically indistinguishable (Willi et al. 1993). Among them, AaΦ23, derived from its natural lysogen *A. actinomycetemcomitans* strain ZIB1023, has been chosen as a representative of this family and has been studied in more detail.

Bacteriophage AaΦ23

The bacteriophage AaΦ23 has first been isolated from *A. actinomycetemcomitans* ZIB1023 (Willi et al. 1993). The virion has an isometric head of about 62 nm and a contractile tail of 112 nm, it belongs to the *Myoviridae* family (Willi et al. 1993). Restriction fragments analysis is in agreement with a physical circular map for the AaΦ23 DNA (Willi and Meyer 1998). Electron microscopy observations revealed that the phage genome comprises about 44 kb and that the DNA molecules contained in phage heads consist of a collection of molecules with a terminal redundancy of 1.6 kb which is circularly permuted (Willi and Meyer 1998). Southern blot experiments showed that the phage DNA integrates, as a prophage, into the host chromosome (Willi and Meyer 1998). The AaΦ23 prophage was found in about 40% of 259 strains derived from European isolates of *A. actinomycemcomitans* (Iff et al. 1997, Haubek et al. 1997, Willi et al. 1997a, Willi et al. 1993). Occurrence of the AaΦ23 prophage was observed in strains isolated from periodontal disease as well as from healthy sites. No correlation between infection with AaΦ23 and severity of the periodontal disease was observed (Willi et al. 1997a). These results suggested that AaΦ23 prophages *per se* do not significantly increase the virulence of its host. Evidence was obtained that phage AaΦ23 is able to transfer antibiotic resistance markers by transduction (Willi et al. 1997b).

AIMS OF THE STUDY

In order to learn more about the biology of this widespread phage, the first goal of the PhD work presented here was the sequencing of the entire double-stranded DNA molecule of bacteriophage AaΦ23. By this approach, genes coding for virulence factor which might have homologies with known virulence determinants would have been quickly identified. Chapter I present the complete genomic sequence of AaΦ23 and its annotation. Additionally, the bacteriophage *att*P and bacterial *att*B sites are reported in this Chapter.

Identification of possible ORFs homologues to proteins that showed previously assigned functions guided our further work.

- In Chapter II, the characterization of an AaΦ23 encoded DNA adenine methyltransferase is reported.
- Chapter III and IV are two preliminary studies of the lytic enzyme and the putative integration elements of AaΦ23, respectively.

CHAPTER I

Complete genomic nucleotide sequence of the temperate bacteriophage AaΦ23 of *Actinobacillus actinomycetemcomitans*

Grégory Resch,[1] Eva M. Kulik,[1] Fred S. Dietrich[2] and Jürg Meyer[1*]

[1]Institute for Preventive Dentistry and Oral Microbiology, University of Basel, Hebelstrasse 3, 4056 Basel, Switzerland.

[2]Department of Applied Microbiology, Pharmacenter, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland.

*Corresponding author. Mailing address: Institute for Preventive Dentistry and Oral Microbiology, Hebelstrasse 3, 4056 Basel, Switzerland. Phone: (41) 61 267 2601. Fax: (41) 61 267 2658. E-mail: juerg.meyer@unibas.ch.

[2]Present address: Duke Center for Genome Technology, Duke University Medical Center, Durham, North Carolina 27710, USA.

ABSTRACT

The entire double-stranded DNA molecule that represents the *Actinobacillus actinomycetemcomitans* bacteriophage AaΦ23 genome was sequenced. DNA contained in the phage particles consists of a collection of linear molecules with a terminal redundancy which are circularly permutated. Therefore, the physical map of the phage genome is circular. Its size is 43,033 bp and it has an overall molar G+C content of 41 mol%. 57 potential open reading frames (ORFs) were identified. A putative function could be assigned to 20 of the 57, i.e. 35% of the ORFs. While 25 another ORFs share homologies with hypothetical proteins present in several bacteria or bacteriophages, 12 seem to be specific for phage AaΦ23. The putative organization of the phage genome shares extensive similarities to that of lambdoid phages. Similarities of AaΦ23 to lambdoid phages are also evident when considering sequence homologies of the lysogeny module (Integrase, C1/Cro regulators), the DNA replication system and the identified DNA adenine methyltransferase. One difference resides in the DNA packaging strategy that is more closely related to the P22 system. The putative lytic system of AaΦ23 seems to correspond to the one described for double stranded DNA bacteriophages of Gram-negative hosts. The attachment sites of AaΦ23 (*att*P) and of several *A. actinomycetemcomitans* strains (*att*B) were determined.

INTRODUCTION

Actinobacillus actinomycetemcomitans is a capnophilic, non-motile Gram-negative bacterium, which has been strongly implicated in the etiology of localized juvenile periodontitis and certain forms of adult periodontitis (Meyer and Fives-Taylor 1997, Zambon 1985). It is thought to play also a role in the establishment of other systemic diseases such as endocarditis (van Winkelhoff and Slots 1999). Although the mechanisms of pathogenicity are still not yet fully understood, several putative virulence factors including cytotoxic and immunosuppressive factors, the lipopolysaccharide, a bacteriocin, a collagenase, a chemotaxis-inhibiting factor and antibiotic-resistance factors have been described (Fives-Taylor et al. 1999). The most extensively studied virulence factor of *A. actinomycetemcomitans* is its leukotoxin. This protein which is capable of lysing human polymorphonuclear leukocytes and macrophages, belongs to the group of repeat in toxin (RTX) family of bacterial cytotoxins (Lally et al. 1989). A large fraction of *A. actinomycetemcomitans* strains isolated from diseased sites were found to produce significantly more leukotoxin than isolates from healthy sites (Zambon 1985). Furthermore, many strains associated with LJP carry a deletion in the promoter region for the leukotoxin operon leading to higher production of the toxin (Contreras et al. 2000).

Lysogeny is widespread in *A. actinomycetemcomitans* (Haubek et al. 1997, Sandmeier et al. 1995). The first inducible bacteriophage, ΦAa17, was isolated from a leukotoxic strain (Stevens et al. 1982). Since then, lysogenic *A. actinomycetemcomitans* have been isolated from periodontal pockets (Iff et al. 1997, Haubek et al. 1997, Sandmeier et al. 1995) as well as from periodontally healthy individuals (Willi et al. 1997). The role of phages in the etiology of periodontal diseases is not yet clear. While some studies indicate a relation between phage-infected *A. actinomycetemcomitans* and rapidly destructive periodontitis

(Preus et al. 1987) others could not find such an association (Willi et al. 1997, Iff et al. 1997, Haubek et al. 1997). In several bacterial species, as *Vibrio cholerae* (Karaolis et al. 1999, Waldor and Mekalanos 1996) and *Escherichia coli* (Plunkett et al. 1999, Neely and Friedman 1998) phage-encoded genes can modulate virulence.

Phages released from several different *A. actinomycetemcomitans* lysogens possess an isometric icosahedral head of 65nm, a contractile tail of 110nm, a baseplate with up to four fibers (Iff et al. 1997, Willi et al. 1993) and therefore belong to the *Myoviridae* family (van Regenmortel et al. 2000). These phages are not only morphologically indistinguishable but also genetically related (Willi et al. 1993). To further our understanding of the biology of *A. actinomycetemcomitans* bacteriophages, we decided to sequence the entire genome of the bacteriophage AaΦ23 that naturally infects the *A. actinomycetemcomitans* strain ZIB1023 (Willi et al 1993). DNA contained in of this temperate phage particles consists of an approximately 45 kb double-stranded linear DNA molecule that is terminally redundant and circularly permuted.

MATERIALS AND METHODS

Bacterial strains, media, and culture conditions. *A. actinomycetemcomitans* ZIB1023 is the natural strain lysogenic for AaΦ23 (Willi et al. 1993). *A. actinomycetemcomitans* ZIB1001 represents a non lysogenic strain (Willi et al. 1993). *A. actinomycetemcomitans* HK1651 is a clinical isolate, the genome of which has been sequenced at the Advanced Center for Genome Technology (University of Oklahoma), obtained from Mogens Kilian (University of Aarhus, Denmark). ZIB1015 is a non-lysogenic strain and ZIB1515 corresponds to the strain ZIB1015 that was experimentally lysogenized with the phage AaΦ23 (Willi et al. 1993). All strains were grown in Todd Hewitt Broth 3% (w/v) (BD

Biosciences, Allschwil, Switzerland) at 37°C in air and 10% CO_2. *Escherichia coli* XL1 Blue electro-competent cells (Stratagene, Zürich, Switzerland) were used as hosts for plasmids pBluescript II SK^-. *E. coli* XL1 Blue were grown in Luria-Bertani medium (LB, Sambrook et al. 1989) containing 100 μg/ml of ampicillin (Sigma, Buchs, Switzerland) at 37°C in air.

Purification, cloning, and sequencing of the AaΦ23 genomic DNA. Phage DNA of AaΦ23 was prepared according to the procedure previously described (Willi et al. 1993). Briefly, after induction of phage release from a log-phase growing culture of *A. actinomycetemcomitans* strain ZIB1023 with mitomycin C (1 μg/ml), supernatant was treated with chloroform 1% (v/v). After addition of DNase (1 μg/ml) and RNase (1 μg/ml), phage particles were precipitated by polyethylene glycol 8000 and recovered by centrifugation. The phage DNA was subsequently extracted by phenol/chloroform and precipitated by ethanol. All chemicals and enzymes were purchased from Sigma.

Phage DNA librairies were constructed by cutting the DNA with *Alu*I, *Cac*8I, *Hae*III or *Rsa*I and shot-gun cloning of the DNA fragments into pBluescript II SK^- digested with *Sma*I. The ligation mixtures were electroporated into *E. coli* XL1 Blue cells. Clones were frozen at –70°C in THB supplemented with 10% (v/v) Hogness Freezing Medium (Werner et al. 1997). Plasmids were recovered with the help of the QIAwell 96 Ultra Plasmid Kit (Qiagen, Basel, Switzerland), following the recommendations of the manufacturer. Sequencing reactions were performed by using the ABI PRISM big-dye ready mix (Perkin Elmer, Langen, Germany) and analysed on an ABI PRISM 377 automated DNA sequencer (Perkin Elmer). Initially, T3 and T7 universal primers were used. The remaining gaps were closed either by primer walking on clones or phage genomic DNA or by sequencing of PCR amplified fragments covering the gaps. All restriction enzymes were obtained from Roche Diagnostics

AG (Rotkreuz, Switzerland) and oligonucleotides were purchased from Microsynth GmbH (Balgach, Switzerland).

Sequence analysis and annotation. Sequences obtained were assembled by using the phred/phrap/consed software package. A total number of 289,652 bp, corresponding to an average coverage of the phage DNA of 6.7, was sequenced. The resulting consensus sequence was analysed by using the following software tools :ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), getorf (http://bioweb.pasteur.fr/seqanal/interfaces/getorf.html), blastn and blastp (http://www.ncbi.nlm.nih.gov/BLAST) against GENBANK and SWISSPROT databases, Webcutter 2.0 (http://www.firstmarket.com/cutter/cut2.html), Compute pI/MW (http://www.expasy.org/tools/pi_tool.html), ClustalW (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_clustalw.html), Interproscan (http://www.ebi.ac.uk/interpro/scan.html), einverted (http://bioweb.pasteur.fr/seqanal/interfaces/einverted.html), tRNAscan-SE 1.21 (http://www.genetics.wustl.edu/eddy/tRNAscan-SE), FAStRNA (http://bioweb.pasteur.fr/seqanal/interfaces/fastrna.html), promoter prediction software (http://www.fruitfly.org/seq_tools/promoter.html) and several others from the GCG Wisconsin package (Genetics Computer Group, Accelrys Inc., Munich, Germany).
Homologies to the genomic sequence of the *A. actinomycetemcomitans* strain HK1651 were determined using the blastn software available on the internet site of the Advanced Center for Genome Technology (Oklahoma University, USA, http://www.genome.ou.edu/act_blast.html).

Determination of the sequences of the attachment sites *att*P, *att*B, *att*L, and *att*R. The *att*B of strains ZIB1001 and 1015 were determined by using a semi-random polymerase chain reaction (SPCR). The protocol for the SPCR was obtained from the Roth Laboratory at the Division of Biological Sciences Section of Microbiology (University of California, Davis, http://rothlab.ucdavis.edu/protocols/semirandom.html). This protocol was adapted from Hermann et al. (2000). Primers used in this study are listed in Table 1. In the first step, the left and right ends of the integrated prophage and the flanking bacterial sequences from *A. actinomycetemcomitans* ZIB1023 and ZIB1515 were determined by SPCR amplification with SPCR primers 1 and 2 and sequencing from SPCR.seq primers 1 and 2. In the second step, SPCR primers 2, 3 and SPCR.seq primers 3 and 4 were designed based on the sequences obtained in the first step, and used to determine the sequence of the *att*B in the non-lysogenic *A. actinomycetemcomitans* strains ZIB1001 and ZIB1015. The *att*P site was determined by comparison of the determined *att*L, *att*R, and *att*B sites with the genomic sequence of AaΦ23. Moreover, the *att* core sequence was compared with the Genbank database and with the *A. actinomycetemcomitans* HK1651 genome.

Table 1. List of the primers used to determine the sequence of the bacteriophage AaΦ23 integration sites.

Primer name	Oligonucleotide sequence of the primer (5'-3')	Hybridization site
SPCR1	TATATGTGCTGCAAGAGTTGGGCGGG	204-229 on AaΦ23
SPCR1.seq	GCGCATTTATCTGTTGAGCAT	153-173 on AaΦ23
SPCR2	CGTTTTATGAGGTTTGCATGTGTGTT	42879-42904 on AaΦ23
SPCR2.seq	TGATGACTACGATAAATTTATAAGCA	42959-42984 on AaΦ23
SPCR3	CAATGGATGTCACCCTGGTCGGCATC	147955-147930 on HK1651
SPCR3.seq	ACGTGCCATTGATTTTATTAAAGCG	147900-147876 on HK1651
SPCR4	CCGTTTAATTTAGACATGTTGTCTTCC	147401-147427 on HK1651
SPCR4.seq	TAACTTACTGAAGTGATTGTCTAG	147556-147579 on HK1651

Southern-blot experiments. Bacterial genomic DNA of *A. actinomycetemcomitans* strains ZIB1001, ZIB1023 and HK1651 was digested either with *Asp*700 or *Eco*RI and the restriction fragments were separated by gel electrophoresis on a 0.7% low-melting agarose gel (Bioconcept, Allschwil, Switzerland). DNA was transferred, by capillary blotting, onto a nylon membrane (Amersham Bioscience, Dübendorf, Switzerland). An Enhanced chemiluminescence method (Amersham Biosciences) was used for probe labeling, hybridization and detection, following the manufacturer protocols. 250 ng of the AaΦ23 genomic DNA served as hybridization probe.

Nucleotide accession number. The nucleotide sequence reported in this paper has been assigned accession number AJ560763 in the EMBL database.

RESULTS AND DISCUSSION

Nucleotide sequence of the AaΦ23 genome. The final consensus sequence obtained with phred/phrap/consed was circular. For convenience, the phage genome was linearized at the first nucleotide of the *att*P core sequence (Figure 1). The sequence of the AaΦ23 genome is composed of 43,033 bp. Virtual restriction maps of this sequence are in good agreement with the restriction maps previously obtained (Willi and Meyer 1998). These observations support the correct assembly of the sequences. The average G+C content of AaΦ23 is 41 mol%. This value is close to the 42.7 mol% reported for the *A. actinomycetemcomitans* genome (Kaplan and Fines 1998). This observation may suggest that phage AaΦ23 co-evolved since a long time with its host.

Gene prediction and organization of the AaΦ23 genome. Analysis of the AaΦ23 genome sequence predicted 56 genes with a minimal length of 300 nucleotides (Table 2). Although *orf14* codes for a protein composed of only 68 amino acids, it was added to the genomic map because it may code for the Cro-like protein of AaΦ23 (Table 2). The genome is very condensed in terms of coding sequences as 93.7% of the entire genome are covered by potential coding regions and 19 gene are overlapping (Figure 1). Putative promoter sequences within 500 bp upstream of 49 of the 57 ORFs were identified (data not shown). No tRNA genes are present on the genome. Fifteen stem-loop like structures, that may represent rho-independent transcription termination signals, have been identified (data not shown). The genome of AaΦ23 appears to be divided in two different regions (Figure 1). Most of the 13 potential genes located in the left part of the genome (nucleotide 1 to 10121) are transcribed leftwards on the genetic map. In the right part of the genome (nucleotide 10122 to 43033) 39 of the 44 potential genes are transcribed rightwards on the genetic map.

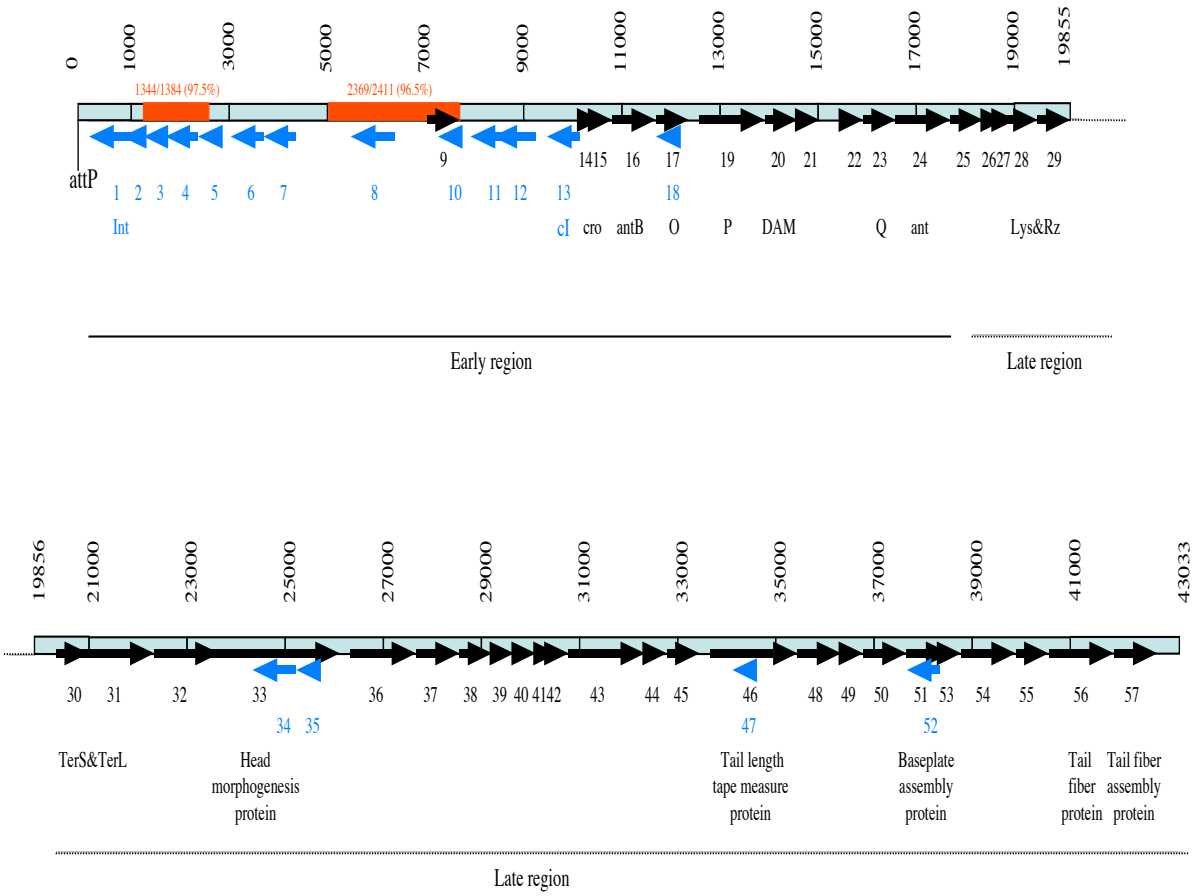

Figure 1. Schematic representation of *Aa*Φ23 genome. *orfs* are numbered consecutively from left to right and indicated by arrows pointing in the direction of transcription.

Black arrows represent genes transcribed rightwards and blue arrows represent genes transcribed leftwards on the map. Putative functions are indicated.

Red boxes indicate regions of homologies to the genome of the *A. actinomycetemcomitans* strain HK1651.

