

**Biosynthese von *N*-Acylglutaminen im Insektendarm:
Beteiligung und Charakterisierung bakterieller Enzyme**

Dissertation

zur Erlangung des akademischen Grades doktor rerum naturalium
(Dr. Rer. Nat.)
vorgelegt dem Rat der Biologisch-Pharmazeutischen Fakultät
der Friedrich-Schiller-Universität Jena

vorgelegt von M. S.
Liyan Ping

geboren am August 21, 1972 in Hebei, VR China

**Biosynthesis of *N*-acyl glutamines in the insect gut: impact
and characterisation of microbial enzymes**

Dissertation

Faculty for Biology and Pharmacy
Friedrich-Schiller-University in Jena

M. S.
Liyan Ping

Date of birth: August 21, 1972 in Hebei, P.R. China

Gutachter:

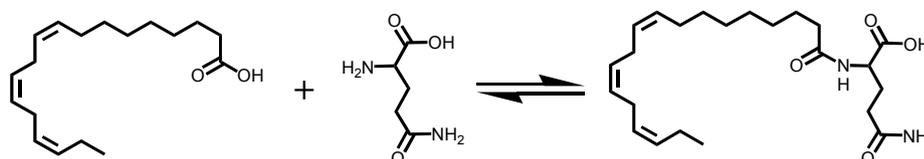
1. Prof. Dr. W. Boland. FRSC
2. Prof. Dr. E. Kothe
3. Prof. Dr. G. Diekert
4. Prof. Dr. J. Wöstemeyer
5. Prof. Dr. A. Brakhage

Supervisors:

1. Prof. Dr. Wilhelm Boland
2. Prof. Dr. Erika Kothe
3. Prof. Dr. Joern Piel

Summary

N-acyl amino acids were first isolated and identified from insect gut as elicitors of the host plant's defence reactions (Alborn *et al.* 1997). Up to date, no experimental proof for a role of these compounds for the physiology of insects has been found. Due to their amphiphilic nature, a reasonable function as bioemulsifiers has been postulated (Spiteller 2002). Accumulating data indicate that the compounds are secreted from a larger pool in the insect's digestive tract. The homeostasis is maintained by a complex cooperation of genuine enzymes of the insect and from commensal gut bacteria. In this work, the mechanism of synthesis/hydrolysis of the amide bond of *N*-acyl amino acids by gut bacteria was studied. The following reaction was used to screen isolated gut bacteria from phylogenetically diverse groups for their ability to synthesize or hydrolyse *N*-acyl amino acids:



The reaction is endergonic. It can not effectively proceed without activation. However, protein fractions from the gut bacteria catalyse this reversible reaction to a certain extent into the synthetic direction. Therefore the active protein was isolated from the most productive strain by an activity-guided fractionation (synthesis). Finally the corresponding gene was cloned and expressed.

Many of the bacterial isolates from the gut of *Spodoptera exigua* were able to synthesize *N*-acyl amino acids by using externally supplied free fatty acids and amino acids as substrates. The enzymatic activity was found to increase at the onset of stationary phase growth.

Compared with the most active gut strain *Microbacterium arborescens* SE14, the reference strain *Microbacterium arborescens* DSM20754 showed very low activity. Furthermore *M. arborescens* SE14 grows faster and reaches a higher cell density in the stationary phase than *M. arborescens* DSM20754. This might be consistent with an adaptation to the conditions in the insect gut lumen (Dillon and Dillon 2004).

An active protein fraction has been isolated from *M. arborescens* SE14. After several steps of purification the specific activity of the protein was raised 186 fold. The active protein, named Afp (Amide-forming-protein), is composed of 12 subunits of 17 kD. According to its catalytic constants, Afp preferentially catalyses the hydrolysis of amide bonds. On the other hand, it can catalyse to a certain extent the reverse reaction, namely the synthesis of *N*-acyl amino acids, if high local concentration of substrates are present. No activation by CoA and ATP is required. The Afp protein catalyses amide formation at pH 8.0 and pH 9.2-12.0, two pH ranges found in the caterpillar stomach and foregut. The temperature optimum of the reaction is about 40 °C. Its thermostability is very high. It can tolerate temperatures up to 48 °C without significant loss of the activity. The spectrum of incorporated amino acids is very broad (amide synthesis). Almost all proteingenous amino acid, except proline, could be used as substrate. This is more coincident with the spectrum of a hydrolytic enzyme. The

spectrum of accepted fatty acids is also broad. The incorporation efficiency of both types of substrates depends on their hydrophobicity.

The mechanism for hydrolysis/biosynthesis is not fully understood. Common proteinase inhibitors could not mock the active centre. According to the UV absorption spectra, a pH-dependent deprotonation of tyrosine residue(s) might be involved in the catalysis. Furthermore, a posttranslational modification, probably an acetylation has been proposed from the MS spectra of the isolated protein.

The cloned *afp* gene is a member of the family of Dps proteins (DNA-binding proteins in starved cell), but represents an evolutionarily distinctive line to other Dps-like proteins identified in eubacteria. Amide-formation and amide-hydrolysis is the first enzymatic activity detected in this protein family. On the other hand, the typical ferroxidase center is conserved in this protein like in other Dps-type proteins. Interestingly, the native protein contains 10 iron atoms per subunit, which is the highest level so far detected in this type of proteins.

The *afp* gene has been overexpressed in *E. coli*. It forms the typical Dps-like hexamer on native PAGE. But unfortunately only a trace activity could be observed when ferric ion was supplemented externally. The *afp* gene has been successfully transformed into *M. arborescens* DSM20754 under the control of the *veg* promoter from *Bacillus subtilis*. The catalytic activity of *M. arborescens* DSM20754 has been significantly increased.

From a biotechnological point of view, a new PCR approach, namely SD-PCR has been developed in this study. The approach can be applied in prokaryotic gene cloning, *esp.* when the DNA template was contaminated by eukaryotic DNA. Furthermore, the protocol of transformation of *M. arborescens* should allow the expression of diverse actinomycete genes in this amenable and fast-growing host. This might eventually lead to the development of a potential gene cloning and expression system for actinomycete genes, which are commercially and clinically important.

Contents

Summary	I
Contents	III
1 Introduction	1
1.1 <i>N</i>-acylamino acid from insect regurgitant	2
1.1.1 Structural diversity of <i>N</i> -acylamino acids	-3
1.1.2 Physiological function of <i>N</i> -acylamino acids	4
1.1.3 Biosynthesis of <i>N</i> -acylamino acids	-5
1.2 Bacterial lipoamino acids and related compounds	6
1.2.1 Lipoamino acids	6
1.2.1.1 Structure of lipoamino acids	9
1.2.1.2 Biosynthesis of lipoamino acids	10
1.2.2 Lipopeptides	-11
1.2.2.1 Classification of lipopeptides	-11
1.2.2.2 Biosynthesis of lipopeptides	13
1.2.3 Lipoproteins	13
1.2.3.1 Secreted proteins	-14
1.2.3.2 Cell wall proteins	14
1.2.3.3 <i>N</i> -Acylation of the prolipoprotein	15
1.3 Bacterial amidation pathways	15
1.3.1 Amide synthase	15
1.3.2 <i>N</i> -Acyltransferases	17
1.3.3 Amidases and proteases	18
1.3.3.1 Proteases and their catalytic mechanisms	18
1.3.3.2 Amidases and their active centers	19
1.3.3.3 Acyl transfer activity in amidases	19
1.4 Goals of this research	20
2 Materials and Methods	21
2.1 Materials	21

Contents

2.1.1	Plasmids	21
2.1.2	Microbial sources and properties	22
2.1.3	Chemicals	22
2.2	Methods	22
2.2.1	Isolation and identification of microorganism	22
	Monitoring of bacteria by FISH	23
2.2.2	Enzyme activity assays	26
2.2.3	Protein purification	26
	Active protein purification	27
2.2.4	Protein characterization	28
	Western blotting	29
2.2.5	Protein sequencing	32
2.2.6	Genomic DNA isolation	32
	Genomic DNA isolation from <i>M. arborescens</i> strains	32
2.2.7	Gene cloning	34
	Detection of restriction-fragments by radioactive probe	34
	<i>E. coli</i> library screening	36
	Plasmid extraction by MINIpip	38
2.2.8	Gene expression	39
	Protocol for gene expression in <i>E. coli</i>	41
	Transformation of <i>M. arborescens</i> DSM20754	42
3	Results	45
3.1	Screening of the gut bacteria	45
3.1.1	Catalytic activities	45
3.1.2	Growth stage-dependent activities	46
3.2	Evidence of selectively enhanced activity	47
3.3	Purification of the amide-forming protein	49
3.4	Characterization of the Afp protein	53
3.4.1	Determination of the molecular mass	53
3.4.2	Determination of the isoelectric point	57
3.4.3	Iron content	58
3.4.4	DNA-binding assay	59

3.5 The catalytic properties of Afp	59
3.5.1 A rapid reversible reaction	59
3.5.2 The temperature optimum	60
3.5.3 The pH optima of Afp	60
3.5.4 Characterization of the active centre	61
3.6 The substrate specificity	63
3.6.1 Specificity on amino acid	64
3.6.2 Fatty acid selectivity	65
3.7 Cloning the Afp-coding gene	66
3.7.1 Protein sequencing	66
3.7.2 Nested degenerate primer PCR	68
3.7.2.1 Cloning the 5'-sequence of <i>afp</i> gene	69
3.7.2.2 SD-PCR	69
3.7.2.3 3rd round PCR with degenerate primer	70
3.7.3 Cloning the entire <i>afp</i> gene	71
3.7.4 Analysis of <i>afp</i> gene	71
3.8 Expression of the <i>afp</i> gene	74
3.8.1 Heterologous expression in <i>E. coli</i>	74
3.8.2 Expression in <i>M. arborescens</i> DSM20754	76
3.8.2.1 The <i>afp</i> gene ortholog in <i>M. arborescens</i> DSM20754	76
3.8.2.2 Plasmids construction	77
3.8.2.3 Transformation of <i>M. arborescens</i> DSM20754	78
3.8.2.4 Increase of <i>N</i> -acyltransferase activity	78
4 Discussion	81
4.1 Screening the commensal gut bacteria	81
4.1.1 Different catalytic properties	81
4.1.2 Growth stage-dependent activities	82
4.1.3 Selection of the most active strain	82
4.2 Evidence of selectively enhanced activity	83
4.3 Purification of the amide-forming protein	84
4.4 Characterization of the Afp protein	85
4.5 The catalytic properties of Afp	86

Contents

4.5.1	A rapid reversible reaction	-86
4.5.2	The temperature optimum	-86
4.5.3	The pH influence on Afp activity	-86
4.5.4	Characterization of the active centre	-87
4.6	The substrate specificity	-87
4.6.1	Specificity on amino acid	-87
4.6.2	Fatty acid selectivity	-88
4.7	Cloning the corresponding gene	-88
4.7.1	Protein sequencing	-88
4.7.2	Nested degenerate primer PCR	-89
4.7.3	Cloning the entire <i>afp</i> gene	-90
4.7.4	Analysis of <i>afp</i> gene	-90
4.7.4.1	DNA binding activity	-90
4.7.4.2	Stable structure	-91
4.7.4.3	Different aggregation states	-92
4.7.4.4	Iron-binding activity	-93
4.7.4.5	An unknown catalytic mechanism	-94
4.7.4.6	Evolution from a distinct line	-94
4.8	Expression of the <i>afp</i> gene	-95
4.8.1	Overexpression in <i>E. coli</i>	-95
4.8.2	Expression in <i>M. arborescens</i> DSM20754	-96
4.8.2.1	The <i>afp</i> gene ortholog in <i>M. arborescens</i> DSM20754	-96
4.8.2.2	Construction of expression plasmid	-97
4.8.2.3	Transformation of <i>M. arborescens</i> DSM20754	-97
4.8.2.4	Increase of <i>N</i> -acyltransferase activity	-98
4.9	The ecological role of Afp	-98
4.10	Perspectives	-100
Acknowledgements		-103
Reference		-105

Appendix I

(Maps of the plasmids constructed in this research)

Appendix II

(Resistance of *M. arborescens* strains against selected antibiotics)

Appendix III

(Instruments)

Appendix IV

(Softwares)

Appendix V

(Primers and probes)

Appendix VI

(Abbreviations)

1 Introduction

The world teems with bacteria. They can live in different environments due to their adaptability. The existence of microscopic organisms and their impact on ecosystems has been always beyond the notice of humans until recent advance of science enabled us to take a closer look into this fascinating field.

One of the questions, that have preoccupied ecologists during the last century, is to address the factors controlling the population of terrestrial herbivorous arthropods, the standard theory of which is the supply of host plant resources (bottom-up effects) *vs* the control of natural enemies (top-down effects) (Figure 1-1) as suggested by Hairston, Smith, and Slobodkin (Hairston *et al.* 1960). This HSS model has been criticized for improper consideration on edibility of defended plants and the ill-defined nature of trophic levels. Likewise, pathogens were simply not considered (Rosenheim 1998). In fact, the role of bacteria in this scenario is much more than solely as pathogen, *e.g.* *Wolbachia* infection caused female-dominated leks in invertebrates (Zimmer 2001) and the aphid fitness and resistance to parasitic wasps enhanced by bacterial symbionts (Oliver *et al.* 2003) definitely add new chapters on the top-down effects.

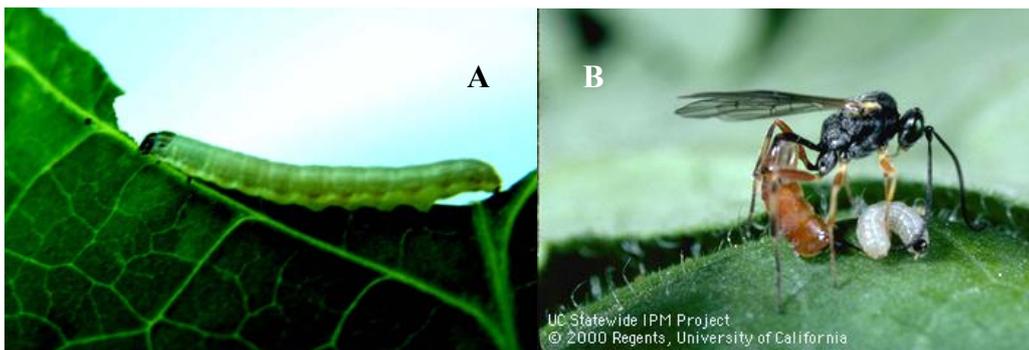


FIGURE 1-1 Factors controlling the population of a cosmopolitan herbivorous insect, beet armyworm (*Spodoptera exigua*) as proposed by HSS model. (A) A beet armyworm larva is feeding on lima bean (*Phaseolus lunatus*) leaf. Lima bean contains cyanogenic glucosides, which break down to toxic hydrocyanic acid (HCN) when chewed. (B) A parasitic wasp (*Hyposoter exiguae*) female is laying an egg in a beet armyworm larva. The parasitic wasps detect their prey by following the volatiles emitted from herbivore-damaged plant. (B) photo by Jack Kelly Clark, reproduced with permission of the University of California Statewide Integrated Pest Management Program.

Plant defence consists of a complex array of responses, classified as direct and indirect defences. Direct defence, such as trichomes and phytoalexins, either constitutive or inducible, plays its role in bottom-up effects. Indirect defences are plant traits that attract predators and parasitoids of herbivores and increase the carnivore's foraging success and thereby facilitate top-down control of herbivore populations (Kessler and Baldwin 2002).

The feeding-inducible plant volatiles are the key signals mediating indirect defence. Under herbivore attack, plants emit volatiles serving as guiding cues of cruising natural

predators of the attacking herbivores (Arimura *et al.* 2000). A family of plant volatile elicitors, *N*-acyl amino acids, has been identified in insect oral secretion (Alborn *et al.* 1997). The amphiphilic molecules are structurally related to lipoamino acids of microbial origin (Figure 1-2) (Boulton 1989). This similarity stimulated the research on the role of insect gut bacteria in the biosynthesis of these plant volatile elicitors. Fossil evidence indicates that plant volatiles, such as terpenoids, are well known repellents for microorganisms, evolved as early as in Paleozoic, long before the evolution of terrestrial insects. They were adopted by organisms of higher trophic levels as semiochemicals during the long arms race of co-evolution (Lamber *et al.* 1998). Logically, a reaction employed in avoidance of harmful microorganism invasion would be induced by some factors of microorganism-related origin.

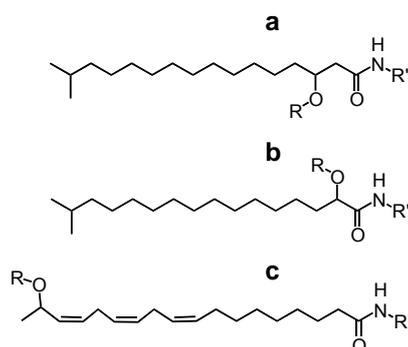


FIGURE 1-2 Structural comparison of the typical lipoamino acids from bacteria (a and b) and the plant volatile elicitor from insect oral secretion (c). In a bacterial conjugate the fatty acid moiety usually carries a hydroxyl group at C-3 (a) or C-2 (b), R is either a hydrogen atom or another fatty acid. R' is the side chain of the amino acids, ornithine, serine, lysine, *etc.* In the larval regurgitant (c), the fatty acid is either hydroxylated at C(17) or non-oxidized. R is either a hydrogen atom, another fatty acid, or a phosphate. R' is the side chain of glutamine or glutamic acid.

The physiological function of *N*-acyl amino acids in caterpillars is still an enigma. The adaptation of bacterial metabolites by insects to exploit novel physiological functions during evolution had been proved in several cases. Bacterial biota in locust gut produces guaiacol (2-methoxyphenol), a key component of a pheromone derived from locust faecal pellets that promotes the aggregation of locusts (Dillon *et al.* 2000). A defence compound, pederin, employed by female *Paederus* beetles, has been demonstrated to be a metabolite of an uncultured bacterial symbiont (Piel 2002). One may speculate that *N*-acyl amino acids were adapted by the plant host as a signal for danger during the long arm race with insect herbivore, no matter whether this compound had once been of benefit to the insect or not.