Table 2. Description of bacteriophage AaΦ23 ORFs, gene product and functional assignments.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf1	1115	69	348	40.8	10.08	Integrase (*int*)	Integrase (phage ST64T, AAL15478.1, 3e-73) Integrase (phage P22, NP_059584.1, 5e-72) Putative phage integrase (*Shigella flexneri* 2a strain 301, AAN41973.1,5e-72) Integrase (phage SfV, AAB72135.1, 2e-71) Integrase (phage DLP12, AAC73638.1, 3e-70) Integrase (phage SfX, AAD10295.1, 6e-69) P38 (phage APSE-1, NP_050999.1, 2e-66) Hypothetical protein (*Xylella fastidiosa* strain Dixon,ZP_00038220.1,4e-63) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00042120.1, 4e-54)
orf2	1336	989	115	13.2	9.38		
orf3	1666	1352	104	12.5	7.67		Hypothetical protein (*Actinobacillus actinomycetemcomitans* strain HK1651,2e-55)
orf4	2377	1721	218	12.5	6.72		Hypothetical protein (*A. actinomycetemcomitans* strain HK1651, e-113)
orf5	2929	2444	161	33.2	5.65		Hypothetical protein (*A. actinomycetemcomitans* strain HK1651, 6e-48)
orf6	3760	3137	207	24.2	4.77		
orf7	4650	3757	297	18.5	9.11		
orf8	6226	5570	218	19.1	9.51		Hypothetical protein (*Haemophilus somnus* strain 129PT, ZP_00123029.1,e-45) ORF14c (phage O395,AAN74016.1, 2e-25) P43 (phage APSE-1, NP_051004.1, e-22) Probable phage protein YPO2093 (*Yersinia pestis* strain CO92,AF0255,4e-20) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00038185.1, 2e-18) Phage-related protein (X.. *fastidiosa* 9a5c, AAF84454.1, 5e-17)
orf9	7138	7662	174	11.9	9.69		Hypothetical protein (*A. actinomycetemcomitans* strain HK651, 6e-78)
orf10	7565	7125	146	13.2	9.38		Hypothetical protein (*A. actinomycetemcomitans* strain HK1651, e-68)
orf11	8491	8018	157	18.3	8.95		
orf12	9347	8460	295	34	6.47		

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, e-value)
orf13	10121	9405	250	28	5.01	Repressor (*c1*)	CI (phage phi-80, S04828, 7e-44) Probable transcription regulator PA0906 (*Pseudomonas aeruginosa* strain PAO1, E83531, 6e-43) Pectin lyase regulation DNA-binding protein (*Pectobacterium carotovorum* strain 71, AAA24865.1,3e-39) Regulatory protein (*Dichelobacter nodosus* strain A198, CAA67160.1,3e-39) Transcriptional regulator, Cro/CI family (phage lambdaSo, AAN56003.1,5e-39) Pyocin regulatory protein prtR (*P. aeruginosa* strain PA01, A47062, 2e-38)
orf14	10231	10429	68	7.2	9.8	Repressor (*cro*)	Hypothetical protein ydaS (*Escherichia coli* strain K12, H64885, 2e-06) Unknown protein (phage CP-933N, G85676, 3e-06) Unknown protein (*Vibrio cholerae*, AAL59694.1, 4e-05) Hypothetical protein XF0501 (*X. fastidiosa* strain 9a5c, AAF83311.1, e-04) Cro protein (phage 21, CAB39982.1, 5.2) Cro protein (phage P22 , AAA32268.1, 6.7)
orf15	10410	10736	108	12.4	5.83		Hypothetical protein (*A. actinomycetemcomitans* strain HK1651, 3e-08)
orf16	10788	11507	239	27.2	9.28	Antirepressor (*antB*)	Hypothetical protein (*Azotobacter vinelandii* strain , ZP_00089317.1, 2e-24) Hypothetical protein HI1422 (*Haemophilus influenzae* Rd strain KW20, E64029, 6e-24) Hypothetical protein (phage VT2-Sa, NP_050579.1, e-22) AntB (*E. coli* strain O157:H7, AAM53132.1, e-22) Putative antirepressor promoter (phage A118, CAB53832.1, 4e-22) Hypothetical protein (phage BP-933W, AAD25485.1, 5e-22) Putative antirepressor (phage ul36, AAM75755.1, 9e-22)
orf17	11662	12444	260	30	7.08	DNA replication protein (gene *o*)	Protein GP18 (phage P22, Z8BPC2, e-18) Replication of DNA (phage Nil2, CAC95088.1, 4e-17) Gp54 (phage HK97, AAF31132.1, e-16)
orf18	12049	11714	111	13.3	9.85		

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf19	12444	13901	485	54.7	5.9	DNA Helicase (gene *p*)	P protein (*E. coli* strain O157:H7, BAA94127.1, e-123) P protein (phage HK022, NP_050526.1, e-123) Replication protein P (phage HK620, NP_112058.1, e-123) Gp12 (phage ST64T, NP_720303.1, e-121) Replicative DNA helicase (*Yersinia pestis* strain KIM, AAM84165.1,4e-72)
Orf20	13904	14398	163	18.9	6.96	DNA adenine methylase (*dam*)	Hypothetical protein (*H. somnus* strain 129PT, ZP_00123655.1, 4e-20) dam (phage HP2, AAK37795.1, 5e-20) Putative adenine-specific methylase (phage HP1, P51715, 5e-20) Putative DNA methyltransferase (phage CP-933V, NP_288678.1, 3e-15) Hypothetical protein (phage Lahn3, CAC95062.1, 4e-15) Hypothetical protein (phage Nil2, CAC95095.1, 4e-15) Gp62 (phage HK97, AAF31137.1, 4e-15) Putative DNA methylase (*E. coli* strain O157:H7, NP_311008.1, 4e-15) DNA methylase (phage VT2-Sa, NP_050531.1, 5e-15)
orf21	14405	14902	165	18.9	9.19	Recombination Protein (*ninB*)	Unknown protein (phage HK620, NP_112060.1, 5e-30) Probable phage protein NinB (*Y. pestis* CO92, AB0255, e-17) Protein NinB (phage HK022, NP_597901.1, 2e-14) NinB protein (phage 21, CAB39988.1, 5e-14) Gp61 (phage HK97, NP_037743.1, 5e-14) Protein NinB (phage lambda, P03765, 9e-14) Hypothetical protein (phage Lahn3, CAC65061.1, 2e-13) NinB (phage P22, AAF75030.1, 3e-13)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf22	15495	16064	189	21.7	9.9	Recombination Endonuclease (*ninG*)	Hypothetical protein of (*H. somnus* strain 129PT, ZP_00122672.1, 6e-32) Recombination endonuclease (phage Nil2,CAC95102.1, 4e-25) Unknown protein (phage CP-933V, AAG57230.1, 3e-24) Protein NinG (phage lambda, P03770, 3e-24) Hypothetical protein (phage I, BAB87988.1, 6e-24) NinG (phage BP-933W, Q9ZWX1, 6e-24) NinG (phage 21, CAB39991.1, e-23) Unknown protein (phage LC159, AAN59917.1, 2e-23) Unknown protein (phage CP-933K, A85581, 2e-23) NinG (phage P22, Caa55163.1, 8e-22)
orf23	16064	16435	123	14	9.61	Anti-termination Protein (gene *q*)	Probable phage anti-termination protein Q (*Y. pestis* strain KIM, CAC90068.1, 7e-02)
orf24	16725	17621	298	34	6.41	Anti-repressor (*ant*)	Putative anti-repressor protein (phage CP-933N, NP_287306.1, 5e-70) Putative anti-repressor protein (*E. coli* strain O157:H7, BAB34980.1, 5e-70) Putative phage anti-repressor (*Y. pestis* strain KIM, NP_669501.1, 3e-36) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041058.1, 4e-36) Gp30 (phage N15, NP_046925.1, 5e-27) Anti-repressor protein Ant (phage P22, NP_059643.1, 2e-25) Putative anti-repressor (*Streptococcus pyogenes* strain MGAS8232, AAL97393.1, 2e-17) Anti-repressor (phage A118, CAC95313.1, 2e-17) Hypothetical protein (*H. somnus* strain 129PT, ZP_00123653.1, 3e-17)
orf25	17785	18231	148	16.3	5.52		
orf26	18228	18596	122	14	8.64		
orf27	18505	18870	121	13.1	9.1		Hypothetical protein HI1413 (*H. influenzae* strain Rd KW20, P44185, e-03)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf28	18939	19520	193	21.6	8.95	Lytic enzyme (*lys*)	Hypothetical protein HI1415 (*H. influenzae* strain Rd KW20, H64028, 2e-56) Lytic enzyme (*P. aeruginosa* strain PA01, BAA83137.1, 2e-33) Lytic enzyme, putative glycohydrolase (phage ST64B, NP_700425.1, 2e-28) Probable bacteriophage protein STY1042 (*Salmonella enterica* serovar typhimurium strain CT18, AI0620, 3e-26) Putative chitinase (phage Fels-1, NP_459884.1, 2e-25) Putative glycohydrolase (*Deinococcus radiodurans* strain R1, NP_285440.1, 6e-23) Hypothetical protein (*Rhodopseudomonas palustris*,ZP_00011491.1, 5e-22) Pyocin R2_PP, lytic enzyme (*P. putida* strain KT2440, NP_745210.1, 9e-21) Phage related lytic enzyme (*Xanthomonas axonopodis* pv. citri strain 306, NP_642961.1, 6e-20)
orf29	19523	19855	110	12.6	9.41	Rz lytic protein (*rz*)	Hypothetical protein HI1414 (*H. influenzae* strain Rd KW20, P44186, e-15) Putative Rz lytic protein (phage SfV, AAL89453.1, e-3)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf30	20475	21035	186	21.1	5.21	Terminase small subunit (*terS*)	Hypothetical protein (*Novosphingobium aromaticivorans,* ZP_00096135.1, 2e-21) Putative terminase small subunit (phage 315.6, AAM80043.1,3e-16) Subunit 1 of the terminase enzyme (phage rho15, CAA47712.1,6e-13) Subunit 1 of the terminase enzyme (phage SF6, CAA47711.1,6e-13) Hypothetical protein HI1411 (*H. influenzae* strain Rd KW20, Q57374, 5e-12) Terminase small subunit (*H. influenzae* strain Rd KW20, AAC23059.1, 5e-12) Terminase (*Staphylococcus aureus* strain Mu50, BAB56962.1, 8e-11) Terminase (phage bIL310, AAK08430.1, 7e-09) Probable terminase (phage phi-gle, T13203, 2e-08) Terminase small subunit (phage SPP1, NP_690652.1, 4e-08) Putative terminase small subunit (phage A118, CAB53787.1,4e-08)
orf31	20980	22443	487	54.7	8.48	Terminase large subunit (*terL*)	Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00042224.1,e-65) Phage-related protein (*X. fastidiosa* strain Temecula1, AAO28853.1,e-65) Hypothetical protein (*X. fastidiosa* strain Dixon , ZP_00039342.1,2e-65) Hypothetical protein (*N. aromaticivorans*, ZP_00094970.1, e-56) Terminase large subunit (phage Bcep781, AAN38019.1, e-50) Phage-related protein (*Listeria innocua* strain Clip11262, CAC96963.1, 5e-50) Phage-related protein XF1675 XF1570 (*X.fastidiosa* 9a5c, A82650,3e-42) Hypothetical protein (*X. fastidiosa* Ann-1, ZP_00042155.1, e-38) Hypothetical protein BH3533 (*Bacillus halodurans* strain C-125, E84091, 6e-32) Putative terminase large subunit (phage ul36, NP_663671.1, 8e-32)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf32	22445	23758	437	49.1	4.71		Hypothetical protein HI1409 (*H. influenzae* strain Rd KW20, P44183, 0.0) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039532.1,2e-48) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00042225.1,5e-48) Hypothetical protein XF1571 XF1676 (prophages XfP3 and XfP4, A82664, 5e-48) Hypothetical protein (*S. flexneri* 2a strain 301, AAN42305.1, e-43) Probable bacteriophage protein STY1048 (*S. enterica* serovar typhimurium strain CT18, AG0621, e-43) Phage-related protein (*X. fastidiosa* strain Temecula1, AAO28846.1, 3e-41) Hypothetical protein (*Mesorhizobium loti*, BAB53655.1, e-38) Hypothetical protein Lin1731 (*L. innocua* strain Clip11262, AB1649, 4e-20)
orf33	23679	26081	800	89.6	9.49	Head protein	TraN-related protein HI1407 (*H. influenzae* strain Rd KW20, B64122, e-109) Plasmid-related protein (*X. fastidiosa* strain Temecula1, AAO28843.1, 2e-34) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041695.1,4e-34) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039531.1,7e-33) Plasmid-related protein XF1574 XF1679 (prophages XfP3 and XfP4, E82650, 4e-32) Minor head protein (*Enterococcus faecalis* strain v583, AAO81248.1, e-15)
orf34	25361	24462	299	34.3	9.88		Hypothetical protein MJ0314 (*Methanococcus jannaschii,* NP_247287.1, 2e-06)
orf35	25723	25391	110	11.8	4.35		

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf36	26713	27825	370	40.3	4.97		Hypothetical protein HI1405 (*H. influenzae* strain Rd KW20, AAC23055.1, e-171) Hypothetical protein (*M. loti*, NP_108196.1, 8e-23) Hypothetical protein XF1575 XF1680 (prophages XfP3 and XfP4, F82650, 3e-22) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039530.1,5e-22) Phage-related protein (*X. fastidiosa* strain Temecula1, AAO28842.1, e-21) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041694.1,e-21) Hypothetical protein (*N. aromaticivorans*, ZP_00096131.1, 6e-21) Probable bacteriophage protein STY1050 (*S. enterica* serovar typhimurium strain CT18, AI0621 , 2e-18) Putative bacteriophage protein (*S. flexneri* 2a strain 301, AAN42303.1, 6e-18)
orf37	27839	28288	149	15.3	5.39		
orf38	28300	29214	304	33.2	7.03		Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041692.1,4e-11) Hypothetical protein XF1577 XF1682 (prophages XfP3 and XfP4, H82650, 7e-11) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00038226.1,e-10) Phage-related protein (*X. fastidiosa* strain Temecula1, AAO28840.1, e-10)
orf39	29225	29584	119	12.9	5.32		Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00038228.1,e-15) Hypothetical protein (*X. fastidiosa* strain Temecula1, AAO28838.1,e-15) Hypothetical protein XF1579 XF1684 (prophages XfP3 and XfP4, A82665, 2e-14) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041690,4e-14)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf40	29585	30031	148	16.6	9.49		Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041689.1,5e-13) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039526.1,e-11) Phage-related protein (*X. fastidiosa* strain Temecula1, NP_779188.1, 7e-11) Hypothetical protein XF1580 XF1685 (prophages XfP3 and XfP4, C82651, 7e-11)
orf41	30028	30399	123	13.9	9.43		Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041688.1,e-19) Hypothetical protein (*X. fastidiosa* strain Temecula1,AAO28836.1,8e-19) Hypothetical protein XF1581 XF1686 (prophages XfP3 and XfP4, Q9P9T2, e-17) Hypothetical protein (*X. fastidiosa* strain Dixon,ZP_00039525.1,4e-15)
orf42	30185	30826	213	23.8	9.04		Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041687.1,e-18) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039524.1,e-16) Hypothetical protein (*X. fastidiosa* strain Temecula1,AAO28835.1,7e-12) Hypothetical protein XF1582 XF1687 (prophages XfP3 and XfP4, E82651, 9e-12)
orf43	30831	32339	502	54.5	5.72		Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039523.1,e-127) Phage-related protein (*X. fastidiosa* strain Temecula1, AAO28834.1, e-126) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041686.1,e-125) Hypothetical protein XF1583 XF1688 (prophages XfP3 and XfP4, F82651, e-123) Putative phage protein STY2029 (*S. enterica* serovar typhimurium strain CT18,AI0622,4e-42) Hypothetical protein (*Magnetospirillum magnetotacticum*, ZP_00049744.1, 3e-36)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf44	32389	32820	143	15.7	5.58		Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041684.1,2e-46) Hypothetical protein XF1584 XF1689 (prophages XfP3 and XfP4, F82665, 9e-46) Phage-related protein (*X. fastidiosa* strain Temecula1, AAO29040.1, e-45) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041331.1,2e-42) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039522.1,7e-41)
orf45	32820	333239	139	15.3	6.17		Hypothetical protein (*X. fastidiosa* strain Temecula1,NP_779390.1,2e-20) Hypothetical protein XF1585 XF1690 (prophages XfP3 and XfP4, H82651, 5e-20) Hypothetical protein (*X. fastidiosa* strain Ann-1,ZP_00041683.1,2e-19)
orf46	33390	35498	702	76.6	4.67	Tail length tape measure protein	Hypothetical protein (*X. fastidiosa* strain Dixon,ZP_00039519.1,2e-51) Hypothetical protein (*X. fastidiosa* strain Ann-1,ZP_00041329.1,3e-51) Phage-related protein (*X. fastidiosa* strain Temecula1, AAO28831.1, 4e-51) Hypothetical protein XF1587 XF1692 (prophages XfP3 and XfP4, B82652, e-49) Probable phage protein STY1061 (*S. enterica* serovar typhimurium strain CT18, AD0623, 2e-11) Phage protein similar to tail length tape measure protein (*Clostridium tetani* strain E88,NP_781744.1,8e-09)
orf47	34735	34133	200	21.5	8.67		
orf48	35502	36269	255	27.2	9.21		Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00038456.1,2e-38) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041679.1,6e-38) Hypothetical protein XF1591 (prophage XfP3, B82661, e-37) Phage-related protein (*X. fastidiosa* strain Temecula1, NP_779179.1, 4e-37) Probable phage protein STY1062 (*S. enterica* serovar typhimurium strain CT18, AE0623,2e-06)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf49	36280	36594	104	11.9	4.97		Hypothetical protein XF1628 (*X. fastidiosa* strain 9a5c, C82648,2e-20) Hypothetical protein (*X. fastidiosa* strain Temecula1, NP_779178.1,2e-18) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00038457.1,8e-10)
orf50	36881	37735	284	31.7	9.29		Hypothetical protein XF1593 (prophage XfP3, D82661, 4e-78) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039058.1,6e-77) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00042047.1,e-76) Phage-related protein (*X. fastidiosa* strain Temecula1, NP_779384.1,e-76) Hypothetical protein (*M. magnetotacticum*, ZP_00051100.1, e-20) Probable phage protein STY1064 (*S. enterica* serovar typhimurium strain CT18,AG0623,3e-16)
orf51	37732	38376	214	22.5	5.38	Baseplate protein	Hypothetical protein XF1594 (prophage XfP3, E82661, 7e-33) Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041677.1,2e-32) Phage-related protein (*X. fastidiosa* strain Temecula1,NP_779176.1,2e-32) Probable phage protein STY2019 (*S. enterica* serovar typhimurium strain CT18, AH0732,4e-32) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039057.1,5e-31) Hypothetical protein (*M. magnetotacticum*, ZP_00051101.1, 6e-17) Orf16 similar to V gene of P2 : baseplate (phage phi-CTX, BAA36243.1, 3e-06)
orf52	38195	37776	139	14.5	9.47		
orf53	38373	38732	119	14	5.09		Hypothetical protein (*X. fastidiosa* strain Ann-1, ZP_00041676.1,5e-27) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00038460.1,e-26) phage-related protein (*X. fastidiosa* strain Temecula1,AAO28824.1,7e-26) Hypothetical protein XF1701 (prophage XfP4, F82648, 9e-26) Probable phage protein STY2019 (*S. enterica* serovar typhimurium strain CT18, AD0624,4e-12)

* Isoelectric point

Table 2. Continued.

Gene	Start position	Stop position	size (aa)	Molecular mass (kDa)	pI*	Potential function (gene name)	Significant matches (source, accession number, blastp e-value)
orf54	38763	39905	380	42.4	5.03		Hypothetical protein (*X. fastidiosa* strain Ann-1,ZP_00041674.1,2e-97) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00038461.1,6e-97) Phage-related protein (*X. fastidiosa* strain Temecula1,AAO29028.1,2e-96) Hypothetical protein XF1704 (prophage XfP4, A82649, 5e-96) Probable phage protein STY2016 (*S. enterica* serovar typhimurium strain CT18, AE0732,7e-35) Hypothetical protein (*M. magnetotacticum*, ZP_00050433.1, e-17) Hypothetical protein homolog Lin1710 (*L. innocua* strain Clip11262, AE1646, e-17) Hypothetical protein (*Leptospira interrogans* serovar lai strain 56601, NP_711243.1, 2e-17) Hypothetical protein (phage HF2, NP_542578.1, 4e-08) Hypothetical protein (phage Bcep781, NP_705659.1, 6e-08)
orf55	39905	40486	193	22.1	9.59		Hypothetical protein XF1705 (prophage XfP4, B82649, e-32) Hypothetical protein (*X. fastidiosa* strain Dixon, ZP_00039052.1,2e-32) Phage-related protein (*X. fastidiosa* strain Temecula1, NP_779171.1,7e-32) Hypothetical protein (*X. fastidiosa* strain Ann-1, NP_00041673.1,e-31) Probable phage protein STY1071 (*S. enterica* serovar typhimurium CT18, AF0624, 3e-12) Hypothetical protein HI1404 (*H. influenzae* strain RdKW20,P44179,5e-08)
orf56	40497	41981	494	55.8	6.08	Tail fiber protein	Hypothetical protein HI1403 (*H. influenzae* strain RdKW20,P44178,e-07) Similar to P2 tail fiber protein H (phage 186, AAC34164.1, 2e-04)
orf57	41972	42592	206	23.5	5.44	Tail fiber assembly protein	Hypothetical protein NMA1328 (*Neisseria meningitidis strain* Z2491, H81901, 4e-17) Hypothetical protein NMB1119 (*N. meningitidis* strain MC58, A81120, e-16) Hypothetical protein (*S. flexneri* 2a strain 301, AAN41969.1, 5e-05) Tail fibre assembly protein P37 (phage SfV, AAL89429.1, 2e-04)

* Isoelectric point

Similarities of the AaΦ23 potential ORFs to other sequences present in databases and functional assignments. The 57 potential genes and the corresponding protein sequences were compared to the Genbank and Swissprot databases. Characteristics of the gene products and the significant homologies are detailed below and are also listed in Table 2. Twelve ORFs (2, 6, 7, 11, 12, 18, 25, 26, 35, 37, 47 and 52) may be specific for AaΦ23 as no homology to other sequences could be found (Table 2). When analysing putative gene functions, an "early region" and a "late region" can be defined on the AaΦ23 genome (Figure 1). Genes located in the early region (from nucleotide 1 to 17621) code for proteins involved in recombination, replication, and modification of the phage DNA. The region between nucleotide 18939 and 43033 represent the "late genes" region coding for the lytic system, DNA packaging and structural proteins. The same genetic organization has also been observed in other bacteriophages such as lambda, P22, L and LP-7 (Campbell 1994).

Early region

AaΦ23 integrase. It is likely that ORF1 may represent the integrase of phage AaΦ23. In fact it shows significant homologies to other phage integrase sequences (Table 2). The alignment of the AaΦ23 ORF1 with the integrase amino-acid sequences of these bacteriophages illustrates the observed homologies (Figure 2). Additionally, the Interproscan software identified a domain conserved among phage integrases from amino acid 145 to 324 (PFAM accession number PF00589, e-value of $1.7e^{-38}$). Thus, this enzyme may be a new member of the large family of integrase/recombinases (Nunes-Düby et al. 1998) which catalyses the site-specific integration and excision of the AaΦ23 genome into and out of the host chromosome. As for phage lambda, the putative AaΦ23 integrase gene is located in a cluster of genes transcribed leftwards on the genetic map, i.e. in the opposite orientation to the majority of the phage genes (Figure 1). No gene coding for an excisionase has been identified on the

genome. This situation is also encountered in *Streptococcus pneumoniae* phage MM1 (Obregon et al. 2003), *Listeria monocytogenes* phage A118 (Loessner et al. 2000) and *Pseudomonas aeruginosa* phage D3 (Kropinski 2000).

Identification of the *att*P and *att*B integration sites of AaΦ23. Results of the SPCRs are shown on Figure 3. By sequencing of the SPCRs products, a 49 bp-long direct repeat was identified at the junction of prophage and bacterial DNA in the strain ZIB1023. Moreover, this repeat was found in the chromosome of the non lysogenic strains ZIB1001 and ZIB1015 and was also evident as a single copy on the AaΦ23 DNA sequence. When strain ZIB1001 was lysogenized to result in ZIB1515, the prophage was flanked by direct copies of the repeat, this representing the *att* site (Figure 4). Bio-informatics analysis revealed also a single copy of this repeat in the genome of HK1651 and *Haemophilus influenzae* Rd strain KW20. In AaΦ23 the *att* site is located 19 bp downstream from the putative integrase gene. The location of *att*P close and upstream to the integrase gene seems to be usual (Shimizu-Kadota et al. 2000, Auvray et al. 1999, van Mellaert et al. 1998). Similarly, the size of 49 bp that is rather long for an *att* site is not unusual (Shimizu-Kadota et al. 2000, Auvray et al. 1999, van Mellaert et al. 1998). We showed that the *att*B homologous sequence found in the genome of *H. influenzae* strain Rd KW20 is located at 111 nucleotides downstream from the ORF HI0274 coding for the glutamyl-tRNA synthetase (data not shown, Fleishmann and Venter 1995). Identically, in the genome of HK1651, this sequence is located 109 nucleotides downstream from an ORF that present homologies with the ORF HI0274 of *H. influenzae* strain Rd KW20. Location of *att*B sequences at the 3' end or even overlapping a tRNA gene has been frequently described (Auvray et al. 1999, McShan et al. 1997, Hauser and Scocca 1992).

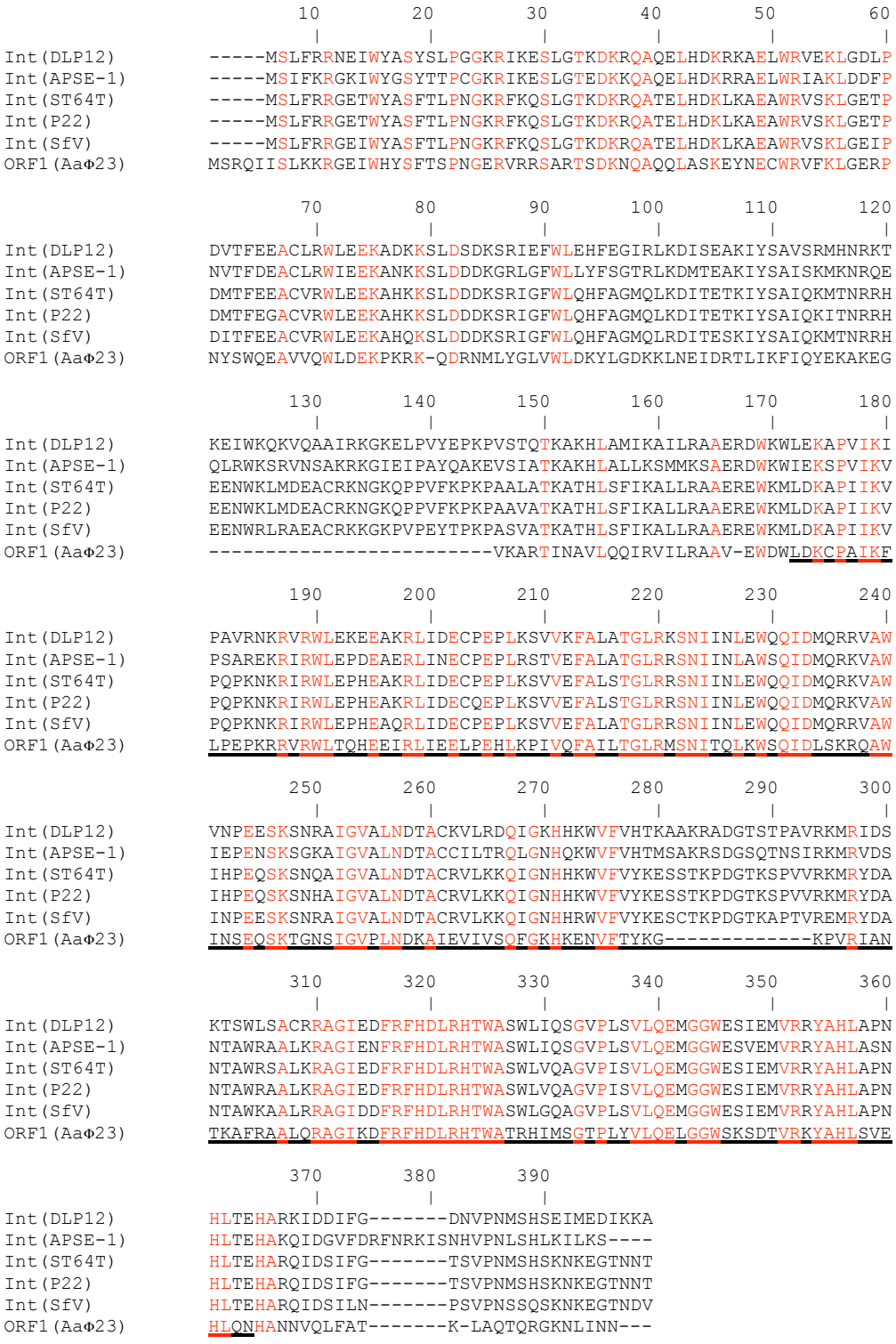

Figure 2. Alignment of AaΦ23 ORF1 with the amino acid sequences of several other bacteriophages Int. Identical residues are in red. The domain matching to the protein family PF00589 is underlined. Int accession numbers : AAC73638 (DLP12); NP_050999 (APSE-1); AAL15478 (ST64T); NP_059584 (P22); AAB72135 (SfV).

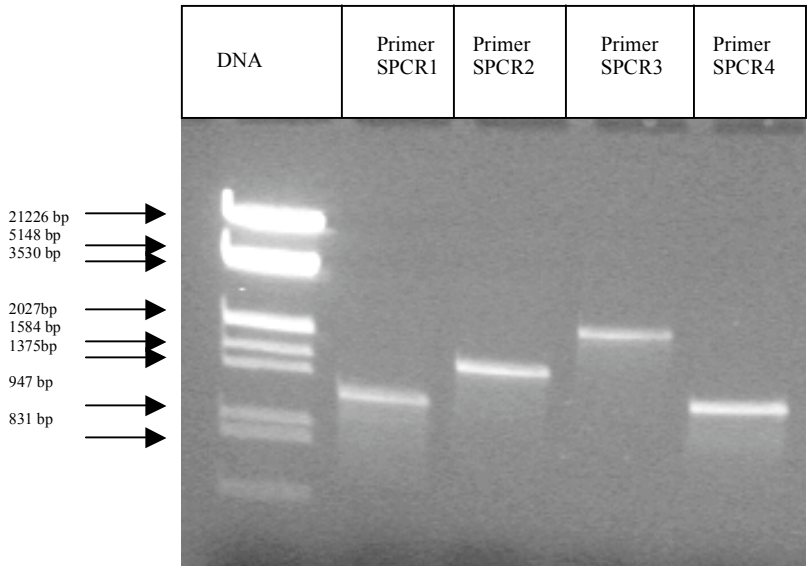

Figure 3. PCR products of the semi-random PCRs performed with the primers SPCR1, 2, 3, and 4, respectively.

>*att*L of *A. actinomycetemcomitans* strain ZIB1023
TAGCTCAGTTGGGAGAGCGCTTGA*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*TTATAGCTTTATCAATTAGTTAGTTATTGATA
* N N I

>*att*R of *A. actinomycetemcomitans* strain ZIB1023
TGATTTATGTATAGGCTTTCAAAT*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*CCAAAATCCTATCAATAAAAATTGATAGGATT

>*att*P of bacteriophage AaΦ23
TGATTTATGTATAGGCTTTCAAAT*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*TTATAGCTTTATCAATTAGTTAGTTATTGATA
* N N I

>*att*B of *A. actinomycetemcomitans* strain HK1651
TAGCTCAGTTGGGAGAGCGCTTGA*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*CCAAAATCCTATCAATAAAAATTGATAGGATT

>*att*L of *A. actinomycetemcomitans* strain ZIB1515
TAGCTCAGTTGGGAGAGCGCTTGA*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*TTATAGCTTTATCAATTAGTTAGTTATTGATA
* N N I

>*att*R of *A. actinomycetemcomitans* strain ZIB1515
TGATTTATGTATAGGCTTTCAAAT*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*CCAAAATCCTATCAATAAAAATTGATAGGATT

>*att*B of *A. actinomycetemcomitans* strain ZIB1015
TAGCTCAGTTGGGAGAGCGCTTGA*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*CCAAAATCCTATCAATAAAAATTGATAGGATT

>*att*B of *A. actinomycetemcomitans* strain ZIB1001
TAGCTCAGTTGGGAGAGCGCTTGA*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*CCAAAATCCTATCAATAAAAATTGATAGGATT

>*att*B of *H. influenzae* strain Rd KW20
TAGCTCAGTTGGGAGAGCGCTTGA*ATGGCATTCAAGAGGTCGTCGGTTCGATCCCGATTATCTCCACCAAATT*AATAAAGCAAGGTAAAGCACCTTGCTGTTTTA

Figure 4. *att*P, *att*L, *att*R, and *att*B sites on the bacteriophage AaΦ23 genome and on several bacterial genomes. The *att* core sequence is boxed. The C-terminal end of the putative integrase is underlined. * represent the stop codon of the integrase gene.

Region between genes *int* and *c1*. The region located between nucleotide 1115 and 9405 contains eleven putative ORFs, 10 of them being transcribed in the same orientation as the putative integrase gene (Figure 1). Five ORFs (2, 6, 7, 11 and 12) seem to be specific for AaΦ23, as they do not share any homologies with other sequences present in databases (Table 2). Five ORFs (3, 4, 5, 9 and 10) have homologies with hypothetical proteins of *A. actinomycetemcomitans* strain HK1651 while ORF8 may have a homologue in several other bacterial species (Table 2).

Gene expression regulation proteins, lysogeny control. Five ORFs coding for putative proteins involved in regulatory functions have been identified on the AaΦ23 genome. The potential ORF13 present strong homologies to the repressor protein CI of the *Siphoviridae* lambda-like phage phi-80, which infects *E. coli* (Eguchi et al. 1988) and several other probable transcriptional regulators found in different bacterial species (Table 2). Moreover, this protein contains a helix-turn-helix motif from amino acid 12 to 66 (PFAM accession number PF013081, e-value of $1.1e^{-08}$) which is a signature of DNA binding proteins. In bacteriophage lambda this repressor protein is responsible for maintaining the prophage in the lysogenic state (Ptashne 1992). The product of *orf14* is a protein of 68 amino acids that shares significant homologies with unknown proteins (Table 2) and weak homologies to the transcription regulatory protein Cro of several bacteriophages and bacteria (Table 2). Identification of potential cI and Cro homologues on the AaΦ23 genome may be an indication that AaΦ23 uses the same lysogeny control system as phage lambda. In bacteria lysogenic for lambda the expression of lytic genes is repressed by binding of cI to the O_L and O_R operators (Ptashne 1992). O_R is located between the *cI* and *cro* genes and is composed of three oligonuleotide sites of two-fold symmetry separated by short spacers. As for the *Shigella flexneri* phage SfV (Allison et al. 2002), three regions containing inverted repeats

are evident in the *cI* and *cro* intergenic region of AaΦ23 (Figure 5). These sequences are named IR1, IR2 and IR3 and may be analogues of the three sequences (O_{R1}, O_{R2}, O_{R3}) that compose the O_R operator of phage lambda. Bio-informatics analysis revealed two putative promoter sequences in this region (Figure 5). The first, from nucleotide 10054 to 10104 may correspond to the promoter for repressor maintenance (P_{RM}) and the second, from nucleotide 10099 to 10147 could represent the promoter $P_{R'}$, both of them being described in lambda (Ptashne 1992). *Orf24* codes for a putative anti-repressor belonging to the protein family of Ant which is expressed in several phages and has been extensively characterized in P22 (Sauer et al. 1983). Ant could play a role in the regulation of lysogeny. It prevents the prophage P22 cII repressor protein from binding to its operators and also inhibits the action of other prophage repressor proteins (Sauer et al. 1983). The deduced gene product of *orf16* may be a second anti-repressor protein similar to the one named antB in *E. coli* O157:H7 (Table 2). It shows significant sequence similarities to the gene HI1422 of *H. influenzae* strain Rd KW20 and to putative anti-repressors from different phages (Table 2). The gene product of *orf23* shares significant homologies with a probable anti-termination protein Q of *Yersinia pestis* strain KIM (Table 2). In bacteriophage lambda, this rho-dependent termination protein allows RNA-polymerase to over-ride the termination signal downstream of promoter $P_{R'}$ so that the lytic and structural genes can be transcribed (Petrov et al. 1981). Finally, no homologies to other regulatory proteins like cII and cIII repressors and anti-termination protein N have been identified on the AaΦ23 genome.

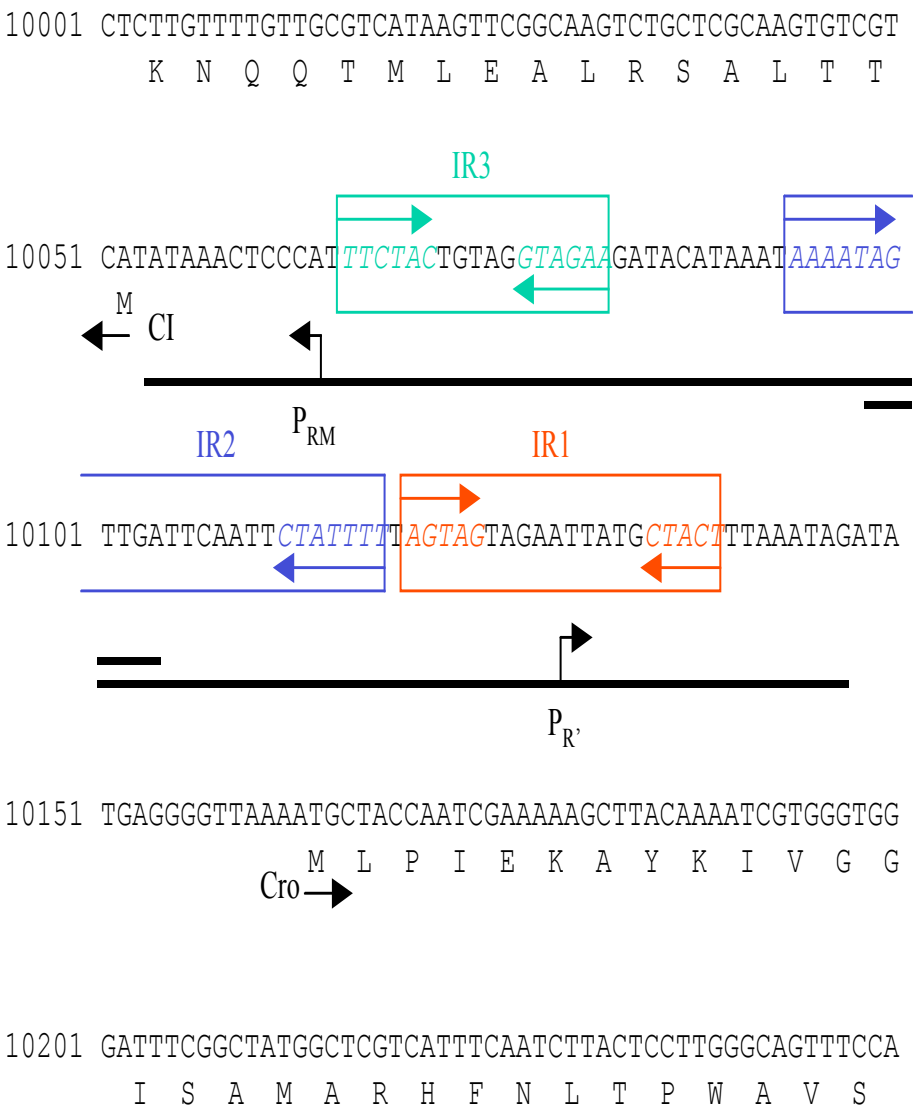

Figure 5. Inverted repeats IR1, IR2 and IR3 in the *cI* and *cro* intergenic region. Regions of inverted repeats are boxed in green, blue and red respectively for IR3, IR2 and IR1. Inverted repeats are italicized within boxes. Putative promoters P_{RM} and $P_{R'}$ are underlined.

DNA replication. The DNA replication machinery of AaΦ23 seems to be closely related to that of the lambdoid bacteriophages. ORF17 is similar to the protein GP18 of phage P22 which is related to the replication protein O of phage lambda (Swissprot accession number P03688). This protein is necessary for the initiation of the bi-directional DNA replication. Protein O interacts with the lambda origin of replication and protein P (Zylicz et al. 1984). Indeed, ORF19 present strong homologies to a replication P protein found in several bacteriophages and bacteria (Table 2). By its DNA helicase activity, this enzyme plays an essential role in a variety of processes such as DNA replication, repair, recombination and transcription (Waksman et al. 2000).

DNA recombination. ORFs 21 and 22 show homologies, respectively, to the ninG and ninB proteins expressed by several different phages (Table 2). In lambda, these two genes are known as *orf* (*ninB*) and *rap* (*ninG*), two recombination genes located in the *ninR* region. A recent study showed that orf and rap participate in Red recombination, the primary pathway operating when wild-type lambda grows lytically in rec^+ cells (Tarkowski et al. 2002). Additional experiments are needed to determine if AaΦ23 induces a similar pathway when it grows lytically in *A. actinomycetemcomitans*.

Methylase. DNA methylases participate in regulatory events of DNA replication, methyl-directed mismatch repair and transposition (Marinus 1996). These enzymes are associated with bacterial DNA restriction-modification systems whose are responsible for the degradation of foreign DNA such as conjugative plasmids and transposons (Meselson et al. 1972) or bacteriophage DNA (Bickle and Krüger 1993). It has been speculated that some bacteriophages express their own DNA adenine methylases in order to overcome this bacterial protection. DNA adenine methylases play also a role in regulating bacterial

virulence (Heithoff et al. 1999). One putative DNA adenine methylase (ORF20) is possibly expressed by AaΦ23 (Table 2). This ORF contains a D/N-P-P-Y/F motif, from amino acid 61 to 64, which is conserved in this family of enzymes (PFAM accession number PF02086, Timinskas et al. 1995). An *A. actinomycetemcomitans* DNA adenine methylase has been shown to be active in vivo (Eberhard et al. 2001). Nevertheless, the role of DNA adenine methylases in the regulation of *A. actinomycetemcomitans* virulence genes remains to be determined.

Late region.

Lytic enzyme complex. The lysis system of AaΦ23 seems to be composed of, at least, two different components. *Orf28* probably codes for the lytic enzyme of AaΦ23 which degrades the *A. actinomycetemcomitans* peptidoglycan barrier in order to allow the liberation of viral progeny into the extracellular media as this protein shares strong homologies with several lytic enzyme found in bacteriophages and bacteria (Table 2). This putative ORF matched also with amino acid sequences of members of the glycoside hydrolase family 19 (PFAM accession number PF00182, e-value of $1.2e^{-2}$). Additionally, ORF28 shows significant homologies with a putative chitinase (Table 2). The previous observation that some chitinases and several lysozymes from animals and bacteriophages present structural similarities (Holm and Sander 1994), support the possibility that the gene product of *orf28* represent the phage AaΦ23 lytic enzyme. ORF29 present significant homologies to the putative R_z lytic protein expressed by the *S. flexneri* bacteriophage SfV (Table 2). In many phages of Gram-negative hosts, R_z together with another short protein called R_{z1}, plays a role as an auxiliary lysis protein that is thought to interact with the outer membrane or with its links to the cell wall (Young et al. 2000). Thus, although no holin and R_{z1} genes could be identified, we can

speculate that AaΦ23 uses the same lysis strategy as decribed for many double-stranded DNA phages (Young et al. 2000).