1.1 *N*-Acylamino acids from insect regurgitant

The feeding of an insect herbivore on plant foliage consists of two separate processes, one is the mechanical damage caused by chewing of the jaws. The open wound resulting from this action is then exposed to chemicals from the insect oral secretion. The second factor, namely the chemicals in the secretion, plays an important role in the induction of volatile emission, quantitatively and qualitatively distinguishable

from an undamaged control plant. The induced volatiles attract natural enemies of the feeding insect (Turlings *et al.* 1990; Kessler and Baldwin 2002). This larval regurgitant is composed of secretion from the saliva gland and enteric fluids regurgitated from the first part of digestion tract.

Several low- and high-molecular weight components have been identified in these insect oral secretions. Some of them are reported as potent plant volatile elicitors. Among the macromolecules are proteinaceous lytic enzymes, such as the β -glucosidase from *Pieris brassicae*, which elicits the release of terpenoid volatiles from cabbage leaves (Mattiacci *et al.* 1995). The family of *N*-acyl amino acids is a representative for those low molecular elicitors. They are fatty acids linked by an amide bond to an amino acid, usually glutamine, in other cases glutamic acid. These small molecules are ubiquitously present in the regurgitants of herbivorous insects (Pohnert *et al.* 1999). These compounds have been found in the regurgitant of larvae of the Lepidopteran families, Sphingidae, Noctuidae, and Geometridae (Kessler and Baldwin 2002).

1.1.1 Structural diversity of *N*-acylamino acids

The first reported *N*-acyl amino acid was isolated from beet armyworm (BAW, *Spodoptera exigua*) by a bioassay-guided fractionation process. This *N*-(17-hydroxy-linolenoyl)-glutamine was named volicitin, because it was found to be able to trigger volatile emission from corn plants (*Zea mais*) (Alborn *et al.* 1997). The chiral centre of the hydroxylated linolenic acid moiety of volicitin was determined later to be more than 94% (*S*)-configured (Spiteller *et al.* 2001).

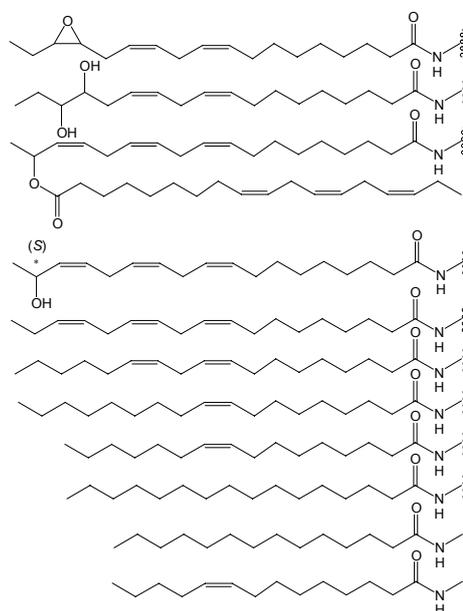


FIGURE 1-3 Fatty acid moieties in *N*-acyl amino acids from lepidopteran larvae. The amide bond is shown, the side chain of glutamine and glutamic acid is omitted. Fatty acids with fewer double bonds than those shown in the picture have also been detected.

Not only the 17-hydroxyl fatty acid conjugates but also the nonhydroxylated fatty acid conjugates were found in the same species (Pohnert *et al.* 1999; Alborn *et al.* 2000). Without the hydroxyl group, the conjugates are still active with respect to volatile induction (Alborn *et al.* 2000). The nonoxidized fatty acid conjugates are more abundant in different insect species (Pohnert *et al.* 1999). The fatty acids with 14-, 16-, and 18-carbon atoms with different degrees of unsaturation (Figure 1-3), are generally detected in total lipid extracts of insects, and they all are found in conjugation with amino acid moieties (Figure 1-3) in the lepidoptera larval oral secretion (Pohnert *et al.* 1999).

In the oral secretions of some insects, for example the tobacco hornworm (*Manduca sexta*), the dominant conjugates are fatty acid conjugates with glutamate instead of glutamine (Pohnert *et al.* 1999; Halitschke *et al.* 2001). When this glutamic acid conjugates were applied to tobacco (*Nicotiana attenuata*), defence reactions identical to that induced by the insect feeding were observed (Halitschke *et al.* 2001).

The fatty acid moiety is further modified by oxidative transformations to secondary alcohols and epoxides (Spiteller and Boland 2003). There are more complex products when the functionalized fatty acid moieties are hydrolysed (epoxides) or acylated (alcohols). The acylation of volicitin-type precursors results in very unpolar and rather unstable compounds such as *N*-(17-linolenoxy-linolenoyl)-glutamine (Spiteller and Boland 2003).

1.1.2 Physiological function of *N*-acyl amino acids

The typical *N*-acyl amino acids are amphiphilic compounds which can reduce surface tension of aqueous solutions (Spiteller 2002). Therefore they were assumed to act as biosurfactants to solubilize nutrients from the diet (Figure 1-4).

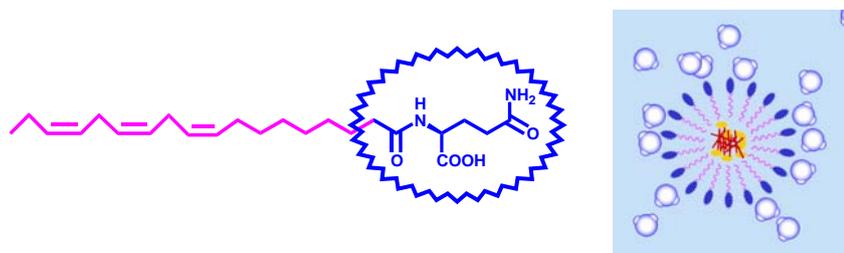


FIGURE 1-4 The commonly detected *N*-acyl amino acid, *N*-linolenyl glutamine has an unpolar tail (purple) and a polar head (blue). In aqueous systems it tends to form micelles (right) to disperse hydrophobic materials (orange).

There is evidence that volicitin is an active plant defence activator. Indole is a major component of the feeding-induced volatiles in maize. Indole-3-glycerol phosphate lyase *Igl* is the structural gene that catalyses its biosynthesis. The transcription of this gene could be activated by treatment of both volicitin and *N*-linolenoyl-L-glutamine. Only the L-stereoisomers show this biological activity (Frey *et al.* 2004). Volicitin treatment also leads to a fast transcription of plant genes involved in defence reaction, for example allene oxide synthase and allene oxide cyclase (Frey *et al.* 2004). Recently, the plasma

membrane fractions isolated from maize leaves had been demonstrated to be able to bind volicitin, suggesting the existence of a specific protein receptor of *N*-acyl amino acids in plant cell (Truitt *et al.* 2004).

On the other hand, ¹⁴C labelled volicitin was not readily transported to undamaged leaves. The insect-produced compounds may not directly serve as a mobile messenger *in planta* triggering the emissions of volatiles systemically (Truitt and Pare 2004). New data show that this kind of compounds only elicits volatile release from intact maize seedlings when applied at midnight and that the release from excised seedlings was much greater than that from intact plants (Schmelz *et al.* 2001). Furthermore, none of the separated enantiomers of volicitin were active in lima bean which is also fed by the BAW (Spiteller *et al.* 2001). The non-functionalized fatty acid amides were found more active than volicitin on induction of expression of volatile biothynthesis gene (Frey *et al.* 2004). It seems to be the amphiphilic structure of these molecules responsible for this bioactivity.

It is worthwhile to notice that acylated volicitin exhibits striking structural similarities to the mayolenes, previously isolated from defensive secretions of the lepidopteran *Pieris rapae* (Smedley *et al.* 2002), but until now, no toxic effect had been reported for *N*-acyl amino acids (Figure 1-3).

1.1.3 Biosynthesis of *N*-acyl amino acids

Remarkably, there is no functionalized fatty acid-amino acid amides detected in regurgitant of *M. sexta* and *Manduca quinquemaculata* (Halitschke *et al.* 2001). Besides volicitin and other *N*-acyl glutamines, the oral secretion of BAW also contains free 17-hydroxylinolenic, and 17-hydrolinoleic acid as well as free linolenic and linoleic acid (Alborn *et al.* 2000). Oxidation of the fatty acid and its incorporation into conjugates must not happen simultaneously. Broader investigation indicates that conjugate formation in caterpillars does not require for specific fatty acids from the diet but reflects the pattern of fatty acids commonly found in insect diets (Pohnert *et al.* 1999). Considering the structural diversity of conjugates, they might be assembled by either a family of selective enzymes or by an enzyme with broad substrate tolerance (Pohnert *et al.* 1999).

Paré *et al.* (1998) had reported that the linolenic acid moiety of volicitin was taken from the plant, while the amino acid was synthesized by the insect. The enzyme catalysing the amide bond formation between them has not been revealed so far. Recently, a rapid CoA- and ATP-independent conjugation of free linolenic acid and glutamine had been detected *in vitro* in microsomal fractions of *M. sexta* alimentary tissue homogenates (Lait *et al.* 2003). However the amide bond formation between linolenic acid and glutamine is a highly endergonic reaction that can only occur under special circumstances without previous activation.



$$\Delta G = 269.26 \text{ kJ/mol}$$

An enzyme catalyses the formation of a structurally related compound, acyl-CoA:glycine *N*-acyltransferase had been characterized from a number of mammalian

sources (Merkler *et al.* 1996). In the reaction catalysed by this enzyme, the acyl group is activated *via* a high-energy linkage with coenzyme A (see 1.3.2 for detail). The *N*-acylglycine is made in liver mitochondria and then secreted into the blood. Such fatty acid primary amides are known as mammalian hormones, but are not involved in digestion processes.

A CoA- and ATP-independent pathway had been found also in the mammalian central nervous system to synthesize another hormone, namely arachidonoyl ethanolamide, a cannabinoid (Kruszka and Gross 1994). But this amidation system is highly selective for arachidonic acid and ethanolamine as the substrates. In fact, cannabinoids are composed of fatty acids and ethanolamine, replacement of ethanolamine with any amino acid has never been detected (Reggio 2002). Furthermore, cannabinoid receptor had been demonstrated to be absent in insects (McPartland *et al.* 2001). Therefore the analogs of the two mammalian amidation systems mentioned above might not exist in insects. Whereas, it was well known at least two decades ago, that bacteria produce lipoamino acids (Boulton 1989).

1.2 Bacterial lipoamino acids and related compounds

Microorganisms produce a variety of surface-active agents. They have very different chemical structures and surface properties, each may provide an advantage in a particular ecological niche (Ron and Rosenberg 2001). Many of them are glycolipids or lipopolysaccharides, in which the hydrophilic moiety is linked to the neutral polar head via an *O*-ester or ether bond (Lang 2002). Other biosurfactants contain fatty acids linked to a charged polar head via an amide bond. The structures may be relatively simple, an acyl moiety linked to a single amino acids. In more complex lipids, the fatty acids are linked to a peptide or even a protein (Boulton 1989). These three kinds of molecules bearing an acylamido side chain will be discussed in the following sections.

1.2.1 Lipoamino acids

Bacterial amino-acid-containing lipids are sometimes referred to as aminolipids (Kawai and Yano 1983) or *N*-acyl amino acids (Yagi *et al.* 1997), but lipoamino acids is the name adopted by most of the authors. They are identified in taxonomically diverse species (Table 1-1), but production of lipoamino acids is by no means a common ability shared by all of the bacteria. In screening for ornithine lipids (Figure 1-5A), which is the most commonly found lipoamino acid, only 2 out of 16 species were found to contain this zwitterionic lipid (Thiele *et al.* 1984). These compounds have been found in many Gram-negative bacteria, but not in Gram-positive bacteria, except in the order Actinomycetales, *e.g.* mycobacteria and streptomyces. The content of lipoamino acids varies dramatically, even between species producing these compounds. Lipoamino acids represent 60% of the total extractable cellular lipids in *Flavobacterium meningosepticum* and *F. indologenes* (Kawai *et al.* 1988), while the extractable lipids of *Hyphomicrobium vulgare* only contain 6% lipoamino acid (Batrakov and Nikitin 1996). Production of lipoamino acids depends on environmental conditions. Some *Pseudomonas* spp. also contain large amounts of lipoamino acids (Kawai *et al.* 1988). Usually 15% of the lipid extract from *Pseudomonas fluorescens* NCMB129 consist of ornithine lipids. Under phosphate starvation, these lipids become almost the sole polar lipid component

Introduction

(Minnikin and Abdolrahimzadeh 1974). In all investigated species, except *Deleya marina* (Yagi *et al.* 1997), the biosynthesis of lipoamino acids is inducible by nutrient starvation.

Table 1-1 Fatty acid composition of ornithine lipids from different bacteria.

Bacteria	Fatty acid													Ref.						
	10:0	12:0	13:0	14:0	iso-14:0	15:0	iso-15:0	Anteiso-15:0	16:0	Iso-16:0	16:1	17:0	Iso-17:0		Anteiso-17:0	18:0	18:1(11)	Cycl-19:0	Iso-19:0	20:1(13)
<i>Shewanella putrefaciens</i>																				(Wilkinson 1972)
Ester-linked						○	⊙		⊙	○	⊙	○	⊙		○	⊙				
Amide-2-OH																				
Amide-3-OH				○		○	○		⊙			○	⊙		○				⊙	
<i>Flavobacterium meningosepticum</i>																				(Kawai <i>et al.</i> 1988)
Ester-linked				○		○	●		⊙						○	○				
Amide-2-OH																				
Amide-3-OH	○	○					⊙		○				●							
<i>Streptomyces olivaceus</i>																				(Batrakov and Bergelson 1978)
Ester-linked				○	○			●	○	○										
Amide-2-OH																				
Amide-3-OH							⊙			●			⊙							
<i>S. globisporus</i>																				(Batrakov and Bergelson 1978)
Ester-linked					○	○		●		○										
Amide-2-OH																				
Amide-3-OH							○			●			⊙							
<i>S. aureoviticillatus</i>																				(Batrakov and Bergelson 1978)
Ester-linked								⊙	○	⊙	○				⊙					
Amide-2-OH																				
Amide-3-OH									⊙	⊙		○	⊙	⊙						
<i>S. 660-15</i>																				(Batrakov and Bergelson 1978)
Ester-linked																				
Amide-2-OH								●		⊙			○							
Amide-3-OH							○		⊙	⊙			●							
<i>Paracoccus denitrificans</i>																				(Thiele <i>et al.</i> 1980)
Ester-linked	○	○	○		○				○						○	●	⊙			
Amide-2-OH																				
Amide-3-OH									○						○				●	
<i>Thiobacillus A2</i>																				(Thiele <i>et al.</i> 1984)
Ester-linked									○						○	●	○			
Amide-2-OH																				
Amide-3-OH															⊙				●	
<i>Achromobacter</i> sp.																				(Thiele <i>et al.</i> 1984)
Ester-linked									●		⊙									
Amide-2-OH																				
Amide-3-OH									○						⊙	●			○	

Table 1-1 (continued)

Bacteria	Fatty acid													Ref.							
	10:0	12:0	13:0	14:0	iso-14:0	15:0	iso-15:0	Anteiso-15:0	16:0	Iso-16:0	16:1	17:0	Iso-17:0		Anteiso-17:0	18:0	18:1	Cycl-19:0	Iso-19:0	20:1	
<i>Pseudomonas fluorescens</i>									●						○	⊙					(Kawai <i>et al.</i> 1988)
Ester-linked		○		○		○			●						○	⊙					
Amide-2-OH		○							●												
Amide-3-OH	○	○							●							○					
<i>P. aeruginosa</i>									●						⊙	○					(Kawai <i>et al.</i> 1988)
Ester-linked		○		○		○			●						⊙	○					
Amide-2-OH		○							●												
Amide-3-OH	○	⊙		⊙					⊙			○			○	○			○		
<i>P. stutzeri</i>									●						⊙	⊙					(Kawai <i>et al.</i> 1988)
Ester-linked				○		○			●						⊙	⊙					
Amide-2-OH									●												
Amide-3-OH	○	○		⊙					⊙			⊙				⊙					
<i>P. apacia</i>									●						⊙	⊙					(Kawai <i>et al.</i> 1988)
Ester-linked		○		○		○			⊙						⊙	⊙					
Amide-2-OH		○							●												
Amide-3-OH									●			○			○	○			○		
<i>Bordetella pertussis</i>									●	○					○						(Kawai and Yano 1983)
Ester-linked				○					●	○					○						
Amide-2-OH									●												
Amide-3-OH				⊙		○			●						○						
<i>B. parapertussis</i>									○												(Kawai and Yano 1983)
Ester-linked									○												
Amide-2-OH									○												
Amide-3-OH				⊙		○			●												
<i>B. bronchiseptica</i>									○												(Kawai and Yano 1983)
Ester-linked									○												
Amide-2-OH									○												
Amide-3-OH				⊙		○			●												

The relative contents of fatty acids is shown by symbols: '●' (>50%); '⊙' (10%–50%); '○' (1%–10%). Trace amount of less than 1% were omitted. If several strains of the same bacterial species have been studied, the data are averaged due to the negligibility of differences between strains. Those fatty acids, which had only been detected in a single case with a concentration less than 10%, were not included in this table.

Lipoamino acids seem to be somehow linked to bacterial pathogenicity. Some of these lipoamino acid-producing bacteria are pathogens, e.g. *Brucella* spp. (Thiele *et al.* 1980), *Bordetella* spp. (Kawai and Yano 1983), mycobacteria (Laneelle *et al.* 1990), *Flavobacterium* spp. (Kawai *et al.* 1988) and *Pseudomonas* spp. (Kawai *et al.* 1988). Among the mycobacteria tested, lipoamino acids were detected only in the pathogenic strains, but not in the avirulent strains (Laneelle *et al.* 1990).

1.2.1.1 Structure of lipoamino acids

The first lipoamino acid was identified by Wilkinson (1972) in *Shewanella putrefaciens* (former *Pseudomonas rubescens*). This compound was an ornithine lipid, the most common type of lipoamino acids found in bacteria. All ornithine lipids, known up to date, contain an amide-linked fatty acid at the α -amino group. Three types of basic structures can be distinguished by the position of second ester-linked fatty acid (Figure 1-5). Type I, the most common structure, in which the terminal fatty acid is linked to the hydroxyl group of the central fatty acid; type II, in which the fatty acid is linked via a connecting diol to the carboxyl group of ornithine; type III, the fatty acid linked via a hydroxyl group on the second fatty acid to the carboxyl group of ornithine. Further modification on the ornithine moiety has also been detected, for example, a taurine that is ester-linked to the carboxy group of ornithine (Thiele *et al.* 1980). Among them, the type I is the most abundant and most widely distributed ornithine lipid (Batrakov and Bergelson 1978).