DNA packaging. Two genes (*orf30* and *orf31*) are possibly coding for the two components of the terminase enzyme (Table 2). This holoenzyme is required for packaging of the phage genomic DNA into the empty capsid shells (Catalano 2000). In AaΦ23, the phage DNA has been shown to have random terminal redundancy of 1.6 kb that are circularly permuted (Willi and Meyer 1998) and it may therefore be packaged by the headfull mechanism. Thus, the AaΦ23 terminase complex may cut concatemeric DNA molecules, resulting from rolling circle replication, at a unique *pac* site to start the first round of packaging. Taking into consideration the 1.6 kb terminal redundancy, about 104% of the genomic DNA molecule is packaged into phage heads on subsequent rounds of packaging.

Structural components of the virion particle. The gene products of *orfs* 33, 46, 56, 57 and 51 seem to be, respectively, elements of the head, tail, tail fibers (ORFs 56 and 57) and baseplate of the AaΦ23phage particle (Table 2). ORF33 contains also a conserved domain, from amino-acid 156 to 256, found in members of a family representing putative minor phage head proteins (PFAM accession number PF04233, e-value of $2e^{-21}$).

Similarities with the genome sequence of the *A. actinomycetemcomitans* strain HK1651. Southern-blot experiments revealed two very distincts patterns when the ZIB1023 and HK1651 genomic DNA was hybridized with the AaΦ23 DNA (Figure 6). Analysis of the phage genomic sequence revealed two segments of strong homologies over 8.7% of the AaΦ23 genome (average of 97% homologies over 3,795 bps) with the *A. actinomycetemcomitans* strain HK1651 genomic sequence (Figure 1). On the protein level,

six AaΦ23 ORFs share significant homologies with ORFs of the *A. actinomycetemcomitans* strain HK1651 genome (Table 2). The very short extent of homologies between the *A. actinomycetemcomitans* strain HK1651 and the AaΦ23 genome and the very different patterns observed in southern-blot experiments suggest that strains HK1651 and ZIB1023 do not carry the same prophage. Therefore it was not surprising that we were not able to induce phages production from *A. actinomycetemcomitans* strain HK1651 by mitomycin C treatment. This DNA of limited homology to AaΦ23 could represent remnants of an AaΦ23-related prophage. However, the presence of an unrelated prophage with limited homologies to AaΦ23 in the genome of the strain HK1651 cannot be excluded.

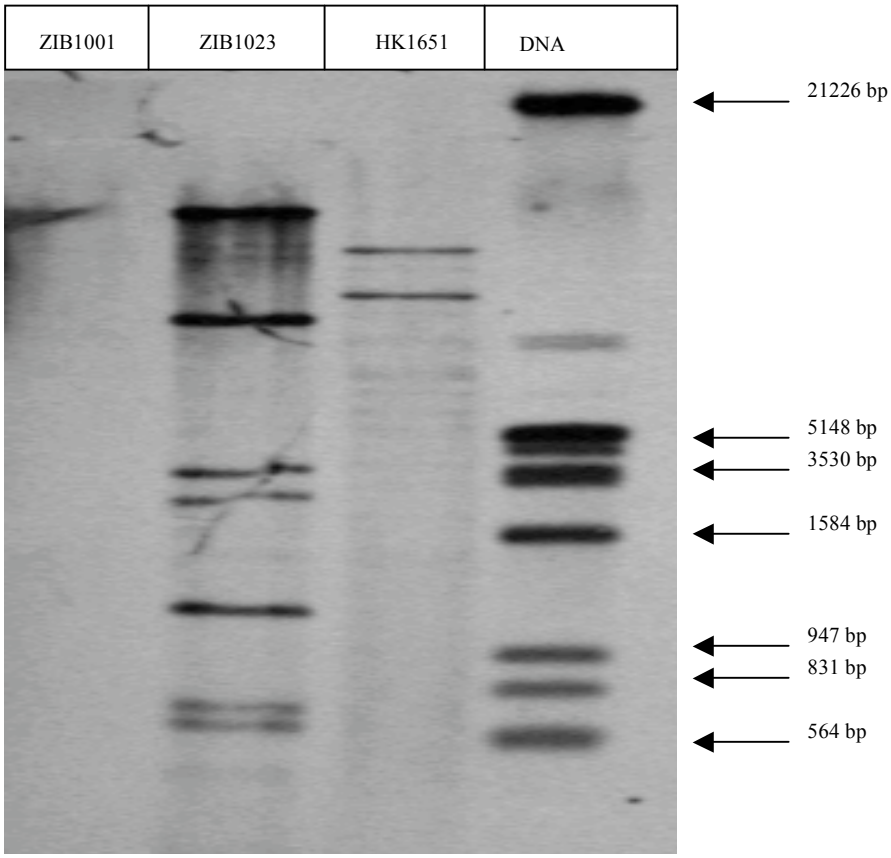

Figure 6. Southern-blot hybridization of genomic DNA of strains ZIB1001, ZIB1023 and HK1651 digested with *Eco*RI using phage AaΦ23 DNA as the probe.

Similarities to the plant pathogenic bacterium *Xylella fastidiosa*. One interesting observation is that 40% of the overall AaΦ23 ORFs (23/57) share homologies with hypothetical proteins from three strains of the plant pathogen *X. fastidiosa* (Table 2). In *X. fastidiosa* strain 9a5c, about 7% of the entire genome is related to double-stranded DNA phage sequences, mostly from the lambda group (Simpson and Setubal 2000). Four putative prophages, XfP1, XfP2, XfP3 and XfP4, were identified on this genome (Simpson and Setubal 2000). XfP3 and XfP4 are exactly similar over 14,317 bp. Homologies of thirteen AaΦ23 ORFs with the *X. fastidiosa* strain 9a5c genome can be found in this duplicated region (Table 2). In addition, three AaΦ23 ORFs (48, 50 and 51) share homologies with ORFs only found in the XfP3 prophage and three other (53, 54 and 55) are closely related to ORFs only present on the XfP4 prophage. These observations suggest that AaΦ23, XfP3 and XfP4 could have had a common ancestor. Moreover, three AaΦ23 ORFs (8, 14 and 49) showed similarities to *X. fastidiosa* strain 9a5c ORFs that are not carried by prophages (Table 2). Nevertheless, the *X. fastidiosa* homologue of AaΦ23 ORF8 is annotated as being a phage-related hypothetical protein (Table 2) and ORF14 may correspond to the Cro repressor protein in AaΦ23 (Table 2). ORF49 is the only one that present homologies with a putative ORF of the *X. fastidiosa* bacterium but not with a phage-related protein (Table 2). Finally, the putative integrase of AaΦ23 has homologies with a hypothetical protein found in the *X. fastidiosa* strains Ann-1 and Dixon but is not represented on the *X. fastidiosa* strain 9a5c genome (Table 2).

Acknowledgements. We are very grateful to Peter Philippsen for providing the sequencing facilities and to Tom Bickle for very fruitful discussions and comments on the manuscript. We thank Mogens Kilian for the *A. actinomycetemcomitans* strain HK1651. We also thank Eric Kofoid and Sophie Lemire-Brachat for their advices concerning the semi-random PCR

protocol and the annotation and submission of the AaΦ23 genomic sequence, respectively. Finally, we thank Sylvia Voegeli and Anita Lerch for technical assistance. The work was support by the Swiss Dental Association (SSO, grant nr. 196).

CHAPTER II

Characterization of the DNA adenine methylase of the *Actinobacillus actinomycetemcomitans* bacteriophage AaΦ23

ABSTRACT

In the following study we reported cloning of the DNA adenine methyltransferase gene (*dam*) of the *Actinobacillus actinomycetemcomitans* bacteriophage AaΦ23, experimental characterization of its product (M.AaΦ23Dam), and comparison with DNA adenine methylases from other phages. Morever, a pairwise alignment of M.AaΦ23Dam and the DNA adenine methylase of *A. actinomycetemcomitans* (M.*Aac*Dam) is proposed. Based on this alignment, comparative analysis of these enzymes and their close homologs revealed, despite an overall sequence divergence, conserved essential cofactor-binding and catalytic residues. However, significant differences were found in the cofactor-binding domain and regions that are thought to contact the target DNA sequence are missing in M.AaΦ23Dam.

INTRODUCTION

Bacterial DNA methyltransferases are associated with bacterial DNA restriction-modification systems to protect against phage infections (Bickle and Krüger 1993). Moreover, they participate in regulatory events of DNA replication, repair and transcription (Marinus 1996). On the basis of different chemical reactions, DNA methyltransferases can be divided into three different groups, generating N6-methyladenine (m^6A), N4-methylcytosine (m^4C), and C5-methylcytosine (m^5C). While nine protein sequence motifs are conserved among members of these groups, a region of essentially higher variability was also identified (Malone et al. 1995). Motifs I-III and X have been shown to be part of the AdoMet-binding subdomain, motifs IV-VIII were assigned to the active-site subdomain, and the variable region was recognized as a separate target-recognizing domain (Dryden 1999, Schlukebier et al. 1998, Kumar et al. 1998, Malone et al. 1995).

DNA adenine methylases catalyse the transfer of a methyl group from S-adenosyl-L-methionine (AdoMet) to the N-6 position of the adenine in a specific DNA sequence (Marinus and Morris 1973). A recent nomenclature for restriction enzymes and DNA methylases has been published (Roberts et al. 2003). DNA adenine methylases are classified within the Type I (as product of *hsdM* gene), Type II or Type III (as product of *mod* gene) restriction-modification systems and are divided in three distinct groups (α, β and γ) in respect to the sequential order of motifs of the protein sequence (Malone et al. 1995).

DNA adenine methylase regulation of gene expression was shown in *E. coli* for *trpR* (Peterson et al. 1985), Tn*10* transposase (Roberts et al. 1985), *dna*A (Braun and Wright 1986). Moreover, DNA adenine methylases were reported to control the expression of a number of virulence genes (Low et al. 2001). The expression of the *Escherichia coli* pyelonephritis-associated pilus (*pap*) genes, which encode pili that are required for infection

of the urinary tract, is reversibly switch between the unexpressed and the expressed states by a methylation-sensitive process termed phase variation (van der Woude et al. 1996, Braaten et al. 1994). *Dam*$^-$ mutants of *Salmonella enterica* serovar typhimurium are dramatically reduced in their virulence as compared to the wild-type strain (Heithoff et al. 1999). DNA adenine methylase overproduction was found to attenuate the virulence of *Yersinia pseudotiberculosis* and *Vibrio cholerae* leading to, in *Yersinia*, a fully protective immune response in vaccinated hosts (Julio et al. 2001). Additionally, DNA adenine methylation exerted by bacterial enzymes plays an important role during development of many phages. In Mu it is involved in the regulation of *mom* genes expression (Hattman 1999) and in P1 it is essential for cleavage of the *pac* site, which initiates the packaging of phage DNA during the lytic cycle (Sternberg and Coulby 1990). Nevertheless, several phages encode their own DNA adenine methylase. Despite evidence that some phage DNA is modified by virus-encoded methylases in order to evade host restriction (Günthert and Trautner 1984), the precise functions of these enzymes are still unclear as no systematic knock-out mutagenesis study has been performed for phages DAM to assess to what degree does their functions overlap with that of the host enzyme.

Recently the GATC-specific DNA adenine methylase of bacteriophages HP1 was characterized (M.HP1Dam, Piekarowicz and Bujnicki 1999) and compared to its *H. influenzae* homologue (M.*Hin*Dam, Bujnicki et al. 2001). Despite overall divergence, both enzymes share the common fold and essential cofactor-binding and catalytic residues. However, significant differences in the cofactor-binding pocket have been identified. Moreover, loops involved in contact with the target DNA sequence are missing in the phage enzyme.

Cloning, sequence analysis and heterologous expression of the GATC-specific DNA adenine methylase of *A. actinomycetemcomitans* (M.*Aac*Dam, Eberhardt et al. 2001)), a Gram-

negative facultative anaerobe that has been strongly implicated in the etiology of localized juvenile periodontitis (Zambon 1985), opened the path for studies on the potential regulatory impact of DNA methylation on gene regulation and virulence in this organism. In this bacterial species, lysogeny is widespread (Iff et al. 1997, Willi et al. 1997a, Loftus and Delisle 1995, Sandmeier et al. 1995, Willi et al. 1993). It has previously been shown that one representative of a family of *A. actinomycetemcomitans* bacteriophages, the bacteriophage AaΦ23, carries a gene coding for a putative DAM (M.AaΦ23Dam, Chapter I). To further characterize DAM occurring in *A. actinomycetemcomitans* we cloned and expressed the gene coding for M.AaΦ23Dam in an *E. coli dam*$^-$ host, compared the amino acid sequence of the corresponding gene product with known DNA adenine methylase sequences of other bacteriophages and aligned it with the recently characterized M.*Aac*Dam.

MATERIALS AND METHODS

Bacterial strains, media and culture conditions. *A. actinomycetemcomitans* ZIB1023 is the natural strain lysogenic for AaΦ23. ZIB1001, ZIB1015 are non lysogenic *A. actinomycetemcomitans* strains. ZIB1504 and ZIB1515 are respectively strains ZIB1015 and ZIB1001 that were experimentally lysogenized with AaΦ23 and *A. actinomycetemcomitans* HK1651 is a clinical isolate, the genome of which has been sequenced (Advanced Center for Genome Technology, Oklahoma's University), obtained from Mogens Kilian (University of Aarhus, Denmark). These strains were grown in Todd Hewitt Broth 3% (w/v) (THB, BD Biosciences, Allschwil, Switzerland) at 37°C in air and 10% CO_2. *E. coli* GM48 (*dam*$^-$, Palmer and Marinus 1994) electro-competent cells were used as hosts for recombinant plasmids. GM48 cells were grown in Luria-Bertani medium (Sambrook et al. 1989) at 37°C in air.

Purification of the AaΦ23 phage DNA. Phage DNA of AaΦ23 was prepared according to the procedure previously described (Willi et al. 1993). Briefly, after mitomycin C (1 μg/ml) induction of bacteriophage from a log-phase growing culture of *A. actinomycetemcomitans* strain ZIB1023, phage particles were recovered by a polyethylene glycol 8000 precipitation and centrifugations. Phage genomic DNA was extracted by phenol/chloroform and precipitated by ethanol. All reagents were purchased from Sigma (Buchs, Swizerland).

Figure 7. Nucleotide sequence of the previously identified gene coding for the putative M.AaΦ23Dam.

Start and stop codons are boxed. Putative M.AaΦ23Dam promoter is underlined. Hybridization sites of the primers *dam*L and *dam*R are represented in bold and underlined with arrows pointing in 5' to 3' orientation.

PCR conditions and PCR product purification. The genomic DNA of AaΦ23 was directly used as template in the PCR reactions. Primers were designed on the basis of the previously established complete genomic sequence of the bacteriophage AaΦ23 (Chapter I and Figure 7). Primers *dam*L (5'-GCGAATTCCTCAGATAAACGACCAATGCCA-3') and *dam*R (5'-GCGAATTCTCATTGCCGCCACTCCAAATTT-3') were chosen in order to amplify a fragment containing the gene coding for the putative M.Aa.Φ23Dam together with 300 bp upstream region to include the putative promoter sequence (Figure 7) identified with the promoter predictor software (http://www.fruitfly.org/seq_tools/promoter.html). Denaturation, primer annealing and extension were done respectively at 94°C, 57°C and 72°C for 35 cycles in a GeneAmp® PCR system 9700 apparatus (Perkin Elmer, Langen, Germany). PCR products were purified using the QIAquick® PCR Purification kit (Qiagen, Basel, Switzerland). Oligonucleotides were purchased from Microsynth GmbH (Balgach, Switzerland).

Cloning of the PCR fragment and sequencing of the recombinant vector. The PCR fragments obtained and the cloning vector pUC19 (Stratagene, Zürich, Switzerland) were digested with *Eco*RI and purified. After treatment of the vector with shrimp alkaline phosphatase, ligation of the PCR fragment into the vector was conducted 16 h at 16 °C in a total volume of 10 µl containing 1U of T4 DNA ligase. The ligation mixture was electroporated in *E.coli* GM48 electro-competent cells and recombinant clones were isolated on ampicillin-containing agar plates. Plasmids were recovered from selected clones and were first screened for insert by *Eco*RI digestion and gel electrophoresis. Plasmids carrying an *Eco*RI insert of the same size as the PCR fragment obtained with primers *dam*L and *dam*R were sequenced using the Big-dye technology (Perkin Elmer). Sequencing reactions were analysed on an ABI PRISM 377 automated DNA sequencer (Perkin Elmer). Enzymes were

purchased from Roche Diagostics AG (Rotkreuz, Switzerland) and other reagents were obtained from Sigma

Isolation of bacterial genomic DNA and methylation-sensitive restriction digestion. Genomic DNA of several *A. actinomycetemsomitans* strains and of *E. coli* strain GM48 was isolated with the QIAamp® DNA Mini Kit (Qiagen) following recommendations of the manufacturer. Methylation-sensitive restriction digestions were performed using *Sau*3AI and *Mbo*I (Bioconcept, Allschwil, Switzerland). Digestions were conducted overnight at 37°C and analysed on a 1% low-melting agarose gel (Seakem® LE Agarose, Bioconcept).

Sequence analysis. The amino acid sequence of the putative M.AaΦ23Dam was compared to the Swissprot database using the blastp software under ORF finder (http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi). Alignment of M.AaΦ23Dam with its homologs was done using ClustalW (http://www.ebi.ac.uk/clustalw) and modified manually, taking into account previously predicted conserved secondary structural elements.

RESULTS AND DISCUSSION

***In silico* identification of M.AaΦ23Dam.**

Comparison of M.AaΦ23Dam with known DNA adenine methylases. Table 3 lists other phage DNA adenine methylases showing significant homologies to M.AaΦ23Dam. An alignment of M.AaΦ23Dam with it phage homologs is represented (Figure 8). The M.AaΦ23Dam is closely related to the M.HP1Dam and M.HP2Dam (Table 3). We therefore tentatively applied the molecular model of M.HP1Dam (Bujnicki et al. 2001) to

M.AaΦ23Dam (Figure 8). The order of the conserved motifs identified on M.AaΦ23Dam indicates that it belongs to the γ group (Malone et al. 1995).

Based on structural superposition of molecular models, M.*Hin*Dam and M.HP1Dam were previously pairwise aligned and compared (Bujnicki et al. 2001). M.*Aac*Dam (Eberhard et al. 2001) and M.*Hin*Dam (Bujnicki et al. 2001) have extended amino acid sequence homologies, especially in the conserved motifs (Figure 9). The order of the conserved motifs identified on M.*Aac*Dam indicates that it belongs to the α group (Malone et al. 1995). As the molecular model of M.*Hin*Dam (Bujnicki et al. 2001) might be applicable to M.*Aac*Dam, we aligned and compared M.AaΦ23Dam with M.*Aac*Dam (Figure 10).

Table 3. Bacteriophage proteins showing significant homologies with M.AaΦ23Dam.

Protein name	Phage name	blastp e-value	Accession number
M.HP2Dam	Phage HP2	5e-20	NP_536818
M.HP1Dam	Phage HP1	5e-20	P51715
M.CP-933VDam	Coliphage CP-933V	3e-15	NP_288678
M.Lahn3Dam	Coliphage Lahn3	4e-15	CAC95062
M.Nil2Dam	Coliphage Nil2	4e-15	CAC95095
M.HK97Dam	Coliphage HK97	4e-15	NP_037744
M.VT2-SaDam	Coliphage VT2-Sa	5e-15	NP_050531

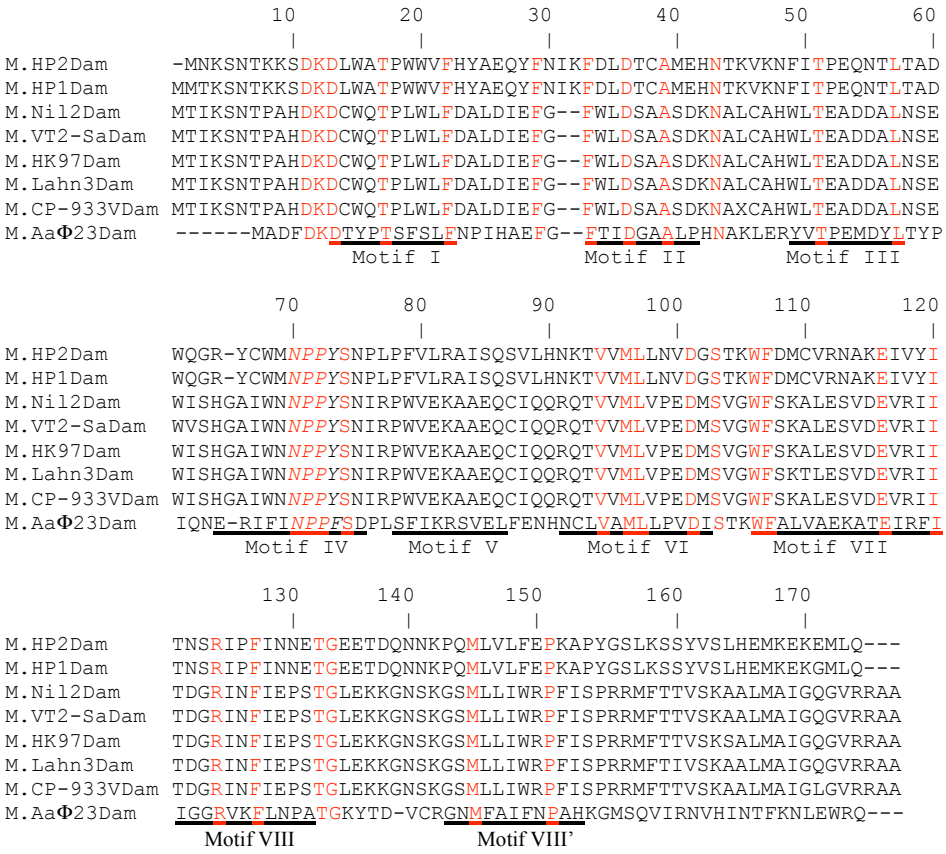

Figure 8. Sequence alignment of M.AaΦ23Dam with its phage homologues. Similar residues are coloured in red. Conserved motifs are underlined and numbered according to Malone et al. (1995). The conserved motif D/N-P-P-Y/F among protein family PF02086 is represented in italic within motif IV.