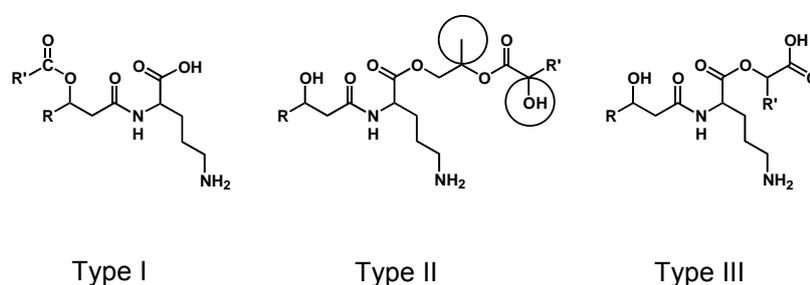


FIGURE 1-5 Three types of ornithine-containing lipids identified. R and R' are the side chains of the amide- and ester-linked fatty acid moieties, respectively. The circles on type II ornithine lipids indicate that the methyl group and hydroxyl group may be replaced by hydrogen atoms in some cases.

In later studies lipoamino acids containing lysine (Boulton 1989), serine (Kawai *et al.* 1988), glycine (Kawazoe *et al.* 1991), leucine and isoleucine (Yagi *et al.* 1997) were found in different bacterial species (Figure 1-6). The basic structure of these lipoamino acids corresponds to the type I ornithine lipids. The amino acid forms an amide bond with the carboxyl group of the first fatty acid, the hydroxyl group of which is linked to the second fatty acid via an ester bond. The first fatty acids were mainly hydroxylated at C-3 (Figure 1-5, 1-6), but 2-hydroxy fatty acids are also quite commonly found (Table 1-1). Besides the typical lipoamino acids mentioned above, in some cases the first acyl substituent is not hydroxylated, and therefore the whole molecule contains only a single fatty acid (Kawai and Yano 1983; Yagi *et al.* 1997).

Many different saturated or unsaturated fatty acids have been found as the acyl moiety of lipoamino acids (Table 1-1). The difference of fatty acid composition between species coincides with their phylogenetic distance. No qualitative difference of fatty acid substitution was observed between strains of same species (Kawai and Yano 1983; Kawai *et al.* 1988). Iso- and anteiso-branched fatty acids predominate in the extract of Gram-positive bacteria. A straight-chain saturated fatty acid, namely hexadecanoic acid, is the major component in Gram-negative bacterial extracts. Straight-chain monounsaturated fatty acids, among which the ω -7 series predominates, were found frequently in Gram-negative species. This coincides with the well-known theory of the elongation pathway of monoenoic fatty acid in bacteria (Thiele *et al.* 1984), while the

occurrence of C_{16:1(9)}, does not fit into this scheme (Thiele *et al.* 1984). A cyclopropane fatty acid was only found in certain strains. This pattern is just reminiscent the free fatty acids distribution among microbial species (Schweizer 1989).

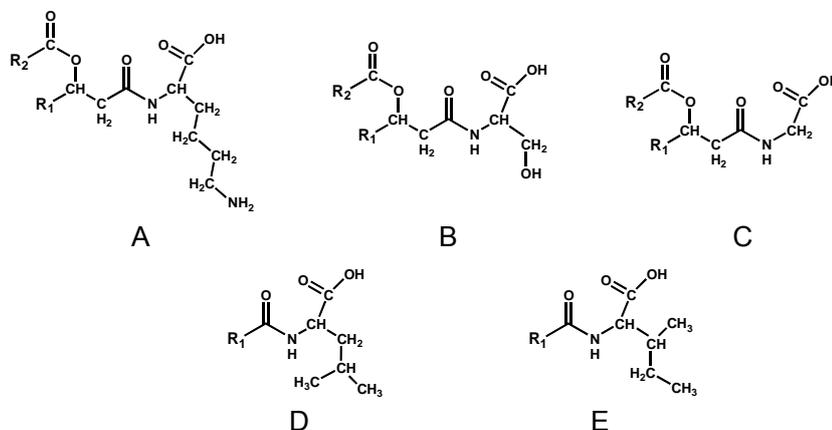


FIGURE 1-6 Non-ornithine-containing lipoamino acids identified in microorganisms. The year of first report is in following brackets. A, lysine lipid (1976); B, serine lipid (1988); C, glycine lipid (1991); D, leucine lipid (1997); E, isoleucine lipid (1997). The alkyl chains of acyl substituents are abbreviated as: R₁, amide-linked fatty acids; R₂, ester-linked fatty acids. The 2-hydroxyl linkage is not shown here.

In most cases the *O*-ester linking fatty acids at the second position are non-functionalized; functionalized fatty acids at this position were detected occasionally. This also reflects the species characteristic. 2- or 3-Hydroxylated fatty acid moiety incorporated into the second position of lipoamino acids were found in *Flavobacterium* spp. (Kawai *et al.* 1988). 11,12-Epoxy octadecanoic acid has been identified in *Thiobacillus* sp. A2 as an ester-linked fatty acid (Thiele *et al.* 1984). Tuberculostearic acid was predominantly found among the *O*-acylated fatty acids in mycobacterial ornithine lipids (Laneelle *et al.* 1990).

1.2.1.2 Biosynthesis of lipoamino acids

Most of the genetic work on lipoamino acids biosynthesis was done in 70th and 80th of the last century. An enzyme involved in ornithine lipid biosynthesis, namely the lyso-ornithine lipid *O*-acyltransferase, was first studied in *Gluconobacter capsulatus* (Thiele *et al.* 1980). The corresponding gene, named *olsA* has been cloned from *Ensifer meliloti* (or *Rhizobium meliloti*). It catalyses the *O*-acylation of lyso-ornithine lipid, strictly requiring acyl carrier protein (ACP) -activated fatty acids (see 1.3.1 for detail) as substrates (Weissenmayer *et al.* 2002). So far, no enzyme activity involved in formation of the amide bond of lipoamino acids has been characterized (Yagi *et al.* 1997; Brady and Clardy 2000; Weissenmayer *et al.* 2002).

In a screening of an environmental DNA library for antimicrobial genes, a putative long-chain *N*-acyl tyrosine synthase gene has been proposed from an unculturable soil proteobacterium (Brady and Clardy 2000). The corresponding protein catalyses the

transfer of an acyl group from ACP to tyrosine. Recently the gene was found to be part of a 13-open-reading-frame gene cluster, coding for the biosynthetic pathway producing acyl phenols (Figure 1-7). The latter represent antibiotic compounds. *N*-acyl tyrosine is, however, not the final product (Brady *et al.* 2002) and therefore this gene does not really relate to lipoamino acid biosynthesis.

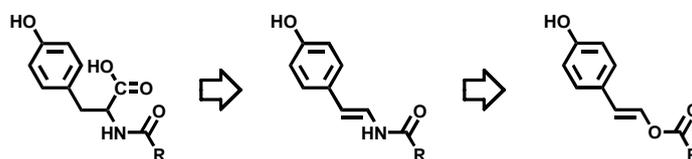


FIGURE 1-7 Putative biosynthetic pathway of acyl phenols in an uncultured soil proteobacterium proposed by Brady and Clardy.

1.2.2 Lipopeptides

Lipopeptides are more complex than lipoamino acids, not only by structure, but also by composition. A defined kind of lipopeptide is always a mixture of components similar by structure, but different by molecular weight. The general structure of these molecules is a cyclic peptide linked to a fatty acid tail. In most of the lipopeptides, the first amino acid is linked to a β -hydroxy fatty acid. The carboxyl-terminal amino acid forms a lactone ring with the β -hydroxyl group of the fatty acid (Figure 1-8). The difference of molecular mass is mainly due to the size of the lipid substituents (Table 1-2). Surfactant BL-86, one of the lichenysin group lipopeptides isolated from *B. licheniformis* 86, is a collection of components ranging in mass from 979 to 1,091 Da (Horowitz and Griffin 1991).

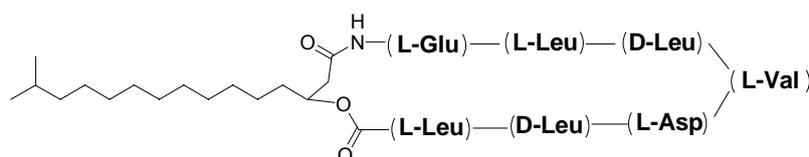


FIGURE 1-8 A representative structure of surfactin-type lipopeptides from *Bacillus subtilis*.

1.2.2.1 Classification of lipopeptides

The *B. subtilis* lipopeptides are classified into surfactin, inturin, and fengycin families. Surfactin (Figure 1-8), the best-known lipopeptide was published in 1968 (Ahimou *et al.* 2000). It is one of the most powerful biosurfactants (Desai and Banat 1997) and exhibits a broad antibacterial spectrum. The peptide ring adopted a “horse-saddle” structure in solution with the two charged residues forming a “claw”, which is a potential binding site for divalent cations (Heerklotz and Seedlig 2001). It is believed that most of these peptides are targeted to biological membranes and increase their

Introduction

permeability by the formation of ion-conducting pores in membrane, which kills the cells (Avrahami and Shai 2004).

Table 1-2 Major groups of identified lipopeptides.

Name	Origin	Amino acids ^a	Fatty acid ^b	Function
Surfactin	<i>Bacillus</i> spp.	7 (7)	β -hydroxy -(C ₁₃ - C ₁₆)	Antibiotic
Iturin	<i>Bacillus</i> spp.	7 (7)	β -amino-(C ₁₄ - C ₁₇)	Anti-fungal agent
Fengycin	<i>Bacillus</i> spp.	10 (8)	β -hydroxy-(C ₁₅ - C ₁₈)	Anti-fungal agent
Polymyxin	<i>Bacillus</i> spp.	10 (8)	Branched chain fatty acid connected 2,4-diaminobutyric acid	Antibiotic
Viscosin	<i>Pseudomonas</i> spp.	9 (7)	β -hydroxy -(C ₁₀)	Surfactant
Serratamolide	<i>Serratia</i> spp.	2 (2)	β -hydroxy-(C ₁₀)	Antibiotic and surfactant
serawettin W2	<i>Serratia liquefaciens</i>	5 (5)	β -hydroxy-(C ₁₀)	Wetting agent
Echinocandin	Coleomycetes Hyphomycetes	6 (6)	Long fatty acid	Anti-fungal agent

a, The first number specifies the number of amino acids in the molecule, the number in brackets specifies the number in the cyclic peptide ring.

b, The length of the fatty acid is shown in brackets.

The synthesis of these lipopeptides during the early stages of sporulation is common to most, if not all members of the genus *Bacillus* (Rosenberg and Ron 1999). Lipopeptide profiles vary greatly between different strains. Among the three lipopeptide types only iturin A is produced by all *B. subtilis* strains (Ahimou *et al.* 2000). The structures of iturin-type lipopeptides are unique on that a β -amino fatty acid forms two amide linkages. Its carboxyl group is linked to the α -amino group of the first amino acid of the peptide chain and its β -amino group is linked to the carboxyl group of the last amino acid of the peptide bond, respectively. In all of the other lipopeptides, the latter linkage is an ester bond. Lichenysin produced by *B. licheniformis* and other related bacilli resembles surfactin. A glutamine residue at position 1 replaced the glutamic acid residue in surfactin. This modification causes a significant reduction of the critical micellar concentration and makes lichenysin a better chelating agent (Lang 2002).

Fengycin is produced by *Bacillus subtilis* F29-3, and is a cyclic lipopeptide with a peptide head of 10 amino acids; a lactone bond connects the third amino acid residue L-Tyr with the tenth residue of a L-Ile (Vanittanakom *et al.* 1986). Plipastatin is a member of the fengycin-type lipopeptide in which one glutamine replaced the glutamate residue of fengycin (Tsuge *et al.* 1999). Polymixin is produced by *B. polymyxa* and other strains. Like fengycin it is a decapeptide in which the amino acids 3 to10 form a cyclic octapeptide. The γ -amino groups of the 2,4-diaminobutyric acid residue, together with the hydrophobic fatty acid chain, give polymyxin-type lipopeptides the cationic surface activity (Ron and Rosenberg 2001).

Most of these biosurfactants were isolated from *Bacillus* spp., but viscosin is produced by soil-inhabiting *Pseudomonas* spp. They can protect the plant root from rot by reducing surface tension to increase motility and lysis of fungal zoospores (de Souza *et al.* 2003). The echinocandin group is produced by fungi, e.g. *Aspergillus* spp. Many of these compounds are water soluble due to the presence of a sulfate residue (Hino *et al.* 2001). Echinocandin-type lipopeptides consist of a cyclic hexapeptide nucleus acylated with a long fatty acid chain. The producing strains were classified into two groups, Coleomycetes and Hyphomycetes group (Lang 2002).

The simplest structure of lipopeptides is represented by serratamolide, which covers the surface of Gram-negative *Serratia* strains (Matsuyama *et al.* 1985). It is a cyclodepsipeptide [cyclo-(*D*-3-hydroxydecanoyl-*L*-seryl)₂]. This compound is different from other lipopeptides in that it contains two fatty acids per molecule forming a bilateral symmetric structure.

Another lipopeptide found in the same genus, serawettin W2 is also simpler than the structures mentioned above. It is composed of five amino acids. The first is linked to the fatty acid via an amide bond, and the last forms a lactone with the β -hydroxyl group of the fatty acid (Lindum *et al.* 1998). They were shown to increase cell hydrophilicity by blocking the hydrophobic sites on the cell surface (Bar-Ness *et al.* 1988) and to promote flagellum-independent mobility (Matsuyama *et al.* 1995).

1.2.2.2 Biosynthesis of lipopeptides

The peptide bone of the lipopeptides is synthesized by non-ribosomal peptide synthase (NRPS). Some of the biosynthesis operons have been cloned. The gene clusters for surfactin biosynthesis is encoded by the 25 kb *urfA* operon (Nakano *et al.* 1991), and that for lichenysin by a highly homologous 26.6 kb *lchA* operon (Konz *et al.* 1999). These clusters code for giant proteins composed of repeating domains whose function and order parallel the consequence of amino acids in the peptide. Additional genes are also required (Desai and Banat 1997). Although the acyltransferase activity had been detected in a fraction of the surfactin synthase multienzyme system *in vitro* (Menkhaus *et al.* 1993), no isolated enzyme responsible for the transfer of the fatty acid to the peptide has been identified so far (Ron and Rosenberg 2001).

By transposon mutagenesis, some other gene clusters responsible for lipopeptide synthesis have been identified. The *pps* operon identified in *B. subtilis* is larger than 100 kb, in which 5 open reading frames (ORF) are organized into 10 amino acids activating domains and 4 putative racemase domains. It had been believed to be the fengycin operon (Tosato *et al.* 1997), but recently was demonstrated to encode plipastatin synthase (Tsuge *et al.* 1999). The biosynthesis gene cluster of viscosin has been localized in *P. fluorescens* by transposon intergrating mutation (de Souza *et al.* 2003). And also based on this method, the *swrA* gene in *S. liquefaciens* had been proven to be involved in serrawettin W2 production.

1.2.3 Lipoproteins

Lipoproteins are the largest biomolecules modified by acylation via amide bonds. They represent a large group of extracellular proteins, on which some special residues are acylated. Some are secreted (Desai and Banat 1997), others are structural components of cell wall, or represent as transporter proteins (Braun and Wu 1994).

1.2.3.1 Secreted proteins

A large number of complex proteinaceous bioemulsifiers has been reported. But in general, little is known about their chemistry due to their structural complexity. For instance, a peptidoglycolipid produced by *P. aeruginosa* was found to consist of 52

amino acids, 11 fatty acids and a sugar unit (Desai and Banat 1997). The data available suggested that the production of proteaceous bioemulsifiers by bacteria is correlated with high bacterial cell density (Ron and Rosenberg 2001).

The bioemulsifier of *Acinetobacter radioresistens*, referred as alasan, is a complex of a polysaccharide and a protein. The active components are proteins. Among the three isolated proteins, the 45-kD protein has highest specific emulsifying activity (Lang 2002). Interestingly, this bioemulsifier has been shown to bind to the surface of other bacteria, e.g. *Sphingomonas paucimobilis* and to change their surface property (Osterreicher-Ravid *et al.* 2000). This horizontal transfer of bioemulsifier from one bacterial species to another has a significant implication in interaction of natural microbial communities as well as biofilm formation.

1.2.3.2 Cell wall proteins

Two different types of lipoproteins have been located in the outer membrane of *E. coli*, Braun's lipoprotein and the so-called peptidoglycan-associated lipoproteins (PAL) (Braun and Wu 1994). Braun's protein is a small 7.2 kDa protein which is dominant in the outer membrane. About one third of them are covalently linked to the murein, the others are associated with the peptidoglycan. The sulfhydryl group of the *N*-terminal cysteine is substituted with a diglyceride, the amino group is substituted with a fatty acid by an amide linkage (Figure 1-9).

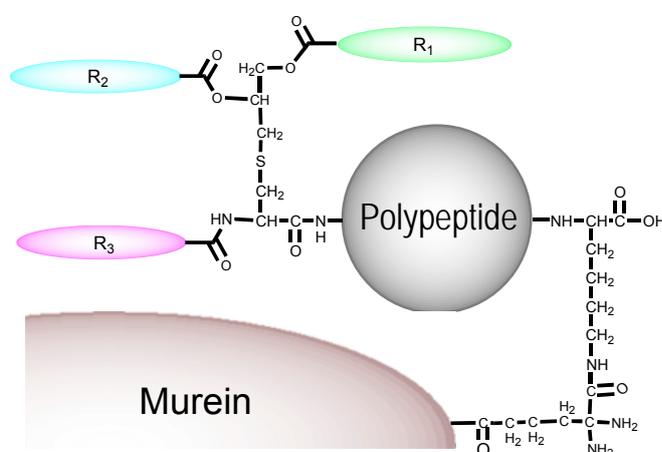


FIGURE 1-9 The terminal amino acid of Braun's protein (grey) serve as lipids (*N*-terminal cysteine) and murein (*C*-terminal lysine) attachment site. R₁, R₂, and R₃ are the side chains of fatty acids.

Lipoproteins are not only membrane-associated, but some are also secreted by pathogenic bacteria. A 38 kd lipoglycoprotein is actively secreted by *Mycobacterium tuberculosis* and is partly attached to the surface of cell wall by a lipid tail. The protein induces strong antibody and T-cell response in mammalian body (da Fonseca *et al.* 1998). Many proteinaceous toxins secreted by pathogenic Gram-negative bacteria are activated by acylation of two critical lysine residues on the protein, such as *E. coli* haemolysin (Stanley *et al.* 1998).

1.2.3.3 *N*-Acylation of the prolipoprotein

The maturation of Braun's protein requires modification of translated prolipoproteins by transfer of glycerol and fatty acids moieties from phosphatidylglycerol and phospholipids. An *in vitro* study suggested that the *N*-acyl group was transferred only when the prolipoprotein was correctly modified with ester-linking of glyceride and removal of signal peptide (Tokunaga *et al.* 1982). The condensation of fatty acids to the terminus of the modified cysteine residues is catalysed by an apolipoprotein *N*-acyltransferase (Sankaran and Wu 1994). This is the only branch of the nitrilase superfamily known to catalyse the reversal amidase reaction *in vivo* (Brenner 2002). No pre-activation of the substrates is required for this reaction. Defects in this acylation pathway render the cell susceptible for copper toxicity (Pace and Brenner 2001).

1.3 Bacterial amidation pathways

Neither the *N*-acyl amino acids-producing crude protein fraction prepared from insect alimentary tissue (Lait *et al.* 2003) nor the protein extract synthesizing *N*-acylethanolamide from mammalian central nervous system (Kruszka and Gross 1994) require substrate activation for the amide bond formation. Considering the energy consumption in such reactions, probably there exists an unknown activation mechanism (for discussion see 1.1.3), or there are enzymes that catalyse reverse amidation like the amidase involved in the maturation of Braun's protein. These atypical pathways seem to be quite different from the mechanisms for amide bond formation characterized so far in bacteria.

1.3.1 Amide synthase

Amide synthase is a kind of prokaryote-specific amidation enzyme. Two kinds of amide synthase genes had been cloned. Both require energy to drive the reaction forward. One type of enzyme is related to the condensation domain of non-ribosomal peptide synthase (NRPS). ATP is hydrolyzed during the activation of the carboxyl group (Figure 1-10). VibH from *Vibrio cholerae* vibriobactin biosynthesis (Keating *et al.* 2000) is a unique enzyme that acts upon an aryl carrier protein domain (ArCP)-bound acyl group and uses a free amine to produce an amide.

In contrast to VibH, which is free-standing, most NRPSs condensation domains are embedded in the giant multifunctional proteins (Stachelhaus *et al.* 1998). The prototype NRPS condensation domain catalyses the condensation between two carrier protein-linked substrates, typically amino acids tethered to a 4-phosphopantetheinyl group (P-pant) of a carrier protein domain (Ehmann *et al.* 2000). VibH doesn't require the downstream substrate tethered to an ACP analog. In this respect, VibH is more like acyl transferases. But VibH uses a similar mechanism as NRPSs catalysing the transfer of an activated acyl group tethered to carrier protein domain to the free amino group of an acceptor.

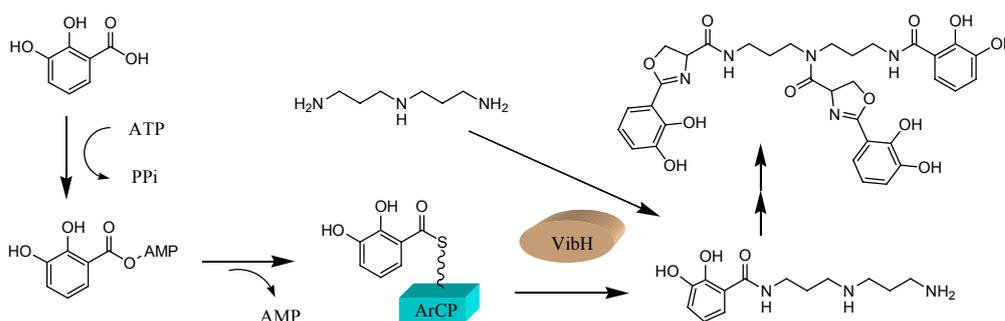


FIGURE 1-10 Part of the biosynthetic pathway to the iron chelator vibriobactin of *V. cholerae*. The 2,3-dihydroxybenzoic acid is activated via thioester linkage to the aryl carrier protein domain (ArCP) of VibB by hydrolysis of ATP. The condensation of this thioester with norspermidine is then catalyzed by VibH. This amidation is a critical step in vibriobactin (upper right) synthesis.

Three amide synthetases, which catalyse both the activation of an acyl component by adenylation and subsequent transfer of the acyl group onto the amino group of the intramolecular receptors, have been identified in aminocoumarin biosynthesis in *Streptomyces* spp., namely NovL involved in novobiocin biosynthesis (Steffensky *et al.* 2000), CouL in coumermycin A₁ biosynthesis (Schmutz *et al.* 2003), and CloL in chlorobiocin biosynthesis (Galm *et al.* 2004). They are large proteins containing adenylation motifs in addition to the acyltransferase modus. Interestingly, these enzymes do not contain the typical P-pant cofactor suggesting a different mechanism for acyltransfer (Steffensky *et al.* 2000). They are probably more homologous to acyl-CoA synthetases (Schmutz *et al.* 2003).

The activation of the acyl group substrate of VibH is driven by hydrolysis of ATP, which is catalysed by a separate domain. The same activation mechanism is operative in mycobacteria. Fatty acids are first activated by adenylation, which is catalysed by FadD. The acyl-adenylates are then transferred to a polyketide synthase (PKS) and act as elongation units (Trivedi *et al.* 2004). Another example is the initiation of rifamycin biosynthesis catalysed by the loading domain of rifamycin synthetase (Figure 1-11) in *Amycolatopsis mediterranei* (Admiraal *et al.* 2001). These two reactions activate the acyl groups via the formation of a high-energy thiol ester bond between the acyl group and the enzyme. Acyl transfer is the following step.

A second amide synthase, Riff is a component of the rifamycin PKS in *A. mediterranei*. This protein catalyses an intramolecular amide formation to release the completed polyketide, in which the chain extension units are CoA derivatives (Pompeo *et al.* 2002). The high energy thiol ester bond is then transferred from CoA to PKS (Floss and Yu 1999). The PKS condensation reaction is similar to that used by the fatty acid synthases (FAS) found in all organisms where fatty acids are required for the biosynthesis of lipids (Hopwood, 1997). Like NRPSs, PKSs as well as FAS enzyme complexes require post-translational modification of their ACP to become catalytically active. The inactive apo-proteins are converted to active holo-enzymes by esterification of a specific serine hydroxyl with the P-pant group of coenzyme A (Wang *et al.* 2001).

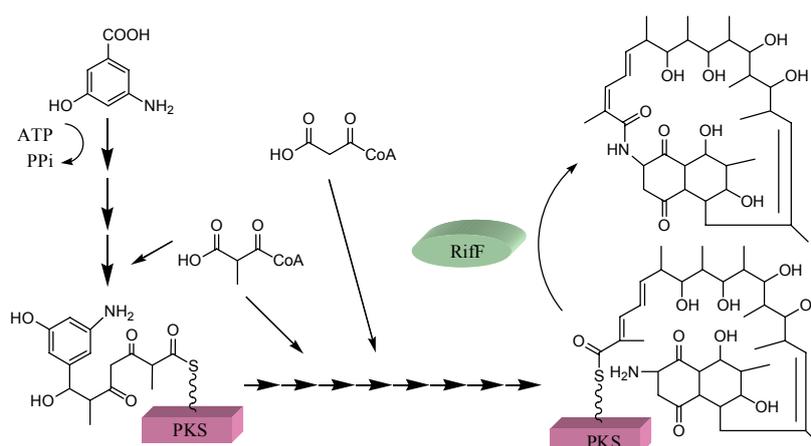


FIGURE 1-11 Part of the biosynthetic pathway to the rifamycins in *A. mediterranei*. The initiation unit is activated by ATP hydrolysis. The malonyl- and methylmalonyl-CoA are chain elongation units. Riff catalyses the release of the polyketide from the PKS. The macrocyclic lactam proansamycin X is further modified to rifamycin B.

Riff catalyses the formation of the macrocyclic ring (Figure 1-11) by utilizing the energy of the PKS thiol ester bond. Based on sequence similarity, Riff is homologous to arylamine *N*-acetyltransferase, which catalyse the transfer of an acetyl group from acetyl-CoA to a free amino group or hydroxyl group of the second substrate. Thus, Riff is a free-standing *N*-acyltransferase domain, that catalyses the transfer of activated carboxyl groups from PKS instead of CoA.

1.3.2 *N*-Acyltransferases

CoA is the major coenzyme of enzymes involved in fatty acid metabolism such as FAS (Schweizer 1989). It is not surprising that acyl-CoA is the donor of acyl group in most of the acylation reactions. Acyl-CoA:ACP acyltransferase domains widely exist in bacterial Type I PKSs (Kuczek *et al.* 1997) and hybrid peptide-polyketide natural product synthetases (Hoffmann *et al.* 2003). These acyltransferases catalyse the chain elongation of the polyketides. The product is a ketone instead of an acyl-amido group. The ACP-activated fatty acids can also be used as substrate in subsequent reactions catalysed by amidase orthologs (see 1.3.3) or acyltransferase orthologs (see 1.3.1). One of such ACP-dependent acyltransferases, HlyC is responsible for the acylation of two critical lysine residues on *E. coli* haemolysin (Stanley *et al.* 1999). Many toxins secreted by pathogenic Gram-negative bacteria are activated by HlyC-type activation (Stanley *et al.* 1998).

In addition to amide bonds, some other linkages can also be formed by the catalysis of acyltransferases (Stanley *et al.* 1998). In palmitoylation of intrinsic and peripherally associated membrane proteins, palmitate or other long chain fatty acids are added to a cysteine residue via a thioester linkage (Linder and Deschenes 2003). *N*-Acyltransferase homologs can also use other acyl-activated substrates in replacement of CoA derivatives to drive the reaction. Lipid A is the essential component of lipopolysaccharide from Gram-negative bacteria. In lipid A biosynthesis, the outer membrane enzyme PagP and

homologs catalyze the transfer of a palmitoyl group from a phospholipid to an hydroxyl group of lipid A (Hwang *et al.* 2002).

1.3.3 Amidases and proteases

Some proteinases and amidases are known to catalyze a low level of the reverse reaction, *i.e.* the human acylxylacyl hydrolase has been demonstrated to have acyltransferase activity (Munford and Hunter 1992). Such a reverse reaction is of low efficiency but energy-saving, since they do not need any activation. Some of them are employed in industrial production of commercially important compounds (Kobayashia *et al.* 1992). Proteases and amidases, as well as nitrilases widely exist in both prokaryotic and eukaryotic organisms.

1.3.3.1 Proteases and their catalytic mechanisms

Proteases are mainly divided into four categories: serine/threonine proteases, cysteine proteases, aspartic proteinases, and metalloproteases (Kobayashi *et al.* 1997). Hydrolysis of peptide bonds is an energetically favourable reaction, but extremely slow (Powers *et al.* 2002). Proteases facilitate fast reactions. The transition states of all proteases involve the formation of a tetrahedral intermediate.

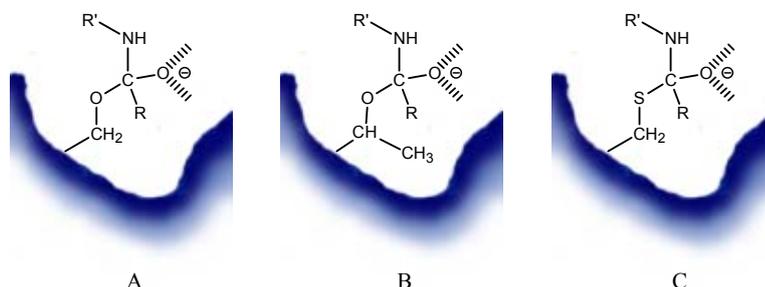


FIGURE 1-12 Schematic diagram of the transition states of the reaction of serine (A), threonine (B), and cysteine (C) proteases. Only the substrate-enzyme covalent linkage and the oxyanion are shown; the oxyanion binding sites are omitted.

The active site residues of serine/threonine and cysteine proteases have many mechanistic features in common (Figure 1-12). Each enzyme has an active site nucleophile and a basic residue, which function as general acids and bases in the catalytic mechanism (Powers *et al.* 2002). Serine proteases are members of the serine hydrolase family. They contain the classical serine-histidine-aspartic acid catalytic triad (Patricelli and Cravatt 2000) and possess an oxyanion stabilizing site (Carter and Wells 1988).

Such an oxyanion stabilization is also important for cysteine proteases. Catalysis of cysteine proteases proceeds through the formation of a covalent intermediate by the involvement of a cysteine and a histidine residue. The nucleophile is a thiolate ion rather than a hydroxyl group. The thiolate ion is stabilized through the formation of an ion pair with neighbouring imidazolium group of histidine. The attacking nucleophile is the thiolate-imidazolium ion pair in both steps (Otto and Schirmeister 1997).

Aspartic acid proteinase has two aspartic acid residues in its catalytic center. In contrary to serine and cysteine proteases mentioned above, anionic intermediates are stabilized sequentially by a complex hydrogen-bonding network at each of the active site aspartic acids in the two intermediates during the catalytic cycle (Veerapandian *et al.* 1992).

Metalloproteases differ widely in their sequences and their structures, but the great majority of enzymes contains a zinc atom, which is catalytically active. In some cases, zinc may be replaced by another metal such as cobalt or nickel without loss of the activity (Dardel *et al.* 1998). A non-covalent tetrahedral intermediate forms in the reaction after the attack of a zinc-bound water molecule on the carbonyl group of the scissile bond, followed by transfer of a proton to the leaving group (Wouters and Husain 2001).

1.3.3.2 Amidases and their active centers

Amidases represent a subgroup of the nitrilase superfamily, which catalyse the hydrolysis of carbon-nitrogen linkages. Most of the members of the superfamily are proposed to utilize a catalytic triad consisting of glutamic acid, lysine and cysteine (Brenner 2002). Major branches of this family are aliphatic amidases, which catalyse the hydrolysis of amides to carboxylates and ammonia (Fournand and Arnaud 2001). All of the enzymes contain a highly conserved motif (GGSS) with unknown function. Some of them are made up four, six, or eight subunits and are classified as sulfhydryl enzymes.

Although the members of the nitrilase superfamily have substrates in common, they are structurally and mechanistically unrelated (Brenner 2002). The so-called signature amidases constitute a large family defined by a conserved stretch of approximately 130 amino acids termed “amidase signature sequence”, which represents a class of serine-lysine catalytic dyad hydrolases whose evolutionary distribution rivals that of the glutamate-lysine-cysteine catalytic triad amidase superfamily (Patricelli and Cravatt 2000). There are some other amidase-signature enzymes utilizing a serine as a catalytic nucleophile and a lysine as a catalytic base for nucleophile activation (Patricelli and Cravatt 2000). Still others like the *Rhodococcus* amidase have a serine and an aspartic acid residue in its active site (Kobayashi *et al.* 1997) reminiscent that of the aspartic acid proteinase.

The majority of the amidases are cytoplasmic proteins. Cell-wall lytic amidases attached to the cell wall surface have been reported to be involved in cell-wall growth, cell-wall turn over, and cell separation. Some of them are even encoded by temperate phages (Saiz *et al.* 2002). Others are extracellular proteins secreted by the bacteria (Neumann and Kula 2002).

1.3.3.3 Acyl transfer activity in amidases

Amidases operate by an acyl-enzyme mechanism similar to that of proteases (Figure 1-12) and show acyl transferase activity (Cravatt *et al.* 1996). As mentioned above, a membrane-bound amidase in *E. coli* is responsible for lipoprotein modification via a reverse amidation process (Sankaran and Wu 1994). 28 homologs of this kind of amidase have been found in prokaryotes by sequence analysis (Brenner, 2002).

It has also been reported that all aliphatic and enantioselective amidases catalyse an acyl transfer to hydroxylamine as the acyl acceptor leading to the formation of hydroxamic acids. They are able to transform diverse amides, acids, esters, or nitriles into the corresponding carboxylic acids, hydroxamic acids or acid hydrazides (Fournand and Arnaud 2001). This reaction pathway is characterised as a Ping Pong Bi Bi mechanism according to the acyl group behavior in reaction.

1.4 Goals of this research

In previous work, some gut bacterial strains had been demonstrated to be able to produce *N*-acyl amino acids (Spiteller *et al.* 2000). From gut segments of *Spodoptera exigua*, *Mamestra brassicae* and *Agrotis segetum*, ca. 30 bacterial strains were isolated and identified by sequencing of their 16S rDNA. About half of them were able to synthesise the typical lepidopteran *N*-acylamino acids from externally added precursors (Spiteller 2002). Some gut bacterial strains provided by Prof. Dettner will be additionally studied in this project.

The *N*-acyl amino acid synthesis activity of some of the gut bacteria will be screened. The most active strain will be selected for further identification and mechanism study. The role of *N*-acyl amino acids and their properties in the interaction with their insect host will be studied. For this purpose, the biological characters of the active strain will be studied with fluorescence microscopy and other techniques. The enzyme activity of the model strain will be correlated with the growth stage to find out corresponding gene regulation by the bacteria. The activity will also be analysed in different fractions of the bacterial extract to localize the active protein. The active protein will be purified by chromatography with different selective columns. After purification, the catalytic properties such as pH optimum, temperature optimum, substrate selectivity will be determined. The catalytic mechanism of the active protein will be studied by use of inhibitors. The purified protein will be sequenced by Edman degradation and tandem mass spectrometry sequencing after trypsin digestion. The protein sequences will be used for looking for homologous proteins. The corresponding gene will be located on the genomic DNA of the model strain. The cloned gene will be analysed by different methods including expression in *E. coli* or other amendable hosts.

2 Materials and Methods

2.1 Materials

2.1.1 Plasmids

[The maps and detailed features of the plasmids constructed in this study are shown as Appendix I.]

Table 2-1. Plasmids and their relevant characteristics.