Comparison of regions interacting with the S-adenosyl-L-methionine. Motifs I-III and X are involved in the binding of the AdoMet (Labahn et al. 1994). A pronounced difference between M.AaΦ23Dam and M.*Aac*Dam is observed in the N terminal region before motif I (Figure 10). The region in M.AaΦ23Dam is shorter and lacks motif X. This motif, which is perfectly similar between M.*Aac*Dam and M.*Hin*Dam, has been shown to be part of the binding pocket for the methionine moiety of AdoMet in M.*Hha*IDam (Cheng et al. 1993). While in M.*Aac*Dam the G-loop of motif I includes the conserved tripeptide FxG (FVG[47]) of DNA adenine methylases of group α, M.AaΦ23 harboures an unusual YPT[11] tripeptide. The G-loop binds to the methionine moiety of the AdoMet. Replacement of the FV[46] dipeptide of M.*Aac*Dam motif I by the YP[10] dipeptide in M.AaΦ23Dam motif may not modify dramatical

the AdoMet binding pocket as both dipeptides have the same spatial bulk. This FxG motif may be conserved only among bacterial enzymes as the M.HP1Dam shows also a very different motif (WAT[17]). One rule of this motif I is respected in M.AaΦ23Dam and M.*Aac*Dam as the residue at position 2 amino acids upstream from the FxG consensus pattern is, respectively, the D[7] and the E[43]. This residue is thought to interact with the G-loop and therefore adding stereochemical constraints in the binding pocket.

```
                        10        20        30        40        50        60
                         |         |         |         |         |         |
M.AacDam        MPEPAKPATP--AKSRPFLKWAGGKYRLMDEINRLLPKRKQCLVEPFVGAGAVFLNSNFK
M.HinDam        MLRPKKQSLKPKLKHRPFLKWAGGKFRLTDEINKAFPNKKNCLIEPFVGAGAVFLNSNFE
                                    Motif X             Motif I

                        70        80        90       100       110       120
                         |         |         |         |         |         |
M.AacDam        RYILADINPDLINLFNIVKLDVERYIQACKPIFFHPEANTEIYYYGKRKEFNQSTDVFQR
M.HinDam        RYILADINPDLINLFNIVKXNVDGYIEDCKPIFFADDANTPDYYYAKRRQFNASTEPFER
                  Motif II

                       130       140       150       160       170       180
                         |         |         |         |         |         |
M.AacDam        AVLFLYLNRFGFNGLCRYNSKNEFNVPFGDYKTHYFPEEELRFFAAKAQSAVFICADFQQ
M.HinDam        SIIFLYLNRFGFNGLCRYNSKNEFNVPFGAYKTHYFPEDELRYFAHKAQSAVFLCCDFQK
                                 TRD                              Motif III

                       190       200       210       220       230       240
                         |         |         |         |         |         |
M.AacDam        TFEMADENSVIYCDPPYAPLSQDSNFTNYAGNEFSIGHQRDLANLAKHTMEQRNIQVLIS
M.HinDam        TFEFADKDSVIYCDPPYAPLQQETNFTGYAGNEFGLAQQRALADLAKSIQKEKQISILIS
                          Motif IV                  Motif V        Motif VI

                       250       260       270       280
                         |         |         |         |
M.AacDam        NHDTPFTREIYQGAKIRRLKVQRSISQAPHKRIKVRELIAMFKQGK
M.HinDam        NHDTKFTREIYNGAKFKRVKVQRSISQNPEKRVKVKELIAIFGARK
                    Motif VII  Motif VIII     Motif VIII'
```

Figure 9. Sequence alignment of the M.*Aac*Dam and M.*Hin*Dam. Similar residues are coloured in red. Conserved motifs are underlined and numbered according to Malone et al. (1995). The target-recognizing domain (TRD) is underlined with a dashed line. The conserved motif D/N-P-P-Y/F conserved among protein family PF02086 is represented in italic within motif IV.

```
M.AacDam        MPEPAKPATP--AKSRPFLKWAGGKYRLMDEINRLLPKRKQCLVEPFVGAGAVFLNSNFK
M.AaΦ23Dam                                                MADFDKDTYPTSFSLFNPIHAE
                                          Motif X             Motif I

M.AacDam      RYILADINPDLINLF(92)AQSAVFICADFQQTFEMADENSVIYCDPPYAP(19)QRDLANL
M.AaΦ23Dam    FG--FTIDGAALPHN----AKLERYV-TPEMDYLTYPIQNERIFINPPFSD----PLSFIKR
                  Motif II             Motif III      Motif IV      Motif V
```

```
M.Aac Dam       AKHTMEQRNIQVLISNHDTPFTREIYQGAKIRRLKVQRSISQAPH----------KRIKVR
M.AaΦ23Dam      SVELFENHNCLVAMLLPVD-ISTKWFALVAEKATEIRFIIGGRVKFLNPATGKYTDVCRGN
                         Motif VI          Motif VII

M.Aac Dam       ELIAMFKQGK
M.AaΦ23Dam      MFAIFNPAHK
               Motif VIII'
```

Figure 10. Sequence alignment of the M.*Aac*Dam and M.AaΦ23Dam derived from sequence comparison of M.*Aac*Dam with M.*Hin*Dam and M.AaΦ23Dam with M.HP1Dam. Structural conserved motifs are underlined and numbered according to Malone et al. (1995). The conserved motif D/N-P-P-Y/F, conserved among the protein family PF02086, is represented in italic within motif IV.

In motif II, a D is conserved among the group α and is also found in M.HP1Dam (D^{36}). This charged residue interacts with the ribose hydroxyl of the AdoMet. In M.AaΦ23Dam it is replaced by a polar T^{26} that may also interact with the hydroxyl group of the AdoMet by hydrogen bonds. The following residue is usually a bulky hydrophobic residue. It is also the case in M.*Aac*Dam and M.AaΦ23Dam as a non-polar I is found at this position in both enzymes. This residue is making van der Waals contacts with the AdoMet adenine (Malone et al. 1995).

Comparison of residues involved in the calalytic activity and target recognition. Motifs IV to VIII are thought to be part of the active domain of DNA adenine methylases. The conserved D/N-P-P-Y/F motif encountered in motif IV of DAM that represents the active site of the enzyme is also present in the putative motif IV of M.AaΦ23Dam ($NPPF^{64}$). Nevertheless, M.AaΦ23Dam, as well as the *Eco*RI associated DNA adenine methylase, represent exception of the γ group which members usually have a NPPY motif (Malone et al. 1995). Putative motif VII of M.AaΦ23Dam begins with a relatively hydrophobic residue (N^{82}) and ends with two hydrophobic residues (LL^{89}) in place of the usual hydrophobic residues SN. In motif VIII, the F^{196} of M.*Taq*IDam that is suggested to make favourable van

der waals contacts to the target DNA adenine (Schluckebier et al. 1995), is replaced by an unusual Q in M.*Hin*Dam and M.*Aac*Dam. In view of the alignment, in M.AaΦ23Dam the corresponding amino acid may be the R^{109} of motif VII. Similarly, an R has been described at this position for the DNA adenine methylase of *E. coli* (Malone et al. 1995). Nevertheless an F, which might also be a good candidate for the residue that contacts the DNA adenine, is present at position 110 on M.AaΦ23Dam. Moreover, this F^{110} is replaced in M.HP1Dam by the Y^{118} residue. The role of the F^{196} of M.*Taq*IDam is possibly played by a corresponding Y residue in motif VIII of several DNA adenine methylases of the α group (Malone et al. 1995). The observation that the TRD is lacking in all bacteriophages DNA adenine methylases studied so far, is also made for M.AaΦ23Dam. Moreover, the loop between motifs IV and V, which makes contact with the target DNA, is also missing in M.AaΦ23Dam as for M.HP1Dam and others (Figure 8). Finally, the putative DNA binding loop described for M.HP1Dam and represented by the insertion between motifs VII and VIII' (Bujnicki et al. 2001) is also present on M.AaΦ23Dam but not in M.*Aac*Dam (Figure 8 and 10).

PCR amplification and cloning of the AaΦ23 *dam* gene. By using the primers *dam*L and *dam*R on the genome of phage AaΦ23, a fragment of approximately 800 bp in size was amplified (Figure 11). The size of this fragment is in good agreement with the size of the gene coding for M.AaΦ23Dam manually determined on the previously established nucleotide sequence (Figure 7). This 800 bp PCR product was successfully cloned into the *Eco*RI site of the cloning vector pUC19 to yield recombinant plasmid pUC19::M.AaΦ23*dam* (Figure 11).

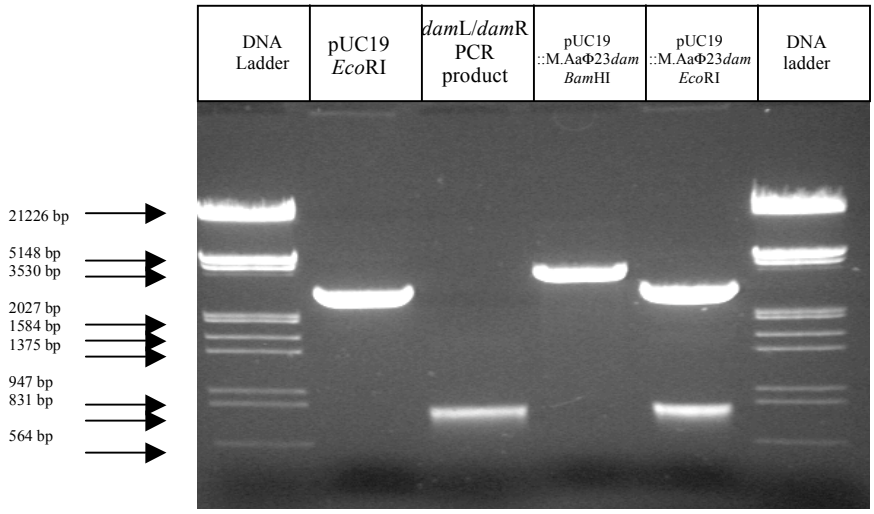

Figure 11. Cloning of the *dam*L/*dam*R PCR product in the vector pUC19 digested with *Eco*RI.

***In vivo* activity of M.AaΦ23Dam in *E. coli* GM48 *dam*⁻ cells.** In order to find out if M.AaΦ23*dam* is expressed in *E. coli* GM48, bacterial DNA was isolated from both GM48 and GM48(pUC19::M.AaΦ23*dam*) and digested with the restriction enzymes *Mbo*I and *Sau*3AI. *Mbo*I recognizes and cleaves only the non-methylated palindromic genomic GATC DNA sequences, while the *Sau*3AI isoschizomer is not sensible to DNA adenine methylation. As shown on Figure 12 the genomic DNA of wild-type *E.coli* GM48 is restricted by both enzymes. This confirms that in *E. coli* GM48 the adenine residue of genomic GATC sequences is not methylated. In contrast, The genomic DNA of GM48(pUC19::M.AaΦ23*dam*) is not cleaved by *Mbo*I. This proves that the genomic DNA has been methylated by the M.AaΦ23Dam activity. Moreover, results suggest that the putative promoter of M.AaΦ23Dam is functional in *E.coli*.

M.AaΦ23Dam may be classified in the Type IIP group of DNA methyltransferase as it recognize and methylate a palindromic DNA sequence (Roberts and Xu 2003). Nevertheless, no genes potentially coding for a cognate endonuclease associated with M.AaΦ23Dam has been identified on the genome of the AaΦ23 bacteriophage (Chapter I).

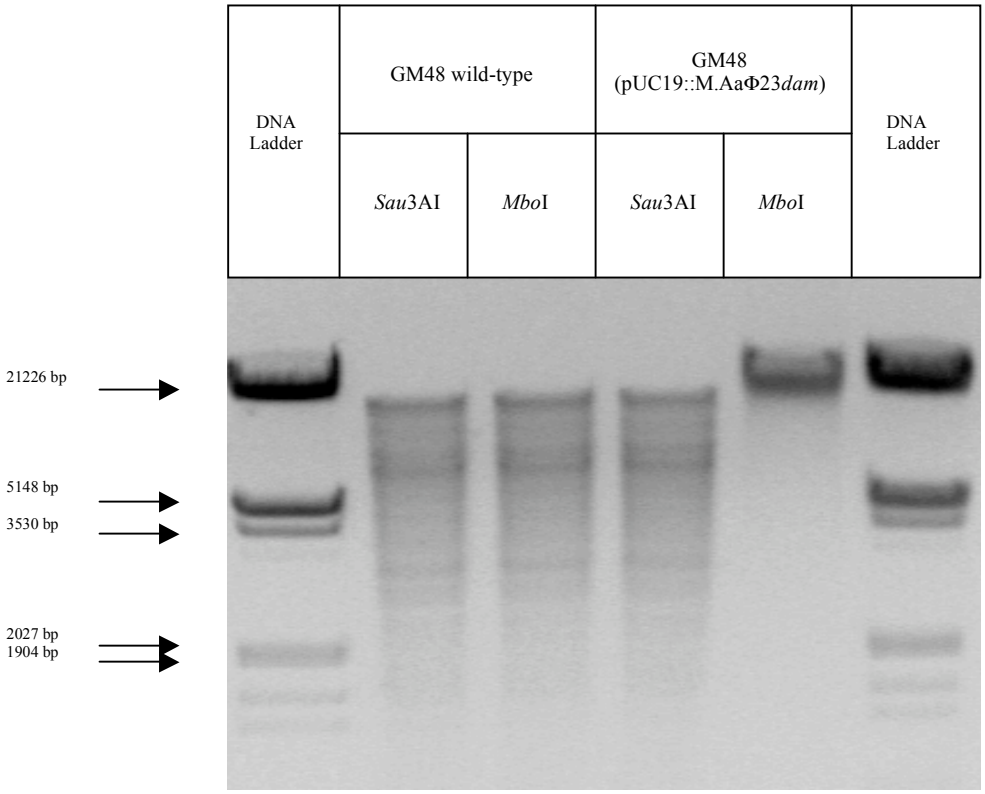

Figure 12. Methylation sensitive digestion of genomic DNA of *dam*⁻ *E. coli* GM48 and GM48(pUC19::AaΦ23Dam) cells.

DNA adenine methylase activity in six clinical isolates of *A. actinomycetemcomitans*. Three non-lysogenic and three strains lysogenic for AaΦ23 were tested. While *Mbo*I was not able to cut the genomic DNA of the strains (Figure 13a), *Sau*3AI successfully restricted all of them (Figure 13b). These results are in agreement with the expression of a functional DNA adenine methylase that methylates the adenine residue of the genomic GATC of all tested strains. Moreover, bio-informatics analysis revealed that the genome of strain HK1651 carries a *dam* gene similar to the one previously cloned (Eberhard et al. 2001). Thus, DNa adenine methylases may be widespread and conserved among *A. actinomycetemcomitans.*

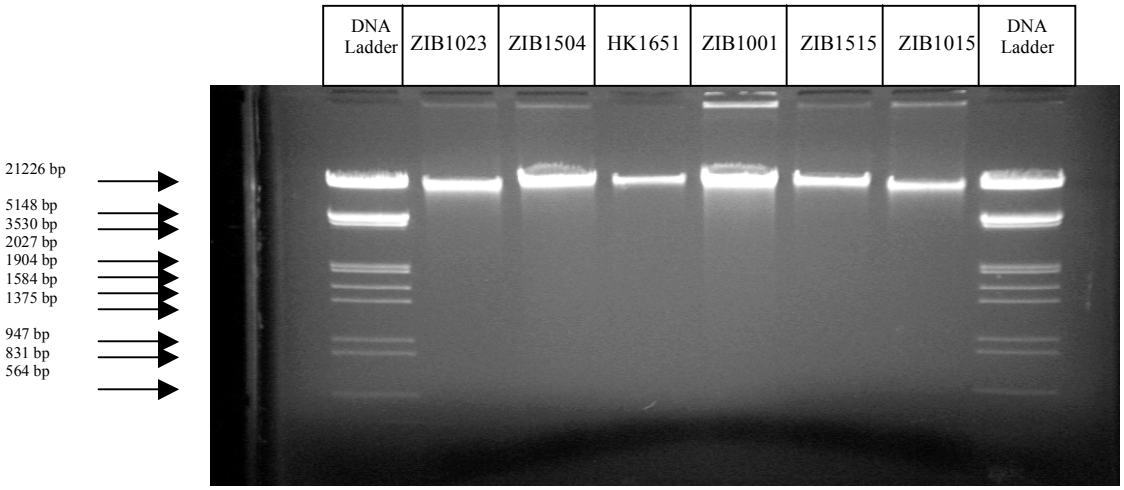

Figure 13a. Digestion of genomic DNA isolated from several *A. actinomycetemcomitans* with the methylation-sensitive restriction enzyme *Mbo*I.

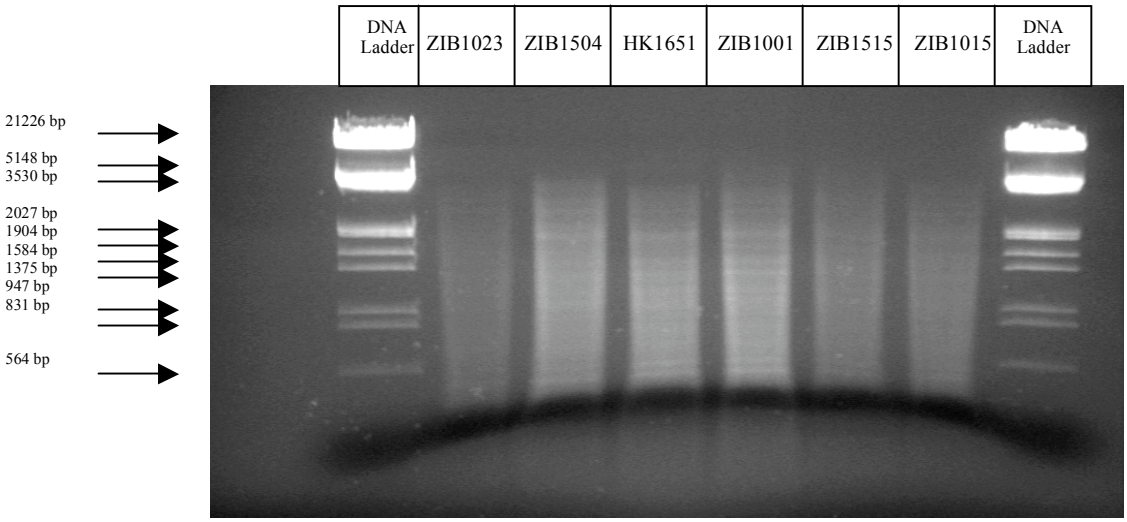

Figure 13b. Digestion of genomic DNA isolated from several *A. actinomycetemcomitans* with the methylation non-sensitive restriction enzyme *Sau*3AI.

CONCLUSIONS

A functional DNA adenine methylase is encoded in the genome of AaΦ23. This enzyme show the same specificity as the DNA adenine methylases expressed by three non-lysogenic and three AaΦ23-lysogenized *A. actinomycetemcomitans* strains. The analysis presented here help to visualize similarities and differences between the M.AaΦ23Dam and its host counterparts. These informations might be very useful for designing futur experiments that could elucidate the roles played by M.AaΦ23Dam in the development of AaΦ23 and possibly in the virulence of *A. actinomycetemcomitans.* Nevertheless, one obstacle for latter study will be the lack of virulence tests adapted to *A. actinomycetemcomitans*. In lysogenic bacteria, not only the chromosomal encoded DNA adenine methylase might be expressed but also homologue enzymes encoded on prophages. Our study emphasizes that similarities and differences between the DNA adenine methylases of bacteria and phages may be crucial for designing new inhibitors. In fact, in bacterial cells lysogenized with a DNA adnine methylase-encoding phage, treatment with specific inhibitors of bacterial DNA adenine methylases might remain ineffective.

Acknowledgements. We thank Thomas Bickle for the *E. coli* strain GM48.

CHAPTER III

The lytic enzyme

of *Actinobacillus actinomycetemcomitans*

bacteriophage AaΦ23

Preliminary results

INTRODUCTION

Bacteriophages were, directly after their discovery, considered as very promising agents for antibacterial therapy. In fact, the large majority of these bacterial viruses is naturally able to lyse its host. However, due to the rapid emergence of phage resistance, this approach had little clinical significance. Recent studies showed that purified lysin was able to kill *Streptococcus pneumoniae* (Loeffler et al. 2001) and *Bacillus anthracis* (Schuch et al. 2002) *in vitro* and eradicate a *Streptococcus pyogenes* infection in mice (Nelson et al. 2001). This approach has only been tested with Gram-positive microorganisms because they have no outer membrane and when the lytic enzyme is administrated from outside, it may be able to reach the peptidoglycan barrier without the help of an holin.