Plasmid	Relevant characteristics	Source
pEM1i2	Expression plasmid in <i>E. coli</i> , <i>afp</i> ORF was introduced into pET11a .	This study
pHER	Intermediate plasmid, a fusion plasmid of pEM1i2 and pHN15 .	This study
pRMM	Expression plasmid in in <i>M. arborescens</i> DSM20754 expression. This plasmid was constructed by introducing a sliced expression cassette from pHER into the vector of pSEK4 (Floss company), a derivative of pRM5 .	This study
pRMC	The control plasmid in which the <i>afp</i> expression cassette has been removed from the <i>Bam</i> HI sites. All other features are identical with pRMM .	This study
pCR2.1@-TOPO	Cloning vector for PCR fragments amplified by normal Taq polymerase.	Invitrogen
pCR@-Blunt II-TOPO	Cloning vector for PCR product amplified with proofreading enzymes.	Invitrogen
pBluescript II SK	Phagemid vector for cloning restriction enzyme digested DNA fragment.	Stratagene
pET11a	Vector for expression in <i>E. coli</i>	Stratagene
pHN15	A plasmid based on pBluescript II KS . A <i>veg</i> p promoter from <i>B. subtilis</i> controls the gene expression of <i>natI</i> , named pBlueneo. The <i>natI</i> gene encodes an enzyme disable the antibiotic clonNAT.	WERNER BioAgents, Jena Germany
pSEK4	Expression vector in <i>Streptomyces</i> spp. (McDaniel et al. 1993). The plasmid pSEK4 (Floss company) used in this research is the derivative of this plasmid (pRM5-actKR)	Floss company

2.1.2 Microbial sources and properties

M. arborescens SE14 and other strains were isolated from Beet armyworm (*Spodoptera exigua*) larvae. *M. arborescens* DSM20754 was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). The antibiotic resistance spectra of the two strains are shown as Appendix II. *E. coli* competent cells were purchased from companies as in table 2-2:

Table 2-2. *E. coli* strains used in this study.

Host strain	Genotype	Company
BL21-CodonPlus(DE3)-RP <i>E. coli</i> strain	B F ⁻ <i>ompT hsdS</i> (r _B ⁻ M _B ⁻) <i>dcm</i> ⁺ Tet ^r <i>gal</i> λ(DE3) <i>endA Hte</i> [<i>argU proL Cam</i> ^r]	Stratagene
INV110 <i>E. coli</i> strain	F' <i>tra</i> Δ36 <i>proAB lacI</i> ^{qZ} ΔM15/ <i>rpsL</i> (Str ^R) <i>thr leu endA thi-1 lacY galK galT ara tonA tsx dam dcm dupE44</i> Δ(<i>lac-proAB</i>) Δ(<i>mcrC-mrr</i>)102::Tn10 (Tet ^R)	Invitrogen
INVαF' <i>E. coli</i> strain	F' <i>endA1 recA1 hsdR17</i> (r _k ⁻ , m _k ⁺) <i>supE44 thi-1 gyrA96 relA1</i> Φ80 <i>lacZ</i> Δ M15 Δ(<i>lacZYA-argF</i>) U169λ	Invitrogen
TOP10 <i>E. coli</i> strain	F' <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 ara</i> Δ139 Δ(<i>ara-leu</i>)7697 <i>galU galK rpsL</i> (Str ^R) <i>endA1 nupG</i>	Invitrogen

2.1.3 Chemicals

Most of the chemicals were purchased from Sigma/Aldrich, except the following: dipeptides (BACHEM Biochemica GmbH, Heidelberg, Germany); mono-, di-, tri-linolenyl glycerol, and linolenoyl-CoA (Larodan Fine Chemicals AB, Malmö, Sweden); Phenyl dodecanoyl glutamine and 17-hydroxyl linolenic acid were synthesized in our lab. The LogP and reaction standard free energy (25 °C, 1 atm, 1 M, ΔG) were calculated with CS ChemOffice 6.0 (CambridgeSoft, USA).

2.2 Methods

2.2.1 Isolation and identification of microorganism

Beet armyworm (BAW) was hatched from eggs and reared on an artificial diet at 23-25°C under a 16-h light/8-h dark cycle as described by Spiteller *et al.* (2000). Fourth instar larvae were surface-sterilized with 70% ethanol, and then dissected. The gut content was diluted with 0.9% NaCl and spread on Brain Heart Infusion medium (BHI)

plates. The plate was kept at 20°C for 7 days. Bacterial colonies were selected primarily based on morphology and standard classification method (Holt 1993).

Universal primer pairs for eubacterial 16S rDNA, (DeLong 1992) were used in PCR reactions to amplify nearly full-length sequences of 16S rDNA genes:

27f

5'-AGAGTTTGATCCTGGCTCAG-3'

and 1492r

5'-GGTTACCTTGTTACGACTT-3'

Bacteria were cultured in BHI broth overnight at 37°C and frozen at -20°C. Thawed cultures were diluted 100 times with water, 1 µl was used as template for PCR reaction. Reaction mixtures contained 1.5 mM MgSO₄, 200 µM dNTP, 2.5 units *HifiTaq* (Invitrogen) and buffer supplied by producer. Thermal cycling was initiated at 96°C for 5 min to lyse the cells, denaturation at 96°C for 30 sec, annealing at 55°C for 45 min, and extension was achieved at 72°C for 2 min, in total 35 cycles. Final elongation was done at 72°C for 3 min. The PCR products were cloned into pCR-Blunt II-TOPO vector (Invitrogen), and sequenced. Monitoring of the growth rate was performed by measuring the OD₆₀₀ by a photometer (Eppendorf). Bacteria were grown in 3 ml liquid medium in glass tubes unless otherwise specified.

Morphological observation was performed with a fluorescent microscope (Axioskop, Zeiss, Germany) after DAPI staining as mentioned in the following FISH protocol (obtained from Prof. R. I. Amann's lab, Max-Planck Institute of Marine Microbiology, Bremen, Germany).

Monitoring of bacteria by FISH

Solutions and buffers:

1. PBS buffer PH 7.3 (Phosphate buffered saline)

(According 'molecular clone', B. 12).

Dissolve:

NaCl	8 g
KCl	0.2 g
Na ₂ HPO ₄	1.44 g
KH ₂ PO ₄	0.24 g

in 800 ml of distilled H₂O. Adjust the pH to 7.3 with HCl. Add H₂O to 1 liter.

Dispense the solution into aliquots and sterilize them by autoclave for 20 minutes at 15 lb/sq. Store at room temperature.

2. 5 M NaCl

NaCl	292.2 g
------	---------

Dissolve in H₂O to final volume of 1 l. Dispense into aliquots and sterilize by autoclaving.

3. 10% SDS (sodium dodecyl sulfate)

Dissolve:

SDS 100 g

in 900 ml of H₂O. Heat to 68 °C to assist solvation. Adjust the pH to 7.2 by adding a few drops of concentrated HCl (37%). Adjust the volume to 1 liter with H₂O. Dispense into aliquots.

4. 1 M Tris buffer

Dissolve:

Tris base 121.1 g

in 800 ml H₂O. Adjust the pH to the required value by adding concentrated HCl.

PH	HCl
7.4	70 ml
7.6	60 ml
8.0	42 ml

Allow the solution to cool to room temperature prior to final adjustment of the pH. Adjust the volume of the solution to 1 liter with H₂O. Dispense into aliquots and sterilize by autoclaving.

5. PFA (paraformaldehyde) fixative 16% :

Dissolve:

PFA 16 g

In 60 ml MQ water, add a drop 2N NaOH to help dissolve, **do not let it overwarm!** Add 33 ml 3 x PBS, adjust pH to 7.2, fill to 100 ml. Filter sterile, store at 4°C.

6. *In situ* hybridization buffer:

Dissolve:

5M NaCl 360 µl

1 M Tris-HCl, pH8.0 40 µl

deionized formamide ~ µl (according table 2-3)

fill to 2 ml with MQ water

add 2 µl 10% SDS

7. Washing buffer:

Dissolve:

1 M Tris-HCl, pH 8.0 1 ml

5 M NaCl ~ µl (according table 2-3)

0.5 M EDTA 500 µl

Fill to 50 ml with MQ water

Add 50 µl 10% SDS

8. DAPI staining solution (1.0 µg/ml):

Materials and Methods

Dissolve 4',6-Diamidino-2-phenylindole (DAPI) in pure ethanol to make stock solution (20 µg/ml), then dilute the stock solution to 1 µg/ml with 1 x PBS buffer just before use.

Procedure:

1. Wash the slide in 95% ethanol for 10 min, allow to air dry vertically. The slide was coated with gelatin by briefly dipping slides in warm solution (70°C) of 0.5% gelatin and 0.25% chrome (III) potassium sulfate. Leave them air dry in a vertical position overnight.
(Gelatin-coated slides can be stored at 4°C for several months.)
2. The insects were dissected, and gut content was taken out. The tissue was resuspended with 1 x phosphate-buffered saline (PBS), and then fixed with paraformaldehyde (PFA) for 30 minutes. The final concentration of PFA is 4% in 1 x PBS solution. The sample was washed twice with 1 x PBS solution, then mix with the same volume of ethanol. The gut content preparation can be stored at -20 °C permanently.
3. 5 µl samples were dipped on the microscope slide, and dried at 48 °C for 10 min.
4. After drying, the slides were dehydrated in ethanol of increasing concentration. 50% → 80% → 96%. Slides were held in each concentration for 2 min. Finally the slides were dried at 37°C for 2 min.
5. The hybridization buffer was prepared according to table 2-3.

Table 2-3. NaCl and EDTA concentrations in the washing buffer

Formamide (hybridization buffer) %	NaCl mM	5 M NaCl µl	0.5 M EDTA µl
0	900	9000	--
5	636	6300	--
10	450	4500	--
15	318	3180	--
20	225	2150	500
25	159	1490	500
30	112	1020	500
35	80	700	500
40	56	460	500
45	40	300	500
50	28	180	500
55	20	100	500
60	14	40	500
65	10	NaCl in EDTA enough	500
70	7	--	350
75	5	--	250
80	3.5	--	175
85	2.5	--	125
90	1.75	--	88
95	1.24	--	62

6. 9 µl hybridization buffer were mixed with 1 µl fluorescence labeled probe.

7. Then about 2 ml hybridization buffer were used to wet a piece of filter paper in a 50 ml tube. The slides were put on the top of the filter paper. The tube was closed and incubated at 37 °C for 4 hours for hybridization.
8. The washing buffer was prepared according to table 5-1. The concentration of NaCl must be changed according to the concentration of formamide in hybridization buffer. The prepared buffer was prewarmed at 48°C.
9. The slides were transferred to washing buffer, and washed at 48°C for 20 minutes. The slides must be kept horizontally before dipping into washing buffer to avoid the contamination of probes.
10. The probes were washed away with flowing water, then the slides were dried by flowing air.
11. The slides were then stained by 10 µl DAPI working solution (1 µg/ml in 1 x PBS) for 10 minutes on ice or at room temperature.
12. The slides were washed with distilled water for 30 seconds, followed by washing with ethanol for 30 seconds. The slides were then dried by blowing with air.
13. A drop of CITI flour (about 4 µl) was used to cover the samples on the slides, and then the cover glass (cleaned with ethanol) was sealed.
14. Observation with fluorescence microscopy.

2.2.2 Enzyme activity assays

The standard activity assay was carried out in 10 mM Tris-HCl buffer of pH 8.0. The reaction mixture contained 3.92 mM linolenic acid, 13.69 mM glutamine, and 0.1 *international unit* (I.U.) protein per liter. Protein and substrates were mixed on ice and then shaken at 220 rpm and 37°C. After reaction of 30 min, 50 µl aliquots were taken out and mixed with 145 µl methanol and 5 µl phenyldodecanoyl glutamine solution (100 µg/ml), which was used as internal standard. The mixture was then centrifuged at 13,000 g for 30 min. The supernatant was analyzed with HPLC-MS. Experiments of different pH, temperature, substrate specificity, *etc.* were carried out under slightly modified conditions as described in the text. Assays with living bacteria were carried out by addition of the substrates to the stationary-phase cultures and incubation for 4 h before sampling. Insect oral secretion samples were measured in the same way after collection.

30 µl samples were injected into HPLC (Agilent 1100), equipped with a RP-18 column (125 mm × 2 mm, 3 µm, Grom, Germany). The HPLC analysis was performed with a flow rate of 0.2 ml min⁻¹ using a binary gradient of two solvent systems (A: H₂O with 0.1 % AcOH, and B: MeCN with 0.1 % AcOH) starting with 100% A (3 min) programmed to 100% B in 27 min maintained for 15 min prior to equilibration to the starting conditions. 10 µl sample was injected into LCQ (Finnigan) equipped with atmospheric pressure chemical ionization (APCI) ion source. MS detection was performed as described before (Pohnert *et al.* 1999).

2.2.3 Protein purification

10 mM Tris-HCl buffer at pH 8.0 was used as the common buffer during purification, unless otherwise specified. The procedures were carried out at 0-4°C unless otherwise specified. Centrifugation was usually done at 14,000 g for 30 min. Concentration and

removal of the salt were performed with a Vivaspin 6 unit (50,000 molecular weight cut off (MWCO), Vivascience, UK).

Bacteria were cultivated on BHI plates 24 h prior to inoculation. Single colonies were inoculated into 100 ml liquid BHI medium. After shaking overnight at 220 rpm and 37°C, pre-cultures were mixed with 900 ml fresh medium, and cultured for further 18 h. When the OD₆₀₀ reached 2.4, bacteria were harvested by centrifugation at 9,000 g for 10 min. The cell pellet was washed twice with same volume of buffer, and resuspended in 50 ml buffer. This cell pellet was used for active protein purification by using an ÄKTA FPLC system (step 4 and 5).

Active protein purification

Buffers:

1. Common buffer

10 mM Tris-HCl buffer pH 8.0

2. Anion exchange chromatography buffer A

50 mM Tris-HCl, pH 8.0

3. Anion exchange chromatography buffer B

50 mM Tris-HCl, pH 8.0

1 M NaCl

4. Size exclusion chromatography buffer

20 mM Tris-HCl pH 8.0

50 mM NaCl

Procedure:

1. Cell lysis

Two kinds of mechanical lysis methods were employed (Cull and McHenry 1990). For large-scale preparations, 4 passes of French Press (SLM Aminco, USA) at 11,000 lb/in² were used. For small-scale preparations 6 minutes of sonication (Bandelin) at 60% of power was usually sufficient. The broken cells were subjected to ultra-centrifugation in order to remove the debris. The supernatant was carefully taken out.

2. Ammonium sulphate precipitation

Ammonium sulphate fractionation was started with slow addition of 17 ml of saturated solution to the cell crude extract to obtain 25% saturation. Next 15.9 g solid powder was added to reach 65% saturation. The fraction precipitating between 25% and 65% saturation was collected for further purification.

2. Ultrafilter cut-off

Viavaspin 20 ultrafiltration columns (100 kd, Vivascience AG) were used to concentrate the solution to 45 ml. The tubes were centrifuged at 6500 g for 30 minutes.

3. Desalting

The precipitate was dissolved in 20 ml buffer and loaded on a Vivaspin 20 ultracentrifugation unit (100,000 MWCO, Vivascience) for removal of small molecules. Concentrated protein solution was applied on Disposable Column PD-10 Desalting Column (Amersham Biosciences) to remove salt completely.

4. Anion exchange chromatography

The desalted protein solution was concentrated and loaded on an anion exchange chromatography column (ResourceQ 6 ml, Amersham Biosciences) connected to an ÄKTA fast protein liquid chromatography system (FPLC, Amersham Biosciences). The column was equilibrated with 18 ml of 50 mM Tris-HCl, pH 8.0. Washing and elution were performed at a flow rate of 2 ml/min. A two-step gradient elution was achieved by addition of 1 M NaCl in the same buffer (Figure 2-10 cyan line). The first 0.33 M NaCl was achieved by elution of 12 ml solvent. The second gradient of 0.33-0.45 M NaCl was eluted with 90 ml solvent. The active fraction eluted at about 0.35 M NaCl was collected.

5. Size exclusion chromatography

The concentrated solution was loaded on a size exclusion column (Superdex 200 HR 10/30, Amersham Biosciences). The column was eluted with 20 mM Tris-HCl pH 8.0 with 50 mM NaCl at a flow rate of 0.5 ml/min. The active fraction eluting at 10.75 ml was collected.

2.2.4 Protein characterization

The protein concentration was determined by the Bradford method (Bradford 1976) and the Micro BC Protein Quantification Kit (Uptima, Itechim, France) using bovine serum albumin (BSA) as standard.

The molecular mass of lyophilized protein was measured by MALDI-TOF (TofSpec 2E, Micromass, UK). The protein was dissolved in 1% trifluoroacetic acid. An 1.5 μ l aliquot was mixed with 1.5 μ l of a mixture of sinapinic acid in 0.1% trifluoroacetic acid and acetonitrile (6/4). 1.5 μ l of the mixed sample was dispensed on the target plate and dried in air at ambient temperature. Positive ions were observed by operation in a linear mode (acceleration voltage 20 kV) using delayed extraction. Desorption and ionization was accomplished by a nitrogen UV laser (337 nm, 4 ns pulses of 180 μ J). Spectra were obtained with averaging at least 50 consecutive scans. The mass charge ratio was externally calibrated using a mixture of porcine insulin, bovine ubiquitin, and BSA with the identical matrix under same analysis conditions.

The freshly purified protein was loaded on size exclusion chromatography column as mentioned above (see 3.4) to estimate its apparent Mr in solution. High Molecular Weight Calibration Kit for Electrophoresis (Amersham Biosciences) was employed as standard.

The flow-field-flow fractionation coupled multi-angle light scattering photometer (FFF-MALS, 18-angle DAWN-EOS, Wyatt Technology Corp.) and an Optilab DSP interferometric refractometer (Wyatt Technology Corp.) was employed to measure Mr of the native protein (Wittgren *et al.* 1998). Data were analyzed using the ASTRA software (Wyatt Technology Corp.) The FFF canal was equipped with a 10 kDa cutoff

Materials and Methods

10 x Running buffer 100 mL
dH₂O 890 mL

3. Membrane transfer buffer

SDS-Running buffer 250 mL
Methanol 200 mL
dH₂O 550 mL

4. Ponceau S Staining Solution

Ponceau S 1g

Acetic acid 50ml

Make up to 1L with ddH₂O

Store at 4°C. (Do not freeze.) Staining solution can be re-used up to 10 times.