In many double-stranded DNA phages of Gram-negative bacteria the lysis process is achieved by an enzymatic complex consisting of a lytic enzyme (encoded by the *lys* gene), a holin and accessory lytic proteins R_z and R_{z1} (Young et al. 2000). The complete genome sequence sequence of phage AaΦ23 reveals a gene possibly coding for a lytic enzyme (*orf*28). No candidate gene coding for a holin was identified (Chapter I). We undertook the

cloning of the *orf*28 into an expression vector in the hope to be able to overexpress and purify the corresponding gene product for subsequently testing its activity.

MATERIALS AND METHODS

Bacterial strains, media, and culture conditions. *A. actinomycetemcomitans* ZIB1023 is the natural strain lysogenic for AaΦ23. This strain was grown in Todd Hewitt Broth 3% (w/v) (THB, BD Biosciences, Allschwil, Switzerland) at 37°C in air and 10% CO_2. *Escherichia coli* XL1 Blue electro-competent cells (Stratagene, Zürich, Switzerland) were used as hosts for the recombinant plasmids and *E. coli* BL21 supercompetent cells (*E. coli* B F^- ompT *hsd*S($r_B^-m_B^-$) dcm^+ Tet^R gal λ(DE3) endA Hte [pLysS Cam^R]) (Stratagene) were used as hosts for the recombinant expression vector. *E. coli* cells were grown in Luria-Bertani medium (Sambrook et al. 1989) at 37°C in air.

Purification of the AaΦ23 phage DNA. Phage DNA of AaΦ23 was prepared according to the procedure previously described (Willi et al. 1993). Briefly, after mitomycin C (1 μg/ml) induction of bacteriophage of a log-phase growing culture of *A. actinomycetemcomitans* strain ZIB1023, phage particles were recovered by a polyethylene glycol 8000 precipitation and centrifugations. Phage genomic DNA was extracted by phenol/chloroform and precipitated by ethanol. All reagents were purchased from Sigma (Buchs, Swizerland).

PCR amplification of the *lys* gene. PCR amplification of a fragment containing the *lys* gene was obtained using primers *lys*L (5'-CCGGAATTCATGTTTATCACACAA-3') and *lys*R (5'-CCGGAATTCTCACCCCAATAAGTT-3') and the phage AaΦ23 DNA as template. Denaturation, primer annealing and extension were done respectively at 94°C, 55°C and

72°C for 35 cycles in a GeneAmp®PCR system 9700 apparatus (Perkin Elmer, Langen, Germany). PCR product was purified using the QIAquick® PCR Purification kit (Qiagen, Basel, Switzerland). Oligonucleotides were purchased from Microsynth GmbH (Balgach, Switzerland).

Cloning of the PCR fragment and sequencing of the recombinant cloning vector. The PCR fragment obtained and the cloning vector pBluescript II SK$^-$ (Stratagene) were digested with *Eco*RI and purified. After treatment of the vector with shrimp alkaline phosphatase, ligation of the PCR fragment into the vector was conducted 16 h at 16°C in a total volume of 10 µl containing 1U of T4 DNA ligase. The ligation mixture was electroporated into *E.coli* XL1 Blue electro-competent cells and recombinant clones were isolated on ampicillin-containing agar plates. Plasmids were recovered from selected clones and were first screened for insert by *Eco*RI digestion and gel electrophoresis. Plasmids carrying *Eco*RI fragments of the same size as the PCR fragment obtained with primers *lys*L and *lys*R (pBluescript::AaΦ23*lys*) were sequenced using the Big-dye technology (Perkin Elmer, Langen, Germany). Sequencing reactions were analysed on an ABI PRISM 377 automated DNA sequencer (Perkin Elmer). After digestion of the selected pBluescript::AaΦ23*lys* with *Eco*RI, the purified insert was subcloned into the expression vector pTYB1 (Bioconcept, Allschwil, Switzerland). Briefly, 1µg of pTYB1 was linearized with *Eco*RI during 8 h at 37°C in the appropriate buffer. After purification and dephosphorylation with shrimp alkaline phosphatase, 50 ng of pTYB1 was ligated with excess of insert in a total volume of 10 µl containing 1U of T4 DNA ligase. Ligation mixtures were incubated for 16 h at 16°C. BL21 cells were mixed with the ligation mixtures and heat-schock pulsed according to the recommendations of the manufacturer. Transformed cells were immediately diluted with prewarmed SOC medium (Sambrook et al. 1989) and incubated 1h at 37°C in air with

shaking. Finally, cells were plated on agar plates containing ampicillin (50 μg/ml). Enzymes were purchased from Roche Diagnostics AG and other reagents were obtained from Sigma.

RESULTS AND DISCUSSIONS

By using the primers *lys*L and *lys*R, a fragment of expected 600 bp in size was amplified (Figure 14). The PCR product was successfully cloned into the prokaryotic cloning vector pBluescript II SK^- (Figure 14) but not in pTYB1. Sequencing of pBluescript::AaΦ23*lys* plasmids recovered from 20 selected clones showed that the *lys* gene was inserted always in the non-transcribed orientation. These results suggest that expression of the *lys* gene, even to background level, may be lethal for *E. coli* cells. Attempts, by using primers with different restriction sites, were made to force the *lys* gene to insert into the cloning and expression vectors in the transcribed orientation. No recombinant clones with both vectors were obtained. We thought that these observations could provide the first indirect evidence that *orf28* gene may code for the lytic enzyme of AaΦ23.

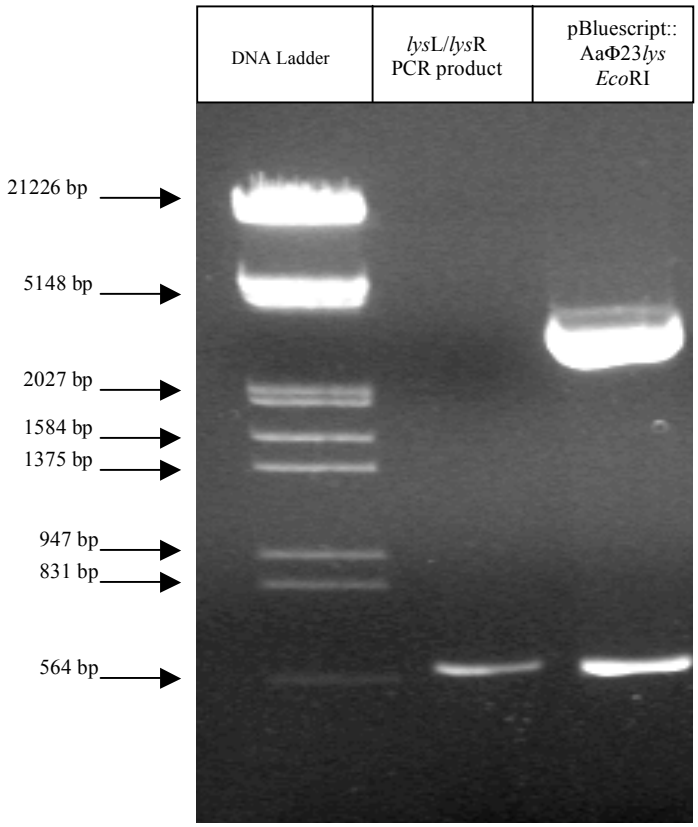

Figure 14. Cloning of the *lys*L/*lys*R PCR products into the vector pBluescript II SK$^-$ digested with *Eco*RI.

CHAPTER IV

The integration system of *Actinobacillus actinomycetemcomitans* bacteriophage AaΦ23. Preliminary results

INTRODUCTION

A. actinomycetemcomitans is a Gram-negative human periodontal pathogen (Zambon 1985). A large portion of clinical isolates are lysogenized by morphologically and genetically related bacteriophages (Haubek et al. 1997, Iff et al. 1997, Willi et al. 1993). The bacteriophage AaΦ23 derived from *A. actinomycetemcomitans* strain ZIB1023 has a 43,033 bp double-stranded genome (Chapter I). In the prophage state, AaΦ23 DNA is integrated in the host chromosome at a specific site. The attachment sites of the phage (*att*P) and the bacterial host (*att*B) have been identified (Chapter I) and suggest a site-specific recombination process (Campbell 1962). Several forms of site-specific recombination are well characterized in Gram-negative bacteria and the best-studied system is that of phage lambda (Landy 1989). The recombination process is catalysed by a phage-encoded integrase (Int). An putative *int* gene has been suggested to be present on the AaΦ23 genome (Chapter I). The gene product was proposed as a new member of the Int family of site-specific recombinases. Similarities and differences between members of this family have been extensively studied (Nunes-Düby et al. 1998). Based on these results, an analysis of the

AaΦ23 Int amino acid sequence is presented. In order to prove the function of the proposed Int, we constructed an integrative vector based on the integration elements of phage AaΦ23.

MATERIALS AND METHODS

Bacterial strains, media, and culture conditions. *A. actinomycetemcomitans* ZIB1023 is the natural strain lysogenic for AaΦ23. ZIB1001 is a non-lysogenic strain of *A. actinomycetemcomitans*. Both strains were grown in Todd Hewitt Broth 3% (w/v) (THB, BD Biosciences, Allschwil, Switzerland) at 37°C in air and 10% CO_2.

Purification of the AaΦ23 phage DNA. Phage DNA of AaΦ23 was prepared according to the procedure previously described (Willi et al. 1993). Briefly, after mitomycin C (1 μg/ml) induction of bacteriophage from a log-phase growing culture of *A. actinomycetemcomitans* strain ZIB1023, phage particles were recovered by a polyethylene glycol 8000 precipitation and centrifugations. Phage genomic DNA was extracted by phenol/chloroform and precipitated by ethanol. All reagents were purchased from Sigma (Buchs, Swizerland).

PCR amplification of the integration region . PCR amplification of a fragment containing the *int* gene, its putative promoter (P*int*) and the *att*P site was done using primers *int*L (5'-GCGAATTCGTATCGCGTCATAAATGCCGTG-3') and *int*R (5'-GCGAATTCTGATTTCATTCAGCAGAAATCA-3') and phage AaΦ23 DNA as the template. Denaturation, primer annealing and extension were done respectively at 94°C, 55°C and 72°C for 35 cycles in a GeneAmp® PCR system 9700 apparatus (Perkin Elmer, Langen; Germany). PCR product was purified using the QIAquick® PCR Purification Kit (Qiagen, Basel, Switzerland).

Cloning of the PCR fragment into pKT210. The PCR product and the broad host range cloning vector pKT210 (Bagdasarian et al. 1981) were digested with *Eco*RI and purified. After treatment of the vector with shrimp alkaline phosphatase, ligation of the PCR fragments into the vector was conducted 16 h at 16°C in a total volume of 10 µl containing 1U of T4 DNA ligase. Enzymes were purchased from Roche Diagostics AG (Rotkreuz, Switzerland) and other reagents were obtained from Sigma (Buchs, Switzerland).

Preparation of *A. actinomycetemcomitans* ZIB1001 electro-competent cells, electroporation with the ligation mixture, and selection of recombinant clones. Electro-competent ZIB1001 cells were prepared according to Sreenivasan et al. (1991). Briefly, a single colony of *A. actinomycetemcomitans* ZIB1001 was cultived over-night in THB at 37°C in air and 10% CO_2. Three additional 1:20 dilutions followed by over-night culture were done. The cell culture was then diluted 1:1 and incubated at 37°C in air and 10% CO_2 till the OD_{600nm} reached a value of 0.4. Cells were washed twice with half a volume of ice-cold buffer EPB (glycerol 15% (v/v), sucrose 272 mM, K_2HPO_4 2,43 mM, KH_2PO_4 0,57 mM). Finally, pelleted cells were resuspend in ice-cold EPB buffer and frozen at -70°C. 60 µl of electro-competent *A. actinomycetemcomitans* ZIB1001 cells were transformed by electroporation with the ligation mixture. Settings of the electro cell manipulator® were V=2.5 kV; R=186 Ω. Electroporated cells were immediately diluted in 1 ml of prewarmed THB and grown for 5H at 37°C in air and 10% CO_2. Finally, cells were plated on blood-agar (BA) plates containing chloramphenicol (50 µg/ml). Antibiotic was purchased from Sigma.

Sequence analysis. The amino acid sequence of the P22 and lambda Int were retrieved manually from NCBI (http://www.ncbi.nlm.nih.gov/) with their accession numbers. AaΦ23

Int and retrieved sequences were aligned with ClustalW (http://www.ebi.ac.uk/clustalw) and annotated manually according to Nunes-Düby et al. (1998).

RESULTS AND DISCUSSION

***In silico* identification of the AaΦ23 Int.** The first interesting observation is that the AaΦ23 ORF1 is closely related to the Int of several *Podoviridae* and a cryptic *E. coli* prophage (Table 2 of Chapter I). Homologies with Int found in *Myoviridae* such as P1, HP1 and HP2 are still significant but weaker (data not shown). In order to find out conserved motifs and residues in M.AaΦ23 Orf, its amino acid sequences was aligned with the P22 and lambda Int (Figure 15 and 16). The conserved boxes I and II and additional patches I, II and III previously described (Nunes-Düby et al 1998) can be delimited on the AaΦ23 ORF1 amino acids sequence (Figures 15 and 16).

The first observation is that the conserved tetrad RHRY described in all Int and which includes the active site tyrosine (Pargellis et al. 1998, Nunes-Düby et al. 1998, Abremski and Hoess 1992) is present in AaΦ23 ORF1 ($R^{221}H^{320}R^{323}Y^{355}$ on Figure 15). In boxes I and II, about 60% of the amino acid residues are identical among AaΦ23 Orf1 and its P22 homologue (Figure 15). The pattern of conserved hydrophobic residues forming the core of the globular structure that was described for all members of the Int family (Kwon et al. 1997), is conserved in the AaΦ23 ORF1. Only three amino acids substitutions in this pattern are not favorable. Lambda Int residue Y (position 197 on Figure 16) is replaced by the acidic residue E (position 187 on Figure 16) in AaΦ23 ORF1. The residue T (position 291 on Figure 16) is replaced by the basic residue K (position 291 on Figure 16) and the residue corresponding to L (position 273 on Figure 16) is missing in AaΦ23 ORF1. Nevertheless the conservation of hydrophobic residues is 87%, meaning that the overall hydrophobic character of this pattern

is conserved. In patch I, the consensus LT-EEV--LL (Nunes-Düby et al. 1998) and is represented by the sequence LT-HEE--LI (Figures 15 and 16) on AaΦ23 ORF1. The second E residue in the LT-EEV--LL consensus is conserved (position 196 on Figure 15 and 185 on Figure 16). This residue is of particular importance as E169K mutated P2 Int are ineffective in recombination (Nunes-Düby et al. 1998).

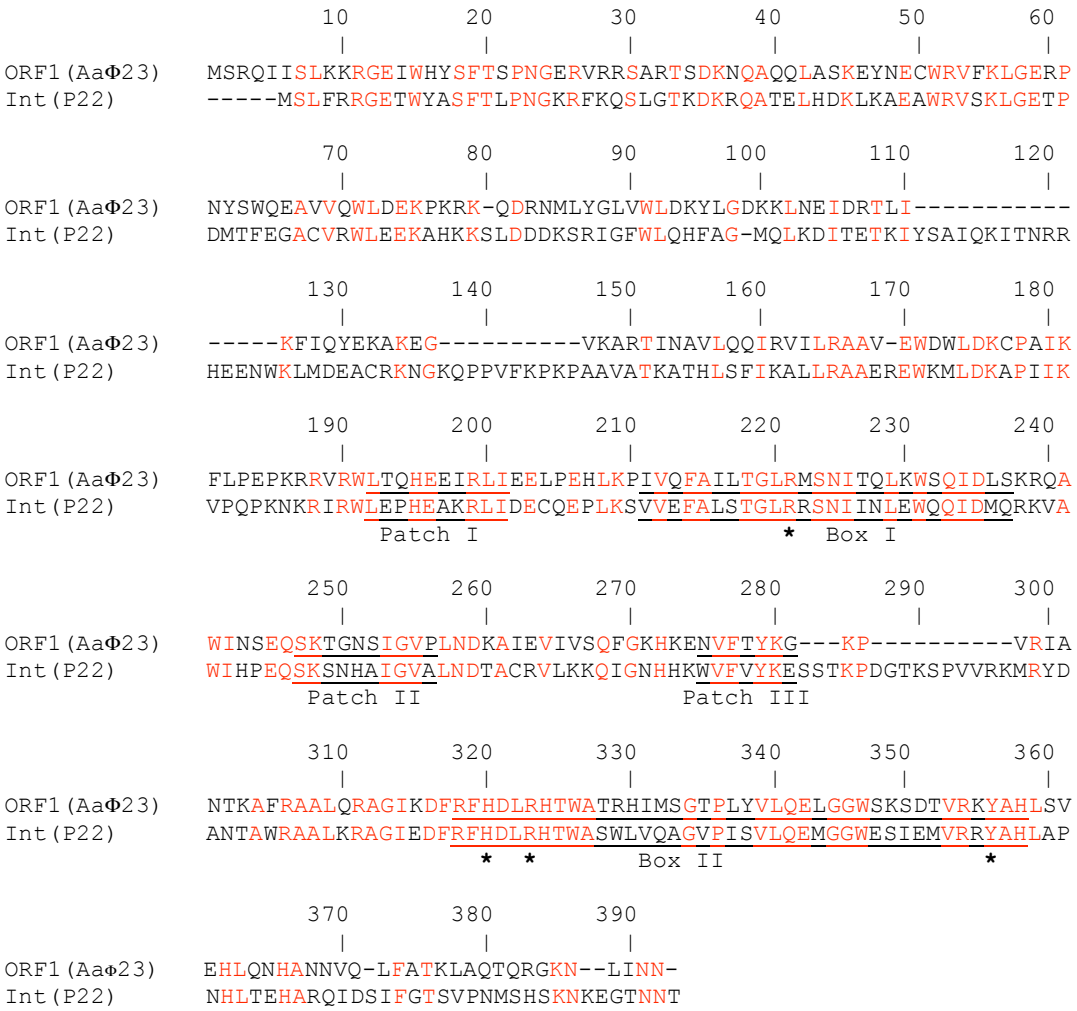

Figure 15. Alignment of AaΦ23 ORF1 with the amino acid sequence of P22 Int (accession number DAA00998). Identical residues are in red. The conserved domains are underlined and numbered according to Nunes-Düby et al. (1998). The highly conserved RHRY tetrad containing the active site tyrosine is indicated by asterisks.

In patch II, a consensus (SKTG--I--P^{249} in AaΦ23 Int, Figure 16) is common for the AaΦ23 ORF1 and lambda Int. The conserved K of this consensus is at one edge of the catalytic

pocket (Kwon et al. 1997). Substitution of the adjacent T in lambda Int (position 242 on Figure 16) provokes a decrease in recombination activity (Han et al. 1994).

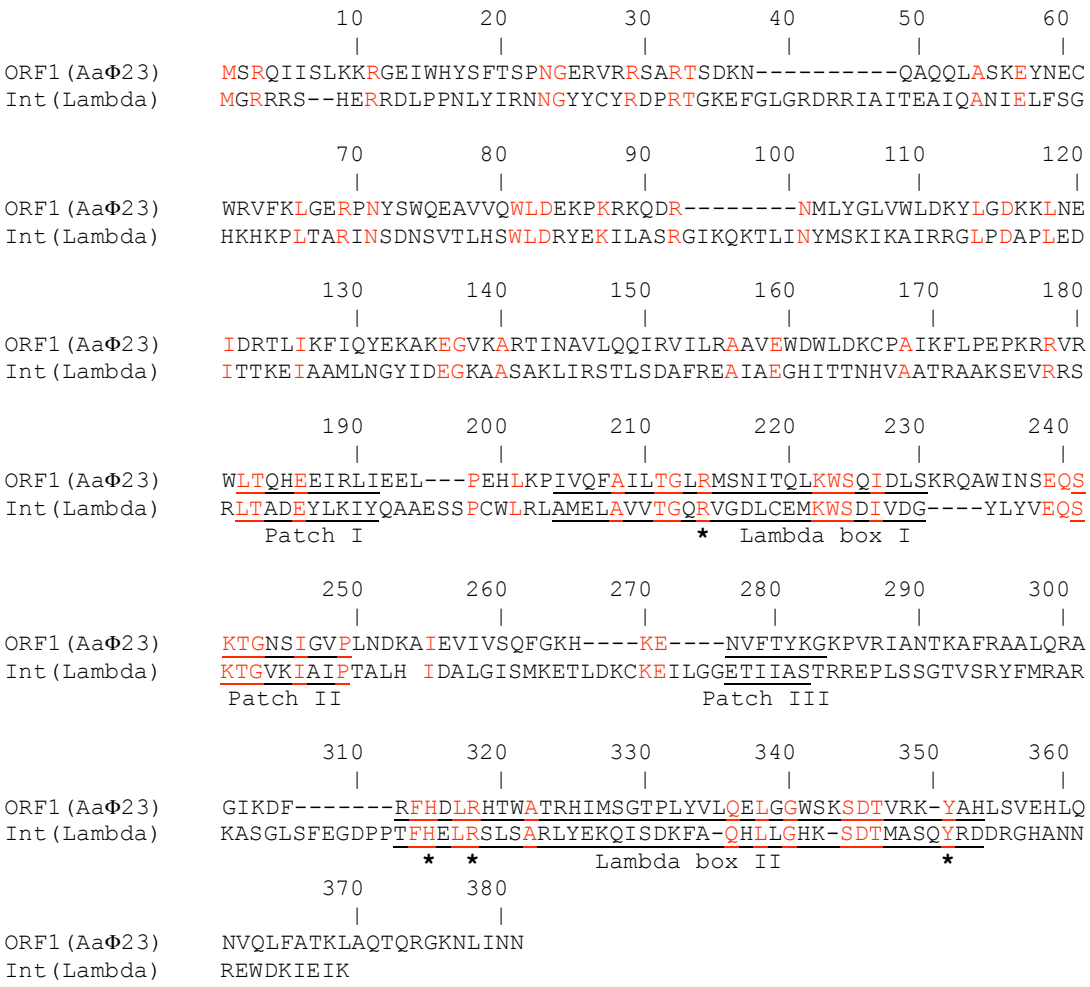

Figure 16. Alignment of AaΦ23 ORF1 with the amino acid sequence of the lambda Int (accession number P03700). Identical residues are in red. The conserved domains are underlined and numbered according to Nunes-Düby et al. (1998). The highly conserved RHRY tetrad containing the active site tyrosine is indicated by asterisks.