5. 0.01M PBS

6. PBST_{0.1}

PBS 1 L

Tween-20 1 mL

(PBS + 0.1% Tween-20)

7. 5% BAS in PBS containing 0.1% NaN₃

Bovine Serum Albumin 5g

Sodium azide 0.1g

in 100 ml PBS

8. Phosphatase substrate solution

SIGMA FAST BCIP/NBT tablet dilute in 10 ml water

(Prepare just before use)

Procedure:

1. 8% native PAGE.

The gel was run at 30 mA (for the stack) and 40 mA (for the resolving gel) for about 2 h, until the tracking dye is ~1 cm from the bottom.

2. Membrane transfer

1. The PAGE gel was soaked in about 100 ml of membrane transfer buffer for 15 min with shaking.
2. Set up the gel transfer cassette.
3. Six pieces of filter papers were cut to a size of a little larger than the size of the membrane. This paper was soaked in membrane transfer buffer. Two pieces of the buffer saturated filter paper was put on the electrode of the gel transfer cassette. The air bubbles between the two pieces of paper were removed by rolling a glass rode on them.

4. A piece of PVDF membrane was activated by first dipping in pure methanol for 5 min, then in membrane transfer buffer for 5 min.
5. The third piece of buffer saturated filter paper was put on the two pieces in gel transfer cassette. Immediately, the activated PVDF membrane was spread on the top of the three pieces of filter paper. Then the gel was quickly put on the PVDF membrane. **Don't move the gel on the membrane.**
6. The other three pieces of filter paper were put on the top of this sandwich. The electric transfer was performed at 1 mA per cm² membrane for 75 min (larger proteins take longer time).
7. The membrane was taken out and washed in water briefly.
8. The membrane was stained with 1 x Ponceau S staining solution with gentle agitation for 15 min.
9. The membrane was scanned and the position of the molecular ladder was marked with pencil prior to removal of the Ponceau S away by washing with water or 1 x PBS.
10. The membrane was destained with distilled water until the background is clean. (The dye can be completely removed from the protein bands by continued washing.)

3. Blocking and detection

1. The membrane was blocked with 5% BSA in 1 x PBS overnight.
2. Before detection, the membrane was washed with PBS for 5 min.
3. The rabbit antiserum was diluted 10,000 times in PBS-T with 0.5% BSA. (This solution can be stored at 4°C for weeks).
4. The membrane and antibody solution was sealed in a plastic bag and incubated for 2 hour with shaking at 200 rpm.
5. The membrane was taken out and washed with 15 ml 1 x PBS for 5 minutes. This washing process was repeated 4 times.
6. The secondary antibody (Goat Anti-Rabbit IgG-Alkaline Phosphatase) was diluted 30,000 times in PBS-T. The membrane was incubated in the second antibody solution in a plastic bag for 1 hour and shaking at 200 rpm.
7. The membrane was washed with 15 ml 1 x PBS for 5 minutes. This washing process was repeated 4 times.
8. The strip was immersed in freshly prepared phosphatase substrate solution (SIGMA FAST BCIP/NBT tablet solution) for 10 min. The solution was gently shaken until the membrane was clearly stained.
9. The membrane was washed in distilled water for 5 min. The water must be changed after every 2 min.
10. The strip was dried between several pieces of filter paper, and stored in dark in a plastic sleeve afterwards.

Iron was detected by staining the gel with $K_3[Fe(CN)_6]$ as suggested by Papinutto *et al.* (2002). Inductive coupled plasma optical emission spectroscopy (ICP/OES) (Optima 3300 DV, Perkin-Elmer) with an axially arranged torch and cross-flow nebulizer was used to determine the exact iron content of the native protein. Simultaneous iron analyses were performed at 259.939 nm, 238.204 nm, 234.349 nm and 239.562 nm, respectively. Results from a two-point calibration 0.25 mg/l and 2.5 mg/l were averaged.

The isoelectric point (pI) was determined on an Immobiline DryStrip (24 cm, pH 4-7NL, Amersham Biosciences), using IPG buffer as suggested by the manufacturer. Strips were stained with 0.1% acid violet.

The DNA binding assay was carried out by running DNA on a 1% agarose gel (Almirón *et al.*, 1992). Freshly prepared protein and protein denatured by heating at 70°C for 20 min were mixed with the plasmid pET11a (Stratagene) DNA and incubated on ice for 2 min prior to loading onto the gel. The same amount of the Dps protein from *E. coli* was used as a control. The *dps* gene was induced and the protein was purified from *E. coli* carrying a plasmid ZK1100 as reported (Almirón *et al.*, 1992). Both Afp and Dps proteins were heated at 70°C for 20 min for the denaturation experiments.

The UV absorption was measured by a Spectrofluorometer (V-550, Jasco). The purified Afp protein was suspended in Tris-HCl buffer (pH 8.0 and 10.0). The absorption spectrum was scanned from 190 nm to 600 nm with data pitch 1.0 nm, band width 1.0 nm. The scanning speed was 200 nm/ min. The resulting spectra were subtracted by the Spectra Analysis software equipped with the instrument.

2.2.5 Protein sequencing

Protein was transferred from SDS-PAGE gel onto an Immuno-Blot PVDF membrane (Bio-Rad). The N-terminal was sequenced by Edman degradation on an amino acid sequencer (494A, Applied Biosystem model, USA).

After in-gel digestion (Shevchenko *et al.*, 1996), peptides were desalted and concentrated by ZipTip columns (C18-RP, Millipore, USA). ESI-MS/MS was performed on a Q-TOF 1.5 hybrid mass spectrometer (Micromass, UK) using “medium” nano ESI-capillaries (Protona, Denmark). The flow rate was regulated by the capillary voltage (650-850V). No desolvation gas was applied. Cone voltage was set to 45V. In MS-mode 25 scans (2.0 sec each) were averaged. In the MS/MS-mode argon was used as collision gas and the collision energy was varied between 16 and 42eV. Spectra were acquired in electrospray positive ionisation mode. The data were processed using MassLynx 3.5 (Micromass) and peptide sequences were calculated manually.

2.2.6 Genomic DNA isolation

Genomic DNA was extracted by a modified CTAB assistant procedure according to (Kieser *et al.*, 2000).

Genomic DNA isolation from *M. arborescens* strains

Solutions and buffers:

1. Lysozyme buffer:

- 10 mM Tris-HCl buffer (pH 8.0)
- 20 % sucrose (v/v)

2. Cell lysis buffer:

- 10 mM Tris-HCl buffer (pH 8.0)

1 mM EDTA
1 % SDS.

3. CTAB/NaCl:

10% CTAB (cetyltrimethylammonium bromide, Sigma H5882)
0.7 M NaCl
(add CTAB slowly to 65°C NaCl solution, autoclave.)

4. 3 M sodium acetate (pH 4.8)

Dissovre:
Sodium acetate trihydrate 40.8 g
in water, adjust volume to 100 ml.

Procedure:

1. The bacteria were cultured in 3 ml BHI or LB liquid medium overnight. The cells were precipitated by centrifugation and washed by 3 ml of 0.1 x SSC.
2. The cells were resuspended in 100 µl lysozyme buffer. 0.25 mg lysozyme was added and mixed with the cell suspension. The mixture was incubated at 37°C for 45 min.
3. Additional 0.25 mg lysozyme were added to the suspension. The mixture was incubated at 37 °C for 45 min.
4. The cell suspension was mixed with 900 µl lysis buffer, 0.15 mg proteinase K, and 0.12 mg RNase. The mixture was incubated at 55 °C for 45 min.
5. The bacterial lysate was mixed with 200 µl prewarmed 5 M NaCl solution (55°C), resulting in a final NaCl concentration of 0.8 M.
6. 150 µl CTAB/NaCl solution was added into the suspension and mixed thoroughly. The mixture was incubated at 55°C for 10 min.
7. The mixture must be cooled to 37°C prior to the following steps.
8. The cell lysate was mixed with ¼ volume of phenol and chloroform (1:1) and centrifuged at 13000 rpm for 7 min. The supernatant was carefully transferred to a new 2 ml tube after centrifugation.
9. The same volume of chloroform was mixed with the supernatant to remove protein and phenol. The mixture was centrifuged at 13000 rpm for 7 min. The supernatant was carefully transferred into another 2 ml tube. Repeat this process 3 times.
10. The supernatant was then mixed with 70% volume of isopropanol and 0.1 volume 3 M sodium acetate (pH 4.8) to precipitate the DNA.
11. The wool-like DNA was spooled out with a pipette tip, and washed with 70% cold ethanol (-20°C).
12. The DNA was completely air dried, and then dissolved in 50 µl water.

2.2.7 Gene cloning

Materials and Methods

Degenerate primers designed from peptide sequences and Shine-Dalgarno sequence or analog were used in PCR (SD-PCR for detail see 3.7.2.2 and 4.7.2) to clone the corresponding gene (Table 2-4).

Table 2-4. Oligonucleotide primers

PCR round	Primer	Sequence 5'→3'	Target sequence	Working concentration
1	Upper: asmf1	ACIGA (C/T)ACIAA (C/T)AT (A/T/C)ACIAC	Peptide eda-1	4.0 μM
	Upper: asmf2	ACIGCIGA (C/T)CCIIGA (A/G)GT		
	Lower: asmr2	CGIAC (A/G)TTCCA (A/G)TG	Peptide msa-2	1.0 μM
2	Upper: sdu	I (A/C) (A/G/T/C)GAGG	Shine-Dalgarno sequence or analog	6.0 μM
	Lower: asmr3	GTTCAGTGCGCCTGCTTGC	1st round DNA	0.4 μM
3	Lower: asmr4	CGACCGAGTGAGGAACTGC	2nd round DNA	0.4 μM
	Upper: asmf4	ACAGCTCGCCGATGGTCACA		
	Upper: asmf5	ATCACCACCCCGCCCTCAC		
	Upper: asmf6	ATGCAGGCGCTCGTCGTCAAC	Peptide msa-4	4.0 μM
	Lower: asmr5	TTIAI (C/T)TC (A/G)TC (C/T)TGCCA (C/T)TG		

Working concentration is their concentration in 50 μl PCR reaction.

A 650-bp fragment representing the first part of the *afp* gene and an upstream sequence was amplified by 3 rounds of degenerate PCR. The primers and their amount in 50 μl PCR reaction are listed in Table 2-4. The reaction also contained 100 ng genomic DNA, 2.5 units Taq polymerase (Invitrogen), 2.5 mM MgCl₂, 300 μM dNTP and the buffer supplied by the manufacturer. Template DNA was melted at 94 °C for 3 min, followed by 40 cycles of 94 °C for 45 sec, 55 °C for 30 sec, 72 °C for 30 sec. Final elongation was achieved at 72 °C for 1 min.

The cloned fragment was used as probe in Southern blotting.

Detection of restriction-fragments by radioactive probe

Solutions and buffers:

1. SSC Buffer (20 x stock):

175.3 g NaCl

88.2 g sodium citrate

(Dissolve in 800 ml of water. Adjust to pH 7.2 with a few drops of a 10 N NaOH solution. Then adjust to 1 liter with water. Sterilize by autoclave.)

2. Depurinate solution

0.25 N HCl

3. Transfer solution

0.5 N NaOH

4. 10 x SSC

5. 2 x SSC

6. Church Buffer:

0.5 M Sodium phosphate buffer (pH 7.2)

7% SDS

1mM EDTA.

(Na₂HPO₄ 46.8g; NaH₂PO₄ 21.8g; add water to 800 ml, adjust pH to 7.2, add 70 g SDS, 2 ml 500 mM EDTA. Just to 1000 ml.)

7. Buffer A

100mM Tris-HCl (pH7.5)

600mM NaCl

Tris base 12.1 g

NaCl 35.064 g

(dilute to 800 ml water, mixed with 6.28 ml concentrated HCl, and adjust volume to 1000 ml with water, autoclave.)

8. SPS buffer:

40 mM sodium phosphate buffer (pH 7.2)

1% SDS

Na₂HPO₄ 3.744g

NaH₂PO₄ 1.744g

(add water to a final volume of 800 ml, adjust pH to 7.2, add 10 g SDS, adjust volume to 1000 ml with water.)

Procedure:

1. Restriction-fragments preparation:

1. The genomic DNA was cut with different restriction enzymes *Bam*H I, *Kpn* I, *Pst* I, *Not* I, *Sac* I, *Sac* II, *Xho* I.
2. The DNA fragments were loaded on a 0.8 % agarose gel. The gel was run at 30V overnight.

2. Transfer to the membrane:

1. The gel was depurinated in 0.25 N HCl for 15 min with gentle shaking.
2. The gel was washed twice with deionized water after removal of the 0.25 N HCl solution.
3. The gel was denatured in 0.5 N NaOH for 30 min (cover the gel with 0.5 N NaOH by gently shaking).
4. The Hybond-N⁺ membrane (Amersham Biosciences) was soaked 10 x SSC in the Vacuum transfer cassette (Bio-Rad).
5. The gel was immediately put onto the nylon membrane.
6. Vacuum transfer was performed at 2.1 x 10⁵ Pa for 90 min as suggested by Bio-Rad protocol.
7. The membrane was washed in 2 x SSC for 5 min.
8. The membrane was air dried between two sheets of filter paper for 2 hours.

3. Labeling the probe:

1. The PCR product was radioactively labeled with NEBlot Kit as suggested by protocol (New England Biolabs)

Materials and Methods

PCR product	15 μ l
Water	18 μ l
Heat to 99°C for 5 min to denature, then quickly place on ice. Add	
labeling buffer	(5 μ l with random octadeoxyribonucleotides)
dATP	2 μ l
dTTP	2 μ l
dGTP	2 μ l
Klenow	1 μ l
α^{32} P dCTP	5 μ l (1.85 MBq, 50 μ Ci)
<hr/>	
total	50 μ l
37°C for 1 hour.	

2. The labeled probe was mixed 5 ml hybridization buffer.

4. Hybridization:

1. The air dried nylon membrane was UV crosslinked for 10 min.
2. The liquid block (NEB) was 20 times diluted by Church Buffer.
3. 50 ml hybridization buffer was preheated at 65°C. The blots were placed in the buffer. Prehybridization was performed in Church buffer at 65°C for 120 min with constant gentle agitation.
4. The probe was briefly denatured at 99°C and placed into Church buffer, (**Avoid placing it directly on the blots**) for hybridization.
5. The blots were hybridized at 65°C overnight with gentle agitation.
6. The Membrane was washed three times in SPS buffer for 10 min at room temperature, and exposed to Kodak scientific imaging film (Sigma) for 4 hours.

A 2.2-kb fragment in the *NotI* digested DNA and a 2.3-kb fragment in the *SacI* digested DNA showed hybridization signals. The digested fragments corresponding to the hybridization signals were purified from gel and ligated into gel-isolated pBluescript II SK (+) Phagemid vector (Stratagene) digested by the same enzymes. The ligation mixture was transformed into *E. coli* INV α F' competent cells (Invitrogen), and screened by ECL Kit.

E. coli library screening

Solutions and buffers:

(Some solutions used in this step are identical to those used for restriction-fragments radioactive hybridization)

1. Solution I (denaturation solution)

0.5 M NaOH
1.5 M NaCl
0.1% SDS

Materials and Methods

(10 g NaOH, 43.83 g NaCl, 0.5 g SDS, dissolve in water, adjust volume to 500 ml.)

2. Solution II (Neutralization solution)

3 M NaCl

0.3 M Tris pH 7.5

(87.66 g NaCl, 18.155 g Tris, mixed with 9.75 ml glacial HCl, adjust volume to 500 ml.)

3. 2 x SSC

4. Church buffer:

As restriction-fragments radioactive hybridization

5. Stringency wash solution (1):

1 x SSC

0.1% SDS (w/v)

(50 ml 20 x SSC in 950 ml water, add 1 g SDS.)

6. Stringency wash solution (2):

0.5 x SSC

0.1% SDS (w/v)

(25 ml 20 x SSC in 975 ml water. Add 1 g SDS.)

7. Buffer A

8. SPS buffer

Procedure:

1. Membrane preparation

1. The transformed *E. coli* colonies were cultivated overnight. 10 μ l of the overnight culture were pipetted on Hybond-N⁺ membrane (Amersham Biosciences). The membrane was placed on the top of filter papers. The membrane was dried in the air.
2. A piece of filter paper was soaked in solution I to saturation. The membrane was put on the top of the filter paper for 15 min.
3. Another piece of filter paper was soaked in solution II. The membrane was treated by solution II for 5 min on the top of the filter paper.
4. The membrane was then transferred on to the filter paper in 2 x SSC solution and soaked for 15 min.
5. The membrane was taken out and dried between two pieces of filter paper.
6. The DNA was crosslinked to the membrane by UV radiation for 10 min. The membrane was then stored at 4°C.

2. Probe labeling [as suggested by the ECL random prime labeling and detection systems (Amersham Life Science)]

Materials and Methods

1. 20 μ l PCR product was denatured by heating at 99°C for 5 min. Denatured template was immediately transferred on ice.
2. In a 1.5 ml tube on ice, add the following: (**Do not change order**)

Water	29 μ l	
Nucleotide mix		10 μ l
Primer	5 μ l	
DNA	5 μ l	
Enzyme solution		1 μ l

3. The reagents were mixed gently by pipetting up and down, and then collected at the bottom of the tube by brief spinning. The mixture was incubated at 37°C for 1 h.
4. The reaction was terminated by addition of 0.5 M EDTA to a final concentration of 20 mM. The probe was stored at -20°C in dark.

3. Hybridization

1. The liquid block (NEB) was 20 x diluted with Church Buffer.
2. 50 ml hybridization buffer was preheated to 60°C.
3. The blot was placed in the buffer. Prehybridization was done in Church buffer at 60°C for 60 min with constant gentle agitation.
4. The probe was denatured briefly at 99°C and then added in Church buffer. (avoid placing it directly on the blots, and mix gentle.)
5. Hybridization was performed at 60°C overnight with gentle agitation.
6. The stringency wash buffer (1) was preheated to 60°C. The blots were carefully transferred to this solution, and washed for at least 15 min with gentle agitation.
7. The membrane was further washed in stringency wash buffer (2) for 15 min.

4. Blocking, antibody incubation, washes and detection (as suggested by protocol)

4 clones, designed 1E12, 1G12, 3F11, and 4C5 were hybridized by the probe. The plasmids were isolated by MINIpip according to Sambrook *et al.* (2001) and sequenced.

Plasmid extraction by MINIpip

Solutions and buffers:

(According 'molecular clone', 1.25)

1. P I solution:

50 mM glucose

25 mM Tris-HCl buffer pH 8.0

10 mM EDTA

(autoclaved for 15 minutes on liquid cycle, after cooling, add 100 μ g/ml

RNase, store at 4°C.)

2. P II solution:

0.2 N NaOH (freshly diluted from 10 N stock)
1 % SDS
(store at room temperature)

3. P III solution:

Mix:
5 M potassium acetate 60 ml
glacial acetic acid 11.5 ml
with 28.5 ml water
(The final concentration is 3 M potassium and 5 M acetate)

Procedure:

1. 1.8 ml of the overnight *E. coli* culture were placed in a 2.0 ml tube, and centrifuged at 10,000 rpm for 30 sec.
2. The supernatant was discarded.
3. The pellet was briefly spun at 10,000 rpm for 30 seconds.
4. The liquid medium was pipetted out.
5. The pellet was resuspended in 200 μ l P I solution by gentle vortex.
6. 200 μ l P II solution was added. The tubes were inverted upside down for 7 times.
7. The solution was mixed with 200 μ l P III solution and mixed up and down for 7 times.
8. The tubes were centrifuged at 13000 rpm for 7 minutes.
9. The supernatant was carefully transferred into another tube. The protein pellet was discarded.
10. The supernatant was mixed with 500 μ l chloroform thoroughly followed by centrifugation at 13000 rpm for 5 minutes.
11. (Cut off the top of blue pipette tip). The supernatant was carefully sucked out and placed into a new tube.
12. The supernatant was thoroughly mixed with 350 μ l isopropanol and kept at -20°C for at least 10 minutes.
13. The DNA was precipitated by centrifugation at 14000 rpm for at least 15 minutes.
14. Discards the supernatant. The pellet was washed with 70 % ice-cold ethanol.
15. The pellet was dried in vacuum for 15 minutes.
16. The plasmid was dissolved in 50 μ l water.

Sequencing data were processed with Lasergene software (DNASTAR Inc.). Homologous protein sequences (to the deduced amino acid sequence) were searched by BLAST at the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov). The protein alignment was performed with free software BioEdit 5.0.6 (North Carolina State University) and ClustalX 1.81 (<ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/>).

2.2.8 Gene expression

Materials and Methods

To detect the *afp* gene in *M. arborescens* SE14 and *M. arborescens* DSM20754, a 506-bp DNA fragment containing the total *afp* gene was amplified by PCR from genomic DNA by using the gene-specific primers Afc1:

5'-GACATGACCGACACCAACATC-3'

and Arc3:

5'-CCGTCGGCGCCGGGTATTAC-3'

and primer asmf4 (Table 2-4).

The reaction contained 100 ng DNA, 2.5 units Taq polymerase (Invitrogen), 2.5 mM MgCl₂, 300 μM dNTP and the buffer supplied by the manufacturer. Template DNA was melted at 94 °C for 3 min, followed by 40 cycles of 94 °C for 45 sec, 55 °C for 30 sec, and 72 °C for 2 min. Final elongation was done at 72 °C for 3 min.

When PCR was performed with Taq polymerase from New England Biolab (NEB), 300 μM dNTP was mixed with 1 x buffer supplied by NEB in a total volume of 50 μl. The PCR thermo-cycle was the same as above, except that the extension was performed at 75°C for 2 min.

In the process of construction of an *afp* gene expression system, all plasmids were propagated by transformation into *E. coli* INVαF'. PCR reactions were performed with *Pfx* (Invitrogen). In order to introduce a restriction site at proper position, the 506-bp DNA fragment containing the total *afp* gene was amplified by PCR from genomic DNA of *M. arborescens* SE14 using primers Afc1 and Arc3.

Reaction mixtures contained 1.5 mM MgSO₄, 200 μM dNTP, 2.5 units *Pfx* (Invitrogen) and its buffer supplied by producer. Thermal cycling was initiated at 96°C for 5 min to lyse the cells, followed by denaturation at 96°C for 30 sec, annealing at 55°C for 45 min, and extension was performed at 68°C for 2 min, for a total of 35 cycles. Final elongation was achieved at 68°C for 3 min. The PCR product of right size was sliced from 1.2 % agarose gel and purified by GFX Purification Kit (Amersham Biosciences). The 100 times diluted DNA fragment was used as DNA template for further restriction site introducing PCR.

For cloning, primers are And1:

5'-CCCATCCATATGACCGACACCAACATC-3'

and Abam3:

5'-AGTCAGGATCCGTCGGCGCCGGGTATTAC-3'

The restriction sites for *Nde*I and *Bam*HI are underlined. PCR was performed as above. Purify the PCR product from gel by GFX kit prior to digestion by the enzymes. The DNA was eluted with 45 μl water.

The purified fragment was double digested and ligated into pET11a which was digested with same restriction enzymes. The plasmid was signed pEM1i2, propagated by transformation into *E. coli* INVαF' (Invitrogen). pEM1i2 was transformed into *E. coli* BL21-CodonPlus (DE3)-RP competent cell (Stratagene) for protein expression.

Protocol for gene expression in *E. coli*

1. Vector preparation

1. Vector pET11a was digested by the restriction enzymes *NdeI* and *BamHI*.
2. After digestion, the phagemid vector was heated at 65°C for 20 min to inactivate the enzymes. The vector was dephosphorylated by Shrimp Alkaline Phosphatase (SAP). 1 µl SAP was mixed with the vector solution, and incubated at 37°C for 30 sec. Then another 1 µl SAP was added and incubated for other 30 min.
3. The digested fragments were purified from agarose gel by GFX purification Kit, and eluted with 20 µl water.

2. Digestion of PCR product

1. PCR product was digested with the same enzymes.
2. The digested fragments were purified by GFX purification Kit and eluted with 20 µl water.

3. Ligation

1. The digested PCR fragment was ligated into the digested vector by T4 DNA ligase (NEB)
2. The mixture was incubated at 15°C overnight.

4. Transformation

1. XL10-Gold β-mercaptoethanol (Stratagene) was diluted 10 times with water.
2. 1.0 µl mercaptoethanol was mixed with 100 µl BL21-CodonPlus(DE3)-RP competent cells with gentle swirling. The cell suspension was put on ice for 10 min.
3. 3 µl ligated plasmid solution were added to 50 µl competent cell. The cells were incubated on ice 30 min and then heat shocked at 42°C for 30 sec. The cells were returned to ice after heatshock and incubated for 5 min.
4. The cells were mixed with 125 µl ice cold SOC medium and incubated at 37°C for 1 hour with shaking at 220 rpm.
5. The transformed cells were spread on LB plate with 100 µg/ml carbenicillin and incubated at 37°C overnight.

5. Induction

1. The colonies of transformed *E. coli* were inoculated into 3 ml LB broth with 100µg/ml carbenicillin and 50 µl/ml chloramphenicol. The culture was incubated at 37°C overnight with shaking at 220 rpm.
2. 300µl aliquots of the overnight culture was inoculated into 6 ml new medium with antibiotics. The bacteria were cultivated for 2 hour until cell density reached OD₆₀₀ of 0.8.
3. The protein expression was induced by addition of 6 µl 1M IPTG into culture. The culture was cultivated at 37°C for 2 hours with shaking, or at 20°C for 20 hours with shaking.
4. The cells were collected by centrifugation at 8000 rpm for 10 min.

The plasmid pEM1i2 was double digested with *NdeI* and *XbaI*. The 6.2-kb fragment was purified from gel. The linearized plasmid pEM1i2 was ligated with a 3.5 kb fragment from plasmid pHN15 (WERNER BioAgents, Jena, Germany) digested with the same enzymes. The produced mosaic plasmid pHER was propagated by transformation into *E. coli* INVαF', and further digested with *EcoRI* and *HindIII*. In this plasmid (see appendix II), the *afp* gene was engineered into an expression cassette, in which the *afp* gene was ligated inbetween a *veg* promoter and a T7 terminator. The 1.6 kb *afp* gene expression cassette was sliced by double digestion of *EcoRI* and *HindIII* and purified. The cassette was then inserted into the unique *EcoRI* and *HindIII* sites of plasmid pRM5 (McDaniel *et al.* 1993) by replacement of the *act* gene cluster. The expression plasmid was named pRMM. A control plasmid, pRMC was constructed by religation of *BamHI* digested pRMM, where *afp* gene was totally removed (see appendix II). The correct sequence and direction of the inserts was proven by both digestion of restriction enzymes listed on the map and sequencing with the universal primers, T7-tem:

5'-GCTAGTTATTGCTCAGCGG-3'

and vegprom:

5'-GAGGTGGATGCAATGGCGAAG-3'

The plasmids pRMM and pRMC were transformed into the methylation-deficient *E. coli* INV110. Isolated unmethylated plasmids were introduced into *M. arborescens* DSM20754 by a modified PEG-assisted protoplast transformation procedure (Kieser *et al.* 2000).

Transformation of *M. arborescens* DSM20754

Solutions and buffers:

[According ref. (Kieser *et al.* 2000)]

1. 10.3 % sucrose

2. Trace element solution

ZnCl ₂	40 mg
FeCl ₃ ·6H ₂ O	200 mg
CuCl ₂ ·2H ₂ O	10 mg
MnCl ₂ ·4H ₂ O	10 mg
Na ₂ B ₄ O ₇ ·10H ₂ O	10 mg
(NH ₄) ₆ Mo ₇ O ₂₄ ·4H ₂ O	10 mg
Dissolve in distilled H ₂ O and adjust to 1 l	

3. TES buffer

5.73% TES (N-Tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid)
(Adjust pH to 7.2 by addition of 2 M NaOH)

4. Tris-maleic acid buffer

1M Tris base solution
adjust to pH 8.0 by addition of solid maleic acid with rapid stirring

5. P (protoplast) buffer

Sucrose	103 g
K ₂ SO ₄	0.25 g
MgCl ₂ ·6H ₂ O	2.02 g
Trace element solution	2 ml
Add water to	800 ml

Dispense in 80 ml aliquots and autoclave. Before use, complete the buffer by adding into each 80 ml.

(Do Not Change Order!)

KH ₂ PO ₄ (0.5%)	1 ml	
CaCl ₂ ·2H ₂ O (3.68%)		10 ml
TES buffer	10 ml	

6. T (transformation) buffer

10.3% Sucrose	25 ml (2.575 g)
Water	75 ml
Trace element solution	0.2 ml
2.5% K ₂ SO ₄	1 ml (25 mg)

To this 9.3 ml solution add:

CaCl ₂ (5M)	0.2 ml
Tris-maleic acid buffer	0.5 ml

(For use, add 3 X volume of the solution to 1 X by weight polyethylene glycol (PEG)1000, previously sterilized by autoclaving)

Procedure:

1. Protoplast preparation

1. A colony of *M. arborescens* DSM20754 was inoculated in 2 ml BHI medium and grown at 37°C overnight.
2. This preculture was then used as inoculant for 20 ml new BHI liquid medium. The bacteria were cultivated for 15 hours in 40 ml falcon tube (**Cap must be open**).
3. Bacteria were collected by centrifugation at 1000 g for 10 min.
4. The pellet was washed twice with 15 ml 10.3% sucrose.
5. The pellet was resuspended in 4 ml lysozyme buffer with 2.5 mg/ml lysozyme and incubated at 37°C for 30 min. Same amount of new lysozyme were added and mixed gently. The mixture was incubated at 37°C for additional 30 min.
6. 5 ml P buffer was added and mixed by gently pipetting up and down.
7. Sediment protoplast gently by brief spinning (no more than 10 sec).
8. The supernatant was discarded. The sedimented protoplasts were resuspended in 1 ml P buffer.
9. The protoplast suspension was aliquoted to 58 µl, and put in 1.5 ml tube for immediate transformation or for storage at – 80°C.
10. To freeze the protoplasts for storage, the tubes were placed in ice in a plastic beaker. The beaker was put at – 80°C overnight. The frozen protoplast aliquots

were stored at -80°C . They can be stored for 3 month without significant loss of viability.

2. Plasmid transformation

1. Take 50 μl fresh protoplast solution into 1.5 ml tube.
2. The plasmid solution containing about 500 ng DNA was added into the 50 μl protoplast suspension and mixed immediately by gently tapping the tube.
3. If frozen protoplasts were used, the protoplasts should be thawed quickly by heating at 25°C with shaking at 200 rpm (or thaw it in 25°C running water). Immediately after thawing, the plasmid solution was added and mixed.
4. The mixture was put on ice for 45 sec for plasmid adhesion.
5. 200 μl T buffer were mixed with protoplast suspension by pipetting up and down once using a tip-cut pipette tip.
6. **(As soon as possible)** 1 ml P buffer working solution was added into the suspension. The protoplasts were spun down for 5 sec. The supernatant was taken out.
7. The protoplasts were resuspended in 200 μl BHI liquid medium. The suspension was incubated at 22°C for 2 min with shaking at 300 rpm.
8. The suspension was spread on BHI plates with thiostrepton 50 $\mu\text{g}/\text{ml}$.
9. The plates were incubated at 37°C for 24 h until colonies emerged on the agar surface.

Transformed cells of *M. arborescens* DSM20754 were grown in BHI liquid medium supplemented with 50 $\mu\text{g}/\text{ml}$ thiostrepton at 37°C overnight with shaking at 220 rpm. Four colonies were selected to detect the *afp* gene and amide forming activity. The whole *afp* gene was confirmed by PCR amplification using primers Afc1 and Arc3. Plasmids were retrieved from transformed bacteria by using a procedure similar to MINIpiprep after digestion of cell wall with lysozyme as mentioned in DNA extraction protocol (see 5.8). The activity of expressed protein was performed as mentioned before (see 5.3).

3 Results

3.1 Screening of the gut bacteria

3.1.1 Catalytic activities

Some strains, representing diverse bacterial groups, were isolated from the gut of BAW (provided by Prof. K. Dettner), and were selected to study the biosynthesis of *N*-acyl glutamines in detail. When free linolenic acid and glutamine were supplied as substrates to overnight bacterial cultures, some of them were able to catalyze the formation of linolenoyl glutamine (Figure 3-1).

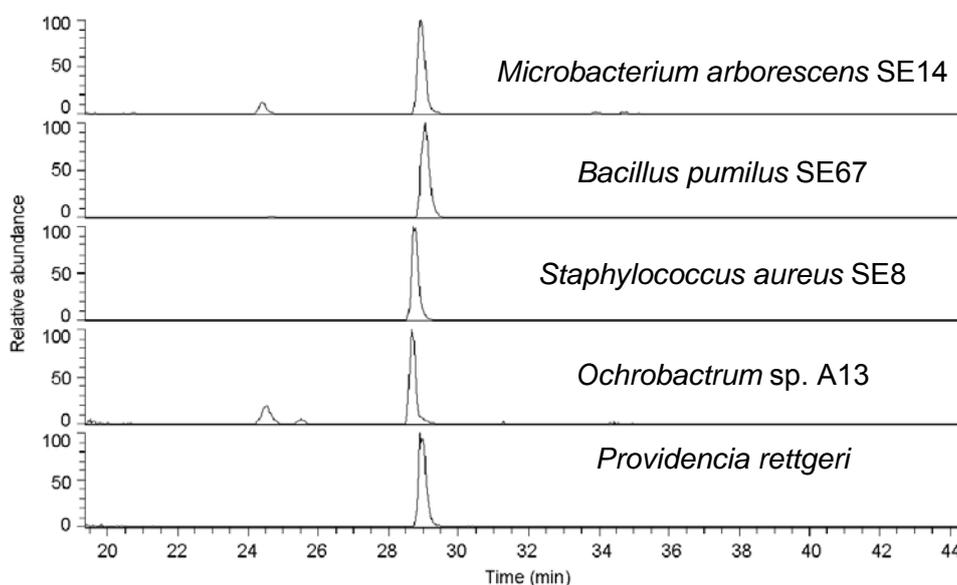


FIGURE 3-1 HPLC-MS chromatographs of linolenoyl glutamine produced by some selected gut bacteria. The substrates, free linolenic acid and glutamine were added to overnight liquid cultures. After 4 hours incubation with shaking at 37 °C, samples were collected and measured. The mass charge ratio shown here is $m/z = 406.0-408.0$.

In this research, two Gram-negative species, *Providencia rettgeri* and *Ochrobactrum* sp. A13, which have been reported to catalyze linolenoyl glutamine production (Spiteller *et al.* 2000) were compared. Three others are Gram-positive. *Staphylococcus aureus* SE8 and *Bacillus pumilus* SE67 contain low GC DNA in their genome, while *Microbacterium arborescens* SE14 is a high GC species.

Different bacteria seem to use different enzymatic systems to produce *N*-acyl glutamine according to the high-pressure liquid chromatography (HPLC) coupled mass spectrometry (MS) measurements (Figure 3-1). In addition to the typical *N*-linolenoyl glutamine with retention time of *ca.* 29.0 min, a small amount of compound with same mass/charge ratio (m/z) was detected in the *M. arborescens* SE14 preparation when

Results

supplied with the same substrates. The lower polarity of this compound leads to its elution at *ca.* 24.5 min. In the *Ochrobactrum* sp. A13 preparation, a third peak of *m/z* of 407 was detected with retention time of 25.5 min. All three peaks showed identical MS profiles. Whether they are isomers is not known and needs further investigation.

The enzymatic activities were located in different protein fractions prepared from cultures of different bacterial strains. Both, *P. rettgeri* and *S. aureus* SE8, do not produce the additional compounds and their catalytic activities were detected in different extract fractions. The activity of *P. rettgeri* as well as those of most of the other bacteria was found mainly in preparations of the bacterial cell (Figure 3-2).