Patch III is less conserved among Int but a consensus pattern (VF-YK) can be designed for AaΦ23 ORF1 and its P22 homologue (Figure 15). Patch III is thought to be important in stabilizating the native folds of the Int recombinases. After patch III, a stretch of 13 amino acids is missing is comparison with P22 Int and the region between Patch III and Box II comprises only 23 residues in AaΦ23 ORF1 (Figure 15). Nevertheless, AaΦ23 putative Int is

not unique for this characteristic as the same region in bacteriophage SF6 contains only 25 residues (Nunes-Düby et al. 1998). One remarkable difference between AaΦ23 ORF1 and its P22 homologue is the deletion of 26 residues in the N terminal region of the protein (residues 111 to 125 and 137 to 146, Figure 15). No functions for the corresponding residues in other Int have been reported, suggesting that this region is not essential.

Identification of conserved Patchs, Boxes and key residues in the AaΦ23 ORF1 strongly suggests that the corresponding gene codes for the Int protein of AaΦ23.

PCR amplification and cloning of the AaΦ23 *int* gene. By using the primers *int*L and *int*R, it has been possible to amplify a fragment of approximately 1,400 bp in size (Figure 17). The size of this fragment is slightly shorter than the size of 1,692 bp manually determined using the previously established nucleotide sequence of the integrase gene (Figure 18). Explanation for this observation may be obtained by the sequencing of the PCR products.

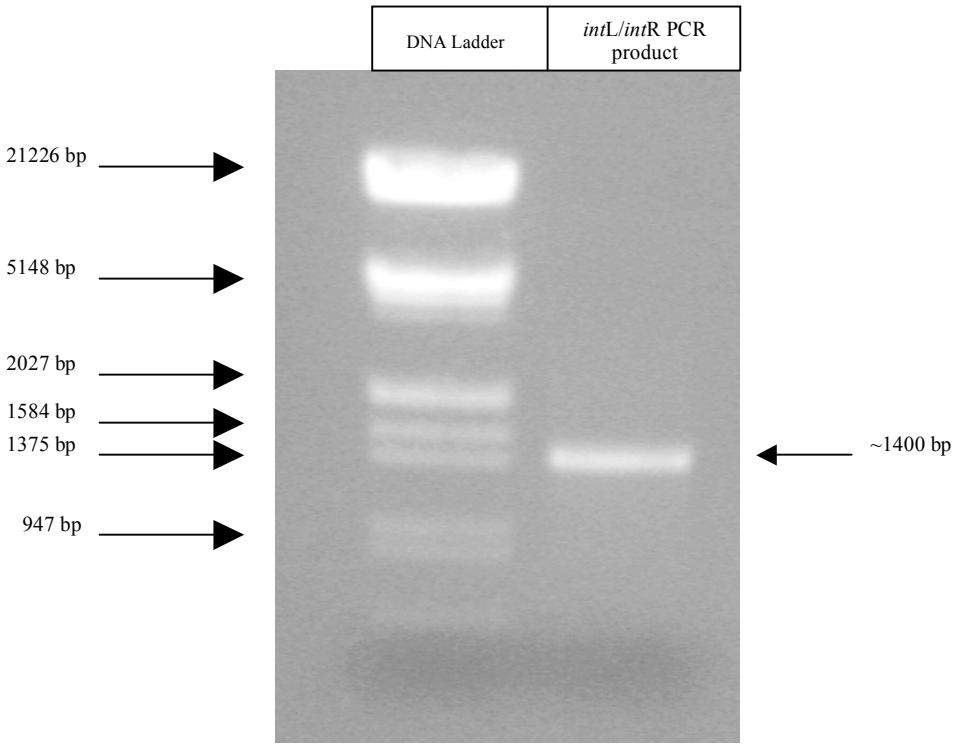

Figure 17. PCR products obtained by amplification of the AaΦ23 genomic region between primers *int*L and *int*R.

GTTTG**GTATCGCGTCATAAATGCCGT**GTTTTTCGACAAAAATCACTGCAAAAAGGCGTACGAAAGCGAAATTAATCGCCGACA
*int*L
ATACACTCAGGACTTTCAATAGCAAAAAAACGGCGGGTAATGTCGTGAAGTGTCATTTTATTTTCCTCGTTTTATGAGGTTTG

CATGTGTGTTGATGTGTGCCATATTTGTGCCATTGATAAACGGCTAACTGCTGCTATTTGCTTCTGATGACTACGATTAATTT

ATAAGCAAGAGGCGTAAGGCATTGTTTTACAATGATTTATGTATAGGCTTTCAAAT***ATGGCATTCAAGAGGTCGTCGGTTCGA***
attP core

TCCCGATTATCTCCACCAAATTTTATAGCTTTATCAATTAGTTAGTTATTGATAAGGTTTTTTCCTCTTTGTGTTTGTGCCAA

TTTTGTGGCAAAAAGTTGAACATTGTTGGCGTGGTTTTGTAAATGCTCAACAGATAAATGCGCATATTTTCTCACGGTGTCTG

ATTTACTCCACCCGCCCAACTCTTGCAGCACATATAACGGCGTTCCCGACATAATGTGTCTTGTCGCCCAAGTATGGCGCAAG

TCGTGAAAACGGAAGTCTTTGATTCCAGCTCGTTGTAATGCCGCACGGAACGCTTTAGTGTTTGCAATGCGAACCGGCTTGCC

TTTATATGTGAACACGTTTTCTTTGTGCTTGCCAAATTGTGATACAATCACTTCAATAGCTTTGTCGTTTAACGGTACGCCAA

TGCTATTACCTGTTTTTGATTGTTCCGAATTTATCCAAGCCTGTCTTTTGGATAAGTCAATTTGCGACCATTTTAGCTGGGTG

ATGTTCGACATTCTTAACCCAGTCAAAATAGCAAATTGAACAATTGGCTTTAAGTGCTCCGGCAATTCTTCTATGAGCCTTAT

TTCTTCGTGTTGGGTCAACCATCGTACACGTCGTTTGGGTTCAGGTAAGAATTTTATTGCTGGGCATTTATCCAACCAATCCC

ACTCAACCGCCGCTCTTAATATCACTCTTATCTGCTGCAAAACGGCATTAATTGTTCTTGCCTTAACGCCTTCCTTCGCTTTT

TCATATTGGATGAATTTTATTAAAGTGCGGTCAATTTCGTTGAGTTTTTTATCACCAAGATATTTATCAAGCCATACTAACCC

GTAAAGCATGTTTCTATCTTGTTTGCGCTTTGGCTTTTCATCAAGCCATTGTACTACGGCTTCCTGCCAGGAATAATTTGGGC

GCTCACCAAGTTTAAATACTCGCCAGCATTCGTTATATTCCTTTGAGGCTAATTGCTGCGCTTGATTTTTGTCTGAAGTCCTA

GCGCTTCGGCGTACGCGCTCACCATTTGGCGACGTGAACGAGTAATGCCAGATTTCCCCTCTTTTCTTGAGGGAGATGATTTG

ACGTGACATATATTCTCCTTATGCCCCGCCTCTTGCTCGGGGTGATCTTTATCCTTTTTCAAG**A**ATTTTTCAAGCTCTTCAAC
*P*i*nt*

ATCAAATTTCCATTTTCTGCCGATCTTGTAAGCGGGTATTTCTTCTTTGTTTGCCATCTCCCGTAGTGTGACGGGGTGCATAT

CAAGGATTCTTGCCGCTTTATCTGTGCCTATTAATTCCATAATTTCTCCAAAAATAAAGCCCATTTGTGCGCTTATTATCTAA

TCAATATCACGGTTTAATATTTCGCTTATCTTGTCACCT**GATTTCTGCTGAATGAAATCA**TTT
*int*R

Figure 18. Nucleotide sequence of the previously identified gene coding for the putative integrase (Chapter I) of AaΦ23. The sequence is represented in the reversed orientation as on the entire AaΦ23 genomic sequence. Start and stop codons are boxed. Putative Int promoter (P*int*) is underlined. The *att*P core sequence is indicated in bold italic. Hybridization sites of the primers *int*L and *int*R are in bold and underlined with arrows pointing in 5' to 3' orientation.

Electroporation of the ligation mixture in *A. actinomycetemcomitans* ZIB1001. After electroporation of ZIB1001 with the ligation mixture (which might contain recombinant vectors pKT210::AaΦ23*int*), no colonies were obtained on BA plates containing chloramphenicol. However, colonies were observed when ZIB1001 was transformed with the

religated plasmid pKT210 indicating that the ligation process worked and that ZIB1001 cells were competents. Further efforts are needed for constructing an efficient integrative vector based on the integration elements of AaΦ23.

CONCLUSIONS

Two approaches were used in order to characterize the Integrase of AaΦ23.

1. Identification of the AaΦ23 Int by *in silico* analysis of AaΦ23 ORF1.
2. Cloning of the putative *int* gene, *att*P site and putative promoter pI into a plasmid.

The indisponibility of a non-replicative vector adapted for cloning within *A. actinomycetemcomitans* obligates use to test our integration fragment with the Gram-negativ broad-host range pKT210 cloning vector. As expected, no colonies were obtained when ZIB1001 competent cells were transformed with pKT210::AaΦ23*int* meaning that this integrative vector integrates with high efficiency into ZIB1001 chromosome. Subsequent replication of the integrated vector within the host chromosome might be lethal for the cells. Nevertheless, no evidence that the pKT210::AaΦ23*int* recombinant vector was generated during the ligation process is available. Further studies using a non-replicative construct (generated for example by removing the origin of replication within pKT210) have to be done in order to characterize the Int activity.

GENERAL CONCLUSIONS AND PERSPECTIVES

1. The AaΦ23 genome is composed of 43,033 bp and has an organization similar to that of lambdoid phages. Of 57 ORFs, 37 have no putative functions and 12 seem to be specific for AaΦ23. As more and more data are available every day it will be very interesting to compare regularly these ORFs with databases. By analogy to lambda, it can be speculated that within these 37 ORFs of unknown functions several structural component might found. In fact, lambda carries 21 genes involved in the formation of the virion protein structure, while in AaΦ23, only 5 genes coding for structural components were identified. Experiments using SDS-PAGE of phage particles may be useful in order to characterize additional structural components.

2. The role of AaΦ23 in the virulence of *A. actinomycetemcomitans* is controversial. The work presented here shows that no protein with homologies to characterized phage-encoded virulence factors is coded on the AaΦ23 genome. Nevertheless, *orf20* codes for a functional DNA adenine methylase. It is well known that these enzymes play a role in the virulence of *Salmonella typhimurium*, *Yersinia pseudotuberculosis* and *Vibrio cholerae.* Whether the identified DNA adenine methylase plays a role in the pathogenesis of *A. actinomycetemcomitans* associated periodontitis remains to be determined. Limitations for this study might be the lack of a simple model for testing the virulence of *A. actinomycetemcomitans*.

3. A putative integrase belonging to the family of integrase/recombinases has been identified on the AaΦ23 genome and the *att*P attachment site has been determined. New strategies may be applied in order to develop an integration vector based on the integration elements of phage AaΦ23. This vector will represent a useful tool for

genetic studies in *A. actinomycetemcomitans*. Moreover, an *att*B homologue sequence has been found in *H. influenzae* meaning that the integration vector could represent an *A. actinomycetemcomitans*/*H. influenzae* shuttle vector.

4. A putative lytic system for AaΦ23 was identified. The *lys* gene was cloned in a cloning vector. All transformants carried the *lys* gene in the orientation which is not transcribed so that the lytic protein was not efficiously produced. This observation represents indirect evidence that *lys* codes for the lytic enzyme of AaΦ23. Further studies of this putative enzyme in systems allowing the complete suppression of background genetic expression might be of particular interest in order to purify this potential lytic protein. However, a therapeutic use of it in the treatment of *A. actinomycetemcomitans* mediated periodontitis is questionable.

LIST OF REFERENCES

Abremski KE and Hoess RH. 1992. Evidence for a second conserved arginine residue in the integrase family of recombination proteins. Protein Eng. 5 : 87-91.

Ackerman HW. 2001. Frequency of morphological phage descriptions in the year 2000. Arch. Virol. 146 : 843-857.

Allison GE, Angeles D, Tran-Dinh N and Verma NK. 2002. Complete genomic sequence of SfV, a serotype-converting temperate bacteriophage of *Shigella flexneri*. J. Bacteriol. 184 : 1974-1987.

Asikainen S, Lai CH, Alaluusua S and Slots J. 1991. Distribution of *Actinobacillus actinomycetemcomitans* serotypes in periodontal health and disease. Oral Microbiol. Immunol. 6 : 115-118.

Auvray F, Coddeville M, Ordonez RC and Ritzenthaler P. 1999. Unusual structure of the *att*B site of the site-specific recombination system of *Lactobacillus delbrueckii* bacteriophage mv4. J. Bacteriol. 181 : 7385-7389.

Bagdasarian M, Lurz R, Ruckert B, Franklin FC, Bagdasarian MM, Frey J and Timmis KN: 1981. Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. Gene. 16 : 237-247.

Barretto-Tinoco EM, Stevens R, Haubek D, Lai CH, Balachandran S and Preus H. 1997. Relationship of serotype, leukotoxin gene and lysogeny in *Actinobacillus actinomycetemcomitans* to periodontal disease status. Eur. J. Oral Sci. 105 : 310-317.

Beckerich A and Haudroy P. 1922. The bacteriolysant therapy of bacillary dysentery in children. Therapeutic application of bacteriolysant; d'Hérelle's phenomenon. Compt. Rend. Soc. Biol. 86 : 168.

Beighton D, Taichman NS, Simpson DL, Dirienzo JM and Johnson NW. 1989. Intra-oral colonization of macaque monkeys by *Actinobacillus actinomycetemcomitans*. Oral Microbiol. Immunol.4 : 35-40.

Bickle TA and Krüger DH. 1993. Biology of DNA restriction. Microbiol. Rev. 57 : 434-450.

Birge EA. Transduction . In Bacterial and Bacteriophages Genetics, Fourth Edition. Springer-Verlag New York Inc. Ed. New York, USA. 293-313.

Biswas B, Adhya S, Washart P, Paul B, Trostel AN, Powell B, Carlton R and Merril CR. 2002. Bacteriophage therapy rescues mice bacteremic from a clinical isolate of vancomycin-resistant *Enterococcuc faecium*. Infect. Immun. 70 : 204-210.

Braaten BA, Nou X, Kaltenbach LS and Low DA. 1994. Methylation patterns in *pap* regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli.* Cell. 76 : 577-588.

Bradley DE. 1967. Ultrastructure of bacteriophages and bacteriocins. Bacteriol. Rev. 31 : 230-314.

Braun RE and Wright A. 1986. DNA methylation differentially enhances the expression of one of the two *E. coli dna*A promoters in vivo and in vitro. Mol. Gen. Genet. 202 : 246-250.

Brunoghe R and Maisin J. 1921. Essais de thérapeutique au moyen du bactériophage du Staphylocoque. Compt. Rend. Soc. Biol. 85 : 1029-1121.

Brüssow H. 2001. Phages of dairy bacteria. Annu. Rev. Microbiol. 55 : 283-303.

Bujnicki JM, Radlinska M, Zaleski P and Piekarowicz A. 2001. Cloning of the *Haemophilus influenzae* Dam methyltransferase and analysis of its relationship to the Dam methyltransferase encoded by the HP1 phage. Acta Biochem. Pol. 48 : 969-983.

Campbell AM. 1962. Episomes. Adv. Genet. 11 : 101-145.

Campbell A. 1994. Comparative molecular biology of lambdoid phages. Annu. Rev. Microbiol. 48 : 193-222.

Catalano CE. 2000. The terminase enzyme from bacteriophage lambda : a DNA-packaging machine. Cell. Mol. Life Sci. 57 : 128-148.

Cheng X, Kumar S, Posfai J, Pflugrath JW and Roberts RJ. 1993. Crystal structure of the *Hha*I methylase complexed with S-adenosyl-L-methionine. Cell. 74 : 299-307.

Coleman DC, Sullivan DJ, Russell RJ, Arbuthnott JP, Carey BF and Pomeroy HM. 1989. *Staphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: molecular mechanism of triple conversion. J. Gen. Microbiol. 135 : 1679-1697.

Contreras A, Rusitanonta T, Chen C, Wagner WG, Michalowicz BS, Slots J. 2000. Frequency of 530-bp deletion in *Actinobacillus actinomycetemcomitans* leukotoxin promoter region. Oral Microbiol. Immunol. 15 : 338-340.

Davison WC. 1922. Le bactériophage dans le traitement de la fièvre typhoïde. Am. J. Dis. Child. 23 : 531.

d'Hérelle F. 1917. Sur un microbe invisible antagoniste des bacilles dysentériques. Compt. Rend. Acad. Sci. Ser. D. 165 : 373.

Dryden DT. 1999. Bacterial DNA methyltransferases. In : S-Adenosylmethionine-dependent methyltransferases : structures and functions. Cheng X and Blumenthal RM. (eds). World Scientific Inc. Singapore. 283-340.

Duckworth DH. 1976. "Who discovered bacteriophage?". Bacteriol. Rev. 40 : 793-802.

Eaton MD and Stanhope BJ. 1934. Bacteriophage Therapy. JAMA 103 : 1769-1776; 1847-1853; 1934-1939.

Eberhard J, Oza J and Reich NO. 2001. Cloning, sequence analysis and heterologous expression of the DNA adenine-(N(6)) methyltransferase from the human pathogen *Actinobacillus actinomycetemcomitans*. FEMS Microbiol. Lett. 195 : 223-229.

Ebersole JL, Taubman MA, Smith DJ, Genco RJ and Frey DE. 1982. Human immune responses to oral micro-organisms. I. Association of localized juvenile periodontitis (LJP) with serum antibody responses to *Actinobacillus actinomycetemcomitans*. Clin. Exp. Immunol. 47 : 43-52.

Eguchi Y, Ogawa T and Ogawa H. 1988. Cleavage of bacteriophage phi-80 CI repressor by RecA protein. J. Mol. Biol. 202 : 565-573.

Fives-Taylor PM, Meyer DH, Mintz KP and Brissette C. 1999. Virulence factors of *Actinobacillus actinomycetemcomitans*. Periodontol. 2000. 20 : 136-167.

Fleishmann RD and Venter JC. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science. 269 : 496-512.

Genco RJ, Van Dyke TE, Park B, Ciminelli M and Horoszewicz H. 1980. Neutrophil chemotaxis impairment in juvenile periodontitis : evaluation of specificity, adherence, deformability and serum factors. J. Reticuloendothel. Soc. 28 : 81-91.

Günthert U and Trautner TA. 1984. DNA methyltransferases of *Bacillus subtilis* and its bacteriophages. Curr. Top. Microbiol. Immunol. 108 : 11-22.

Haffajee AD and Socransky SS. 1994. Microbial etiological agents of destructive periodontal diseases. Periodontol. 2000. 5 : 78-111.

Hammond BF, Lillard SE and Stevens RH. 1987. A bacteriocin of *Actinobacillus actinomycetemcomitans*. Infect. Immun. 55 : 686-691.

Han YW, Gumport RI and Gardner JF. 1994. Mapping the functional domains of bacteriophage lambda integrase protein. J. Mol. Biol. 235 : 908-925.

Hattman S. 1999. Unusual transcriptional and translational regulation of the bacteriophage Mu *mom* operon. Pharmacol. Ther. 84 : 367-388.

Haubek D, Willi K, Poulsen K, Meyer J and Killian M. 1997. Presence of bacteriophage AaΦ23 correlates with the population genetic structure of *Actinobacillus actinomycetemcomitans*. Eur. J. Oral Sci. 105 : 2-8.

Hauser MA and Scocca JJ. 1992. Site-specific integration of the *Haemophilus influenzae* bacteriophage HP1. Identification of the points of recombinational strand exchange and the limits of the host attachment site. J. Biol. Chem. 267 : 6859-6864.

Hayashi T, Matsumoto H, Ohnishi M and Terawaki Y. 1993. Molecular analysis of a cytotoxin-converting phage, phi CTX, of *Pseudomonas aeruginosa* : structure of the *att*P-cos-ctx region and integration into the serine tRNA gene. Mol. Microbiol. 7 : 657-667.

Heithoff DM, Sinsheimer RL, Low DA and Mahan MJ. 1999. An essential role for DNA adenine methylation in bacterial virulence. Science. 284 : 967-970.

Hermann SR, Miller JA, O'Neill S, Tsao TT, Harding RM and Dale JL. 2000. Single primer amplification of flanking sequences. Biotechniques 29 : 1176-1178.

Holm L and Sander C. 1994. Structural similarity of plant chitinases and lysozymes from animals and phages. J. Bacteriol. 340 : 129-132.

Hongo M, Oki T and Ogata S. 1972. Phage contamination and control. In : The Microbial Production of Amino Acids. Yanada K, Kinoshita S, Tsunida T and Aida K Eds. Tokyo : Tokyo Kodansha.

Iff M, Willi K, Guindy J, Zappa U and Meyer J. 1997. Prevalence and clinical significance of a temperate bacteriophage in *Actinobacillus actinomycetemcomitans*. Acta Med. Dent. Helv. 2 : 33-38.

Inouye T, Ohta H, Kokeguchi S, Fukui K and Kato K. 1990. Colonial variation and frimbriation of *Actinobacillus actinomycetemcomitans*. FEMS Microbiol. Lett. 69 : 13-18.

Johnson LP, Tomai MA and Schlievert PM. 1986. Bacteriophage involvement in group A *streptococcal* pyrogenic exotoxin A production. J. Bacteriol. 166 : 623-627.

Julio SM, Heithoff DM, Provenzano D, Klose KK, Sinsheimer RL, Low DA and Mahan MJ. 2001. DNA adenine methylase is essential for viability and plays a role in the pathogenesis of *Yersinia pseudotuberculosis* and *Vibrio cholerae*. Infect. Immun. 69 : 7610-7615.

Kaplan JB and Fines DH. 1998. Codon usage in *Actinobacillus actinomycetemcomitans*. FEMS Microbiol. Lett. 163 : 31-36.

Kaplan JB, Perry MB, MacLean LL, Furgang D, Wilson ME and Fine DH. 2001. Structural and genetic analyses of O polysaccharide from *Actinobacillus actinomycetemcomitans* serotype f. Infect. Immun. 69 : 5375-5384.

Karaolis DKR, Somara S, Maneval DR, Johnson JA and Kaper JB. 1999. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. Nature. 399 : 375-379.

Kiley P and Holt SC. 1980. Characterization of the lipopolysaccharide from *Actinobacillus actinomycetemcomitans* Y4 and N27. Infect. Immun. 30 : 862-873.

Kimizuka R, Miura T and Okuda K. 1996. Characterization of *Actinobacillus actinomycetemcomitans* hemolysin. Microbiol. Immunol. 40 : 717-723.

Klinger R. 1912. Untersuchungen über menschliche Aktinomycose. Zentralbl. Bakteriol. 62 : 191-200.

Kropinski AM. 2000. Sequence of the genome of the temperate, serotype-converting, *Pseudomonas aeruginosa* bacteriophage D3. J. Bacteriol. 182 : 6066-6074.

Kumar S, Cheng X, Klimasauskas S, Mi S, Posfai J, Roberts RJ and Wilson GG. 1994. The DNA (cytosine-5) methyltransferases. Nucl. Acids Res. 22 : 1-10.

Kwon HJ, Tirumalai R, Landy A and Ellenberger T. 1997. Flexibility in DNA recombination : structure of the lambda integrase catalytic core. Science. 276 : 126-131.

Labahn J, GranzinJ, Schlukebier G, Robinson DP, Jack WE, Schildkraut I and Saenger W. 1994. Three-dimensional structure of the adenine-specific DNA methyltransferase M.*Taq*I in complex with the cofactor S-adenosylmethionine. Proc. Natl. Acad. Sci. USA. 91 : 10957-10967.

Laird W and Groman N. 1976. Orientation of the *tox* gene in the prophage of corynebacteriophage beta. J. Virol. 19 : 228-231.

Lally ET, Golub EE, Kieba IR, Taichman NS, Rosenbloom J, Rosenbloom JC, Gibson CW and Demuth DR. 1989. Analysis of the *Actinobacillus actinomycetemcomitans* leukotoxin gene. Delineation of unique features and comparison to homologous toxins. J. Biol. Chem. 264 : 15451-15456.

Landy A. 1989. Dynamic, structural and regulatory aspects of lambda site-specific recombination. Annu. Rev. Biochem. 58 : 913-949.

Löe H, Theilade E and Jensen B. 1965. Experimental gingivitis in man. J. Periodontol. 36 : 177-187.

Löe H, Anerud A, Boysen H and Smith M. 1978. The natural history of periodontal disease in man : the rate of periodontal destruction between 40 years of age. J. Periodontol. 491 : 607-620.

Loeffler JM, Nelson D and Fischetti VA. 2001. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. Science. 294 : 2170-2172.

Loessner MJ, Inman RB, Lauer P and Calendar R. 2000. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes* : implication for phage evolution. Mol. Microbiol. 35 : 324-340.

Loftus A and Delisle AL. 1995. Inducible bacteriophages of *Actinobacillus actinomycetemcomitans*. 30 : 317-321.

Low DA, Weyand NJ and Mahan MJ. 2001. Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. Infect. Immun. 69 : 7197-7204.

Malone T, Blumenthal RM and Cheng X. 1995. Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. J. Mol. Biol. 253 : 618-632.

Mangan DF, Taichman NS, Lally ET and Wahl SM. 1991. Lethal effects of *Actinobacillus actinomycetemcomitans* leukotoxin on human T lymphocytes. Infect. Immun. 59 : 3267-3272.

Marinus MG. 1996. Methylation of DNA. In *Escherichia coli* and *Salmonella* : cellular and molecular biology, 2nd ed., vol. 1. Neidhardt FC, Curtiss III R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M and Umbarger ME ed., ASM Press, Washington. 782-791.

Marinus MG and Morris NR. 1973. Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. J. Bacteriol. 114 : 1143-1150.

Mayer Mp, Bueno LC, Hansen EJ and DiRienzo JM. 1999. Identification of a cytolethal distending toxin gene locus and features of a virulence-associated region in *Actinobacillus actinomycetemcomitans*. Infect. Immun. 67 : 1227-1237.

McShan WM, Tang YF and Ferretti JJ. 1997. Bacteriophage T12 of *Streptococcus pyogenes* integrates into the gene encoding a serine tRNA. Mol. Microbiol. 23 : 719-728.

Meselson M, Yuan R and Heywood J. 1972. Restriction and modification of DNA. Annu. Rev. Biochem. 41 : 447-466.

Meyer DH and Fives-Taylor M. 1997. The role of *Actinobacillus actinomycetemcomitans* in the pathogenesis of periodontal disease. Trends in Microbiol. 5 : 224-228.

Moore WEC and Moore LVH. 1994. The bacteria of periodontal disease. Periodontol. 2000. 5 : 66-77.

Morrison J. 1932. Bacteriophage in the treatment and prevention of cholerae. HK Lewis & Co, London.

Neely MN and Friedman DI. 1998. Functional and genetic analysis of regulatory regions of coliphage H-19B : location of shiga-like toxin and lysis gene suggest a role for phage functions in toxin release. Mol. Microbiol. 28 : 1255-1267.

Nelson D, Loomis L and Fischetti VA. 2001. Prevention and elimination of upper respiratory colonization of mice by group A *streptococci* by using a bacteriophage lytic enzyme. Proc. Natl. Acad. Sci. USA. 98 : 4107-4112.

Nunes-Düby SE, Kwon HJ, Tirumalai RS, Ellenberger T and Landy A. 1998. Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucl. Acid Res. 26 : 391-406.

Obregon V, Garcia JL, Garcia E, Lopez R and Garcia P. 2003. Genome organization and molecular analysis of the temperate bacteriophage MM1 of *Streptococcus pneumoniae*. J. Bacteriol. 185 : 2362-2368.

Olsen I, Namok E and Myhrvold V. 1993. Electron microscopy of phages in serotypes of *Actinobacillus actinomycetemcomitans*. Oral Microbiol. Immunol. 8 : 383-385.

O'Shea YA and Boyd EF. 2002. Mobilization of the Vibrio pathogenicity island between *Vibrio cholerae* isolates mediated by CP-T1 generalized transduction. FEMS Microbiol. Lett. 214:153-157.

Page RC and Schroeder HE. 1976. Pathogenesis of chronic inflammatory periodontal disease : a summary of current work. Lab. Invest. 33 : 235-249.

Page RC and Kornman KS. 1997. The pathogenesis of human periodontitis : an introduction. Periodontol. 2000. 14 : 9-11.

Paju S, Saarela M, Alaluusua S, Fives-Taylor P and Asikainen S. 1998. Characterization of serologically nontypeable *Actinobacillus actinomycetemcomitans* isolates. J. Clin. Microbiol. 36 : 2019-2022.

Palmer BR and Marinus MG. 1994. The *dam* and *dcm* strains of *Escherichia coli*--a review. Gene. 143 : 1-12.

Pargellis CA, Nunes-Düby SE, Moitoso de Vargas L and Landy A. 1988. Suicide recombination substrates yield covalent lambda Int-DNA complexes and lead to the identification of the active site tyrosine. J. Biol. Chem. 263 : 7678-7685.

Peterson KR, Wertman KF, Mount DW and Marinus MG. 1985. Viability of *Escherichia coli* K-12 DNA adenine methylase (*dam*) mutants requires increased expression of specific genes in the SOS regulon. Mol. Gen. Genet. 201 : 14-19.

Petrov NA, Karginov VA, Mikriukov NN, Serpinski OI and Kravchenko VV. 1981. Complete nucleotide sequence of the bacteriophage lambda DNA region containing gene Q and promoter $P_{R'}$. FEBS Lett. 133 : 316-320.

Piekarowicz A, Bujnicki J. 1999. Cloning of the Dam methyltransferase gene from *Haemophilus influenzae* bacteriophage HP1. Acta Microbiol. Pol. 48 : 123-129.

Plunkett G III, Rose DJ, Durfee TJ and Blattner FR. 1999. Sequence of shiga toxin 2 phage 933W from *Escherichia coli* O157:H7 : shiga toxin as a phage late-gene product. J. Bacteriol. 181 : 1767-1778.

Preus HR, Olsen I and Namok E. 1987a. Association between bacteriophage-infected *Actinobacillus actinomycetemcomitans* and rapid periodontal destruction. J. Clin. Periodontol. 14 : 245-247.

Preus HR, Olsen I and Gjermo P. 1987b. Bacteriophage infection-a possible mechanism for increased virulence of bacteria associated with rapidly destructive periodontitis. Acta Odontol. Scand. 45 : 49-54.

Ptashne M. 1992. A genetic switch. Cell Press and Blackwell Scientific Publications, Cambridge, Mass.

Ranney RR, Yanni NR, Burmeister JA and Tew JG. 1982. Relationship between attachment loss and precipitating serum antibody to *Actinobacillus actinomycetemcomitans* in adolescents and young adults having severe periodontal destruction. J. Periodontol. 53 : 1-7.

Roberts D, Hoopes BC, McClure WR and Kleckner N. 1985. IS*10* transposition is regulated by DNA adenine methylation. Cell. 43 : 117-130.

Roberts RJ and Xu SY. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucl. Acids Res. 31 : 1805-1812.

Rosan B, Slots J, Lamont RJ, Listgarten MA and Nelson GM. 1988. *Actinobacillus actinomycetemcomitans* fimbriae. Oral Microbiol. Immunol. 3 : 58-63.

Ruska H. 1940. Uber die Sichtbarmachung der bakteriophagen Lyse im mikroskop. Naturwissenschaften. 28 : 45-46.

Sambrook J, Fritsch EF and Maniatis T. 1989. Molecular cloning : a laboratory manual. Second Edition. Cold Spring Harbor Press.

Sandmeier H, van Winkelhoff AJ, Bär K, Ankli E, Maeder E and Meyer J. 1995. Temperate bacteriophages are common among *Actinobacillus actinomycetemcomitans* isolates from periodontal pockets. J. Periodont. Res. 30 : 418-425.

Sauer RT, Nelson HC, Hehir K, Hecht MH, Gimble FS, DeAnda J and Poteete AR. 1983. The lambda and P22 phage repressors. J. Biomol. Struct. Dyn. 1 : 1011-1022.

Schlesinger M. 1936. The feulgen reaction of the bacteriophage substance. Nature. 138 : 508-509.

Schluckebier G, O'Gara M, Saenger W and Cheng X. 1995. Universal catalytic domain structure of AdoMet-dependent methyltransferases. J. Mol. Biol. 247 : 16-20.

Schuch R, Nelson D and Fischetti VA. 2002. A bacteriolytic agent that detects and kills *Bacillus anthracis*. Nature. 418 : 884-889.

Shimizu-Kadota M, Kiwaki M, Sawaki S, Shirasawa Y, Shibahara-Sone H and Sako T. 2000. Insertion of bacteriophage ϕFSW into the chromosome of *Lactobacillus casei* strain Shirota (S-1): characterization of the attachment sites and the integrase gene. Gene. 249 : 127-134.

Simpson AJG and Setubal JC. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature. 13 : 151-157.

Slots J, Reynolds HS and Genco RJ. 1980. *Actinobacillus actinomycetemcomitans* in human periodontal disease : a cross-sectional microbiological investigation. Infect. Immun. 29 : 1013-1020.

Slots J. 1982. Salient biochemical characters of *Actinobacillus actinomycetemcomitans*. Arch. Microbiol. 131 : 60-67.

Slots J and Rosling BG. 1983. Suppression of the periodontopathogenic microflora in localized juvenile periodontitis by systemic tetracycline. J. Clin. Periodontol. 10 : 465.

Socransky SS. 1979. Criteria for the infectious agents in dental caries and periodontal disease. J. Clin. Periodontol.6 : 16-21.

Socransky SS and Haffajee AD.1994. Evidence of bacterial etiology: a historical perspective. Periodontol. 2000. 5 : 7-25.

Sreenivasan PK, LeBlanc DJ, Lee LN, Fives-Taylor P. 1991. Transformation of *Actinobacillus actinomycetemcomitans* by electroporation, utilizing constructed shuttle plasmids. Infect. Immun. 59 : 4621-4627.

Sternberg N and Coulby J. 1990. Cleavage of the bacteriophage P1 packaging site (*pac*) is regulated by adenine methylation. Proc. Natl. Acad. Sci. USA. 87 : 8070-8074.

Stevens RH, Goncharoff p, Furgang D, Fine DH, Schreiner HC and Firgurski DH. 1993. Characterization and physical mapping of the genome of bacteriophage phiaa from *Actinobacillus actinomycetemcomitans*. Oral Microbiol. Immunol. 8 : 100-104.

Stevens RH, Hammond BF and Lai CH. 1982. Characterization of an inducible bacteriophage from a leukotoxic strain of *Actinobacillus actinomycetemcomitans*. Infect. Immun. 35 : 343-349.

Tarkowski TA, Mooney D, Thomason LC and Stahl FW. 2002. Gene products encoded in the ninR region of phage lambda participate in Red-mediated recombination. Genes Cells. 7 : 351-63.

Timinskas A, Butkus V, and Janulaitis A. 1995. Sequence motifs characteristic for DNA [cytosine-N4] and DNA [adenine-N6] methyltransferases. Classification of all DNA methyltransferases. Gene 157 : 3-11.

Tinoco EMB, Stevens RH, Haubek D, Lai CH, Balachandran S and Preus H. 1997. Relationship of serotype, leukotoxin gene type and lysogeny in *Actinobacillus actinomycetemcomitans* to periodontal disease status. Eur. J. Oral Sci. 105 : 310-317.

Tran LS, Szabo L, Ponyi T, Orosz L, Sik T and Holczinger A. 1999. Phage abortive infection of *Bacillus licheniformis* ATCC9800; identification of the *abi*BL11 gene and localisation and sequencing of its promoter region. Appl. Microbiol. Biotechnol. 52 : 845-852.

Twort FW. 1915. An investigation on the nature of the ultra-microscopic viruses. Lancet. : 1241-1243.

van der Woude M, Braaten B and Low D. 1996. Epigenetic phase variation of the *pap* operon in *Escherichia coli*. Trends Microbiol. 4 : 5-9.

van Mellaert L, Mei L, Lammertyn E, Schacht S and Anné J. 1998. Site-specific integration of bacteriophage VWB genome into *Streptomyces venezuela* and construction of a VWB-based integrative vector. Microbiol. 144 : 3351-3358.

van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR and Wickner RB. 2000. Virus Taxonomy :

The Classification and Nomenclature of Viruses. The Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, 1167pp.

van Steenbergen TJM, Bosch-Tijhof CJ, van Winkelhoff AJ, Gmür R and de Graaff J. 1994. Comparison of six typing methods for *Actinobacillus actinomycetemcomitans*. J. Clin. Microbiol. 32 : 2769-2774.

van Winkelhoff and de Graaff J. 1991. Microbiology in the management of destructive periodontal disease. J. Clin. Periodontol. 18 : 406-410.

van Winkelhoff AJ and Slots J. 1999. *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis* in non-oral infections. Periodontol. 2000. 20 : 122-135.

Waksman G, Lanka E and Carazo JM. 2000. Helicases as nucleic acid unwinding machines. Nat. Struct. Biol. 7 : 20-22.

Waldor MK and Mekalanos JJ. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science. 272 : 1910-1914.

Werner E, Holder AA and Hoheisel JD. 1997. Growth and storage of YAC clones in Hogness Freezing Medium. Nucl. Acids Res. 25 : 1467-1468.

Willi K, Sandmeier H and Meyer J. 1993. Temperate bacteriophages of *Actinobacillus actinomycetemcomitans* associated with periodontal disease are genetically related. Med. Microbiol. Lett. 2 : 419-426.

Willi K, Sandmeier H, Asikainen S, Saarela M and Meyer J. 1997a. Occurrence of temperate bacteriophages in different *Actinobacillus actinomycetemcomitans* serotypes isolated from periodontally healthy individuals. Oral Microbiol. Immunol. 12 : 40-46.

Willi K, Sandmeier H, Kulik EM and Meyer J. 1997b. Transduction of antibiotic resistance markers among *Actinobacillus actinomycetemcomitans* strains by temperate bacteriophage AaΦ23. Cell. Mol. Life Sci. 53 : 904-910.

Willi K and Meyer J. 1998. DNA analysis of temperate bacteriophage AaΦ23 isolated from *Actinobacillus actinomycetemcomitans*. Mol. Gen. Genet. 258 : 323-325.

Wilson MJ, Weightman AJ and Wade WG. 1997. Application of molecular ecology in the characterization of uncultured microorganisms associated with human disease. Rev. Med. Microbiol. 8 : 91-101.

Wommack KE and Colwell RR. 2000. Virioplankton : viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. 64 : 69-114.

World Workshop in Periodontics 1996. 1996. Periodontal diseases : pathogenesis and microbial factors. Consensus report. Ann. Periodontol. 1 : 926-932.

Young R, Wang IN and Roof WD. 2000. Phages will out : strategies of host cell lysis. Trends in Microbiol. 8 : 120-128.

Zambon JJ, Slots J and Genco RJ. 1983. Serology of oral *Actinobacillus actinomycetemcomitans* and serotype distribution in human periodontal disease. Infect. Immun. 41 : 19-27.

Zambon JJ. 1985. *Actinobacillus actinomycetemcomitans* in human periodontal disease. J. Clin. Periodontol. 12 : 1-20.

Zambon JJ, Umemoto T, De Nardin E, Nakazawa F, Christersson LA and Genco RJ. 1988. *Actinobacillus actinomycetemcomitans* in the pathogenesis of human periodontal disease. Adv. Dent. Res. 2 : 269-274.

Zambon JJ, Sunday GJ and Smutko JS. 1990. Molecular genetic analysis of *Actinobacillus actinomycetemcomitans* epidemiology. J. Periodontol. 61 : 75- 80.

Zylicz M, Gorska I, Taylor K and Georgopoulos C. 1984. Bacteriophage lambda replication proteins : formation of a mixed oligomer and binding to the origin of replication. Mol. Gen. Genet. 196 : 401-406.

LIST OF FIGURES AND TABLES

Curriculum vitae

Name:	Resch
Vorname:	Grégory
Geburtsort:	Mulhouse
Geburtsdatum:	10.11.1972
Heimatort:	Wittelsheim
Mutter:	Ginette Yvonne Steiger
Vater:	Francis Resch

Ausbildung:

10/1978-06/1983: Ecole Primaire du Centre (Wittelsheim, France)

10/1983-06/1987: Collège Jean Mermoz (Wittelsheim, France)

06/1987: Brevet des Collèges

10/1987-06/1990: Lycée Laurent Lavoisier (Mulhouse, France)

06/1990: Baccalauréat série D

10/1990-06/1993: DEUG Biologie, Université Louis Pasteur (Strasbourg, France)

10/1993-06/1994: Licence de Biologie, option Immunologie

10/1994-06/1995: Maîtrise de Biochimie, option Immunologie approfondie

10/1995-06/1996: DEA de Pharmacochimie Moléculaire, Université Paris V (Paris, France)

10/1996-10/1997: Scientifique du Contingent (Brétigny-sur-orge, France)

11/1997-12/1998: PhD en Biologie Moléculaire (Pr. K. Geering, Lausanne, Suisse)

01/1999-06/2003: PhD en Microbiologie (Pr. J. Meyer. Bâle, Suisse)

An meiner Universitären waren folgende DozentInnen beteiligt:

Pr. Charles Thompson, Pr. Peter Philippsen, Pr. Werner Arber, Pr. Mark Borodovsky, Dr. Frederic Dietrich, Dr. Sophie Lemire-Brachat, Dr. Mogens Kilian, Dr. David Dyer, Dr. Eric Kofoid, Dr. Ryan Mills, Dr. Jeffrey Lawrence, Dr. Cathy Weber, Dr. Caroline Baumann, Dr. Tuomas Waltimo, Dr. Gilles Kolb, Dr. Frauke Berres, Dr. Daniel Collet, Dr. Karin Willi, Dr. Marc Weber.