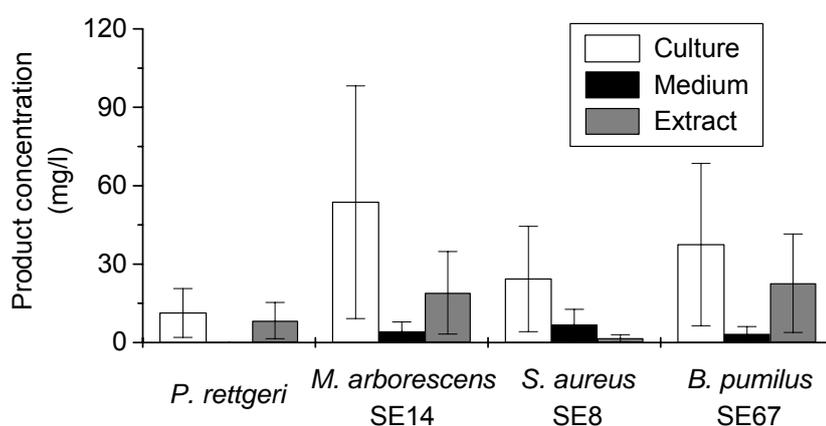


FIGURE 3-2. *N*-linolenoyl glutamine produced by different samples of overnight bacterial cultures. Cells were removed from medium by ultracentrifugation. Extract corresponds to the soluble fraction of broken cells. In the culture control, cells in the medium were disrupted by sonication. After an incubation of 30 min at 37 °C, samples were analysed by HPLC. Mean values of concentrations were calculated (n = 3).

When *S. aureus* SE8 cells were removed from the medium, about 30% of the activity remained in the broth supernatant. Only a trace activity was detected in the cell extract (Figure 3-2) suggesting that the enzyme had been secreted by the bacterium into surrounding medium. It is known that Staphylococci, especially *S. aureus*, produce a wide variety of exoproteins. They are accessory proteins, which are synthesized at the end of the exponential growth and during the early stationary phase (Novick 1993).

3.1.2 Growth stage-dependent activities

Because the accessory proteins from *S. aureus* SE8 are usually synthesized during stationary growth (Novick 1993), the enzymatic activities of the other strains were determined throughout their whole growth phase. Interestingly, the three tested bacteria mainly produce *N*-linolenoyl glutamine within a quite short time window during early stationary phase growth, although they displayed different growth kinetics (Figure 3-3). Not only the pathogenic proteobacterium *P. rettgeri* grows much faster than the two

Results

Gram-positive strains, but its activity also rose earlier than that of the Gram-positive strains.

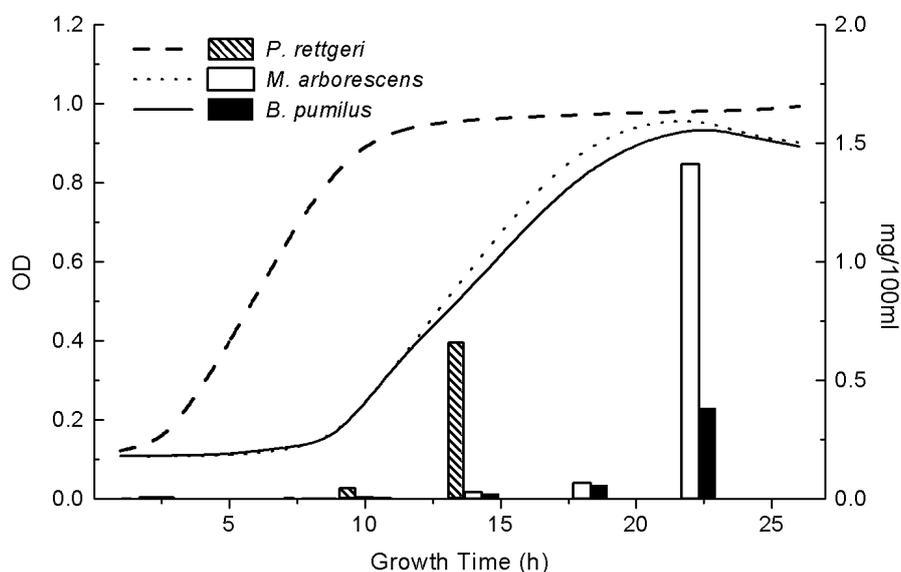


FIGURE 3-3. *N*-linolenoyl glutamine production by different bacteria at different growth stages. Bacterial cell density was monitored by measuring optical density at 600 nm (OD_{600}) in 96-well microplates. Total enzyme activity of 1 ml culture was tested. Samples were collected and measured every 4 hours ($n \geq 2$). Single colonies of bacterial strains were inoculated into 1 ml liquid brain heart infusion medium (BHI). Substrates, namely linolenic acid and glutamine, were added at different times after inoculation. After 4 hours, samples were collected and measured. To draw the growth curve, bacteria were grown in 1 ml culture, 100 μ l was taken at given interval dropped into 96-well microplates, and the cell density was monitored at 600 nm with a SpectraMAX 250 (molecular Devices Corp.)

The *N*-acyl glutamine productivity varies greatly among tested species. The activity of *Ochrobactrum* sp. A13 culture is 10 times lower than that of the other Gram-negative strain *P. rettgeri*. Therefore its activity has not been monitored in the assays of different cell fractions and growth stages. *M. arborescens* SE14 overnight culture showed the highest activity of all the tested bacterial isolates (Figure 3-2 and 3-3). This strain was therefore chosen to further investigate the active protein and corresponding gene.

3.2 Evidence of selectively enhanced activity

The most active gut bacterial strain was identified as *M. arborescens* based on both classical bacteriological tests (Holt 1993) and its 16S rDNA sequence (NCBI, AY649756, first determined by Dr. Piel). This sequence is 99% identical to that of *M. arborescens* DSM 20754 (EMBL, X77443), the type strain of the *Microbacterium* genus (Rainey *et al.* 1994). The strain was isolated from the gut of a *Spodoptera exigua* larva, and named *M. arborescens* SE14. To test whether this activity is a species character, the reference strain *M. arborescens* DSM20754 (German Collection of Microorganism and Cell Cultures, DSM, Braunschweig, Germany) was incubated under the same conditions.

Results

As shown in Figure 3-4, the amount of *N*-acyl amino acid produced by *M. arborescens* SE14 is 20 times higher than that of the reference strain. The difference between the two strains becomes also evident on the relative production ratio of the unknown compound eluting at 24.5 min and the typical α -*N*-linolenoyl glutamine with respect to peak area. The same amount of this compound and of the typical α -*N*-linolenoyl glutamine were produced by *M. arborescens* DSM20754 (Figure 3-4).

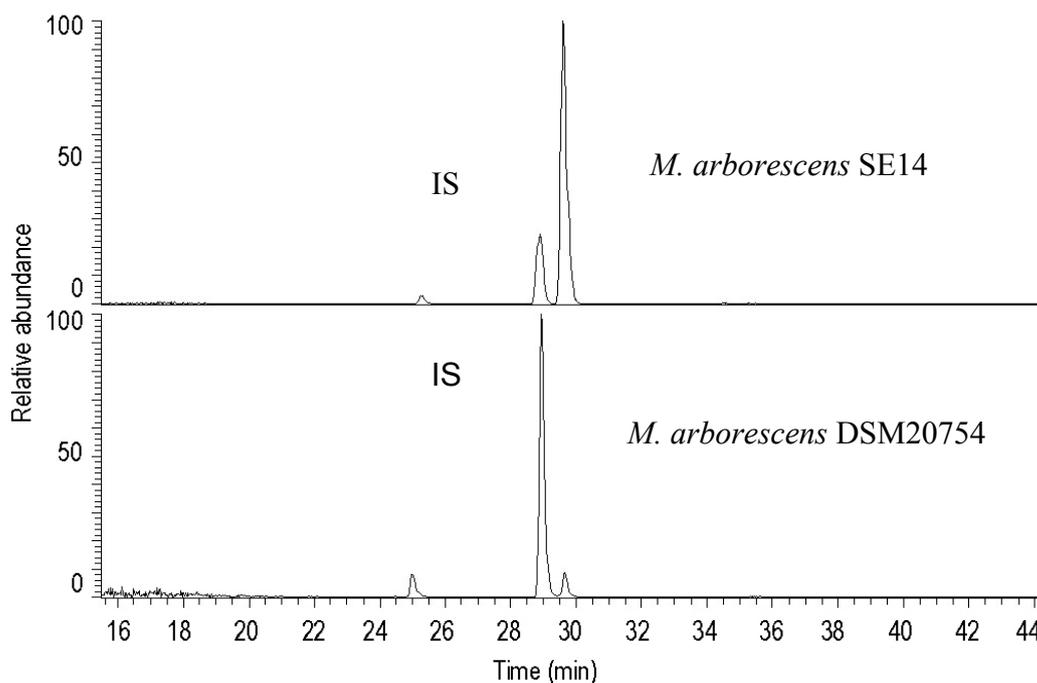


FIGURE 3-4. The enzymatic activity of *M. arborescens* SE14 and DSM20754. Bacteria were inoculated 18 h before addition of the substrates, then incubated at 37°C with shaking at 220 rpm for 4 h before the linolenoyl glutamine was determined. IS, internal standard.

Microbacterium is a high GC Gram-positive bacterial genus. They are aerobic, catalase- and oxidase-negative, irregularly rod-shaped bacteria (Collins and Bradbury 1991). No morphological difference has been observed under the microscope between the two *M. arborescens* strains (Figure 3-5). Only the cell density of *M. arborescens* SE14 during the stationary phase growth is higher than that of *M. arborescens* DSM20754. This has been confirmed by photometer measurements. Based on 4 separate assays, the average OD_{600} of *M. arborescens* SE14 at stationary growth is 3.3258 ± 0.2680 , while that of *M. arborescens* DSM20754 is 2.5325 ± 0.2411 .

Pink-colour pigmentation has been observed in later stationary growth of both strains. This pigment emits yellow fluorescence when excited by laser irradiation at 355 nm (Figure 3-5 A). The interference of the fluorescence from DAPI and that from the pigment resulted in the yellow green colour in Figure 3-5 B. Although *M. arborescens* SE14 grows faster than *M. arborescens* DSM20754 (see 3.3), the pigmentation of this strain is delayed somehow.

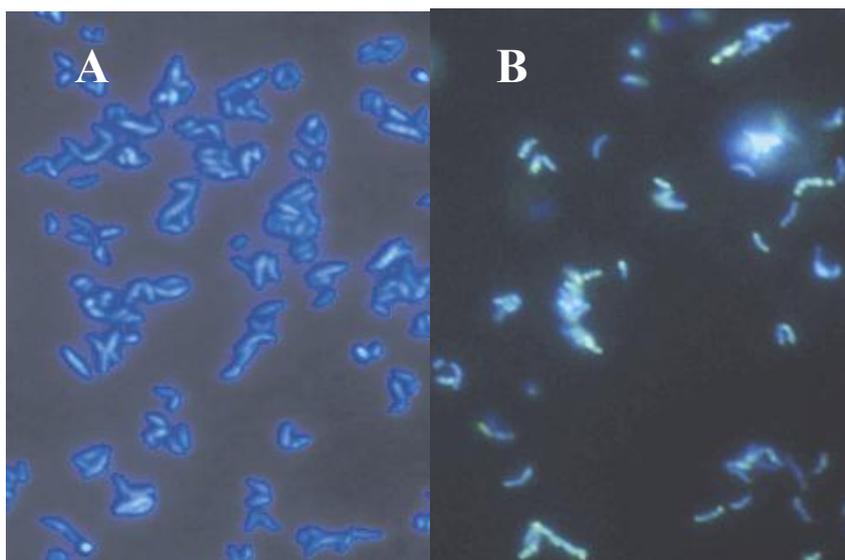


FIGURE 3-5. Microscopic images of *M. arborescens* SE14 (A) and *M. arborescens* DSM20754 (B). A drop of a 24-h-old culture was dried on slide, fixed with paraformaldehyde, and stained with the DNA dye 4,6-diamidino-2-phenylindole (DAPI) (X 1000). The complex of DAPI and DNA emits blue light.

It became clear that the catalytic activity of *M. arborescens* SE14 is strictly dependent on growth stage, raising in a short time of early stationary growth phase (Figure 3-3). Therefore, the activity of *M. arborescens* DSM20754 was monitored throughout the whole cultivation period. Although the amount of *N*-linolenoyl glutamine produced by *M. arborescens* DSM20754 is much lower, the maximum activity was detected when the bacterial growth reached a cell density of OD_{600} 2.0 (Figure 3-6).

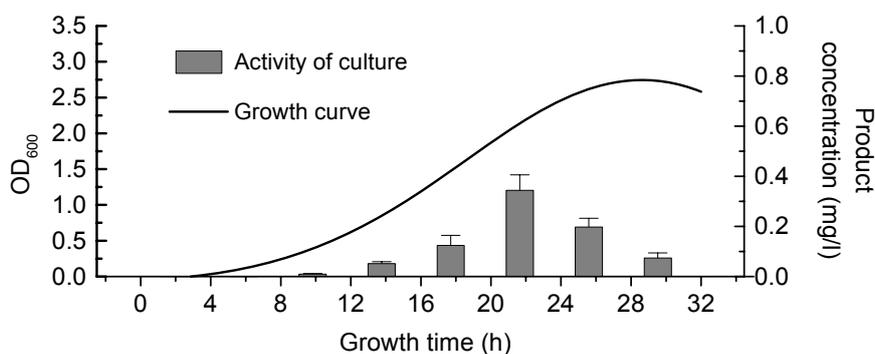


FIGURE 3-6. The enzymatic activity of *M. arborescens* DSM20754. The growth curve and the average concentration and mean error of linolenoyl glutamine synthesized in every 4 h were measured through out the cultivation ($n \geq 4$).

3.3 Purification of the amide-forming protein

The activity of *M. arborescens* SE14 to achieve amide formation has been shown to be growth-stage dependent (see 3.1.2), raising at early stationary growth. The relevant

Results

protein was found mainly in cell extracts (see 3.1.1). The enzyme activities inside and outside the cell were monitored throughout the bacterial growth.

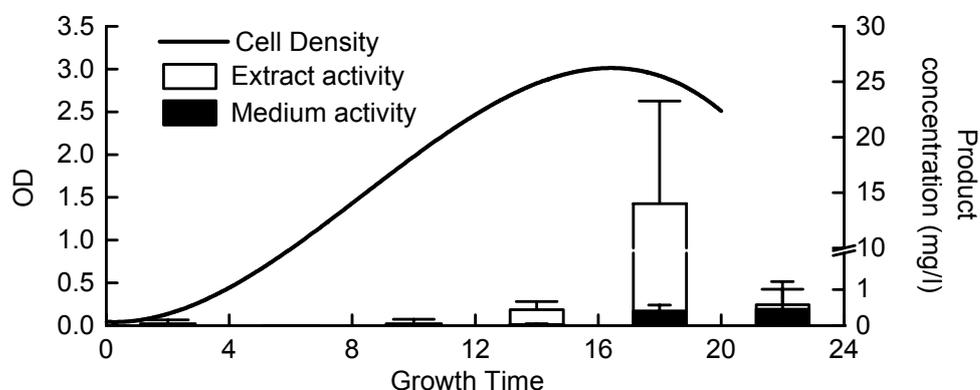


FIGURE 3-7. Enzyme activity varied at different growth stages. Bacterial cell density was monitored at OD₆₀₀. Total enzyme activity of 1 ml culture was tested. Bacteria were removed from medium fraction by centrifugation at 13,000 g for 30 min, and then sonicated in buffer. After removal of debris by ultracentrifugation activity in crude extract as well as medium was measured. Data points represent mean \pm SEM ($n \geq 3$)

The activity was not observed before and during exponential growth, but increased dramatically at the beginning of the stationary phase when OD₆₀₀ reached 2.0 (Figure 3-7). Because *M. arborescens* SE14 forms pellicle-structures during later stationary growth, at this time not only the free cell number in the medium decreased, but also the enzyme activity. Whether this is due to changes of the gene expression level or reduced protein release by the amorphous polysaccharides needs further proof. Normally no enzymatic activity was found in the medium after removal of the bacteria. Some activity in the medium during the later stationary phase might be due to autolysis of dead cells.

Table 3-1. Major steps of protein purification from *M. arborescens* SE14.

Step	Total protein (mg)	Specific activity (unit/mg)	Purification (fold)	Yield (%)
Crude extraction	411.2	0.0119	1	100
(NH ₄) ₂ SO ₄ precipitation	274.1	0.0179	1.50	65
Vivaspin 100 kD cutoff	196.9	0.0236	1.97	58
Anion exchange	0.92	1.7594	147.28	33
Size exclusion	0.58	2.2281	186.51	26

Average results are from 1 liter culture at 18h (Figure 3-7). Data represent mean \pm SEM ($n \geq 5$).

After 18 hours, the enzyme activity reached its maximum (Figure 3-7). The bacteria were harvested at this time for protein purification. After several steps the protein specific activity was increased 186 fold in an overall yield of about 26% of total activity (Table 3-1).

Results

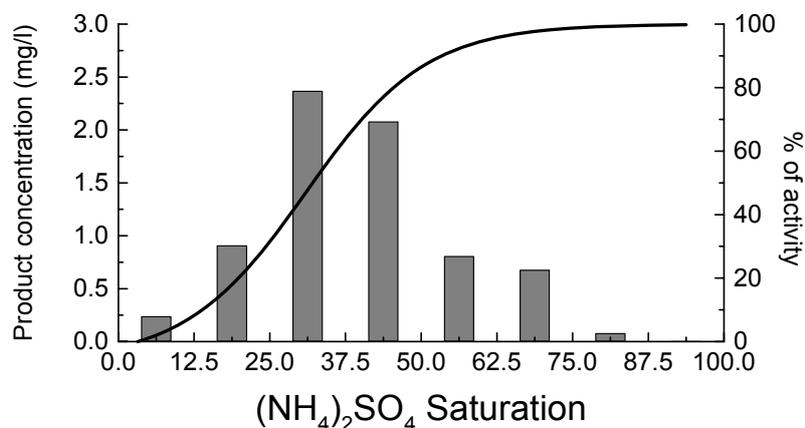


FIGURE 3-8. Activity of protein fractions precipitated at different salt concentrations. Enzymatic activities were measured by production of linolenoyl glutamine in 30 min in 1 ml Tris-HCl buffer pH 8.0. Protein was extracted from 1 ml 18-h-old culture.

The enzymatic activity was found to be very sensitive to ion strength. When the protein was suspended in 50 mM Tris-HCl buffer, the activity was 10 times lower compared to standard buffer (10 mM Tris-HCl). Thus, for the activity assays, proper desalination is required. Ultrafiltration with Vivaspin columns (Vivascience) was commonly used to desalt the protein during the purification procedure. Disposable PD-10 desalting columns (Amersham Biosciences) were used to remove salts completely from concentrated protein solutions.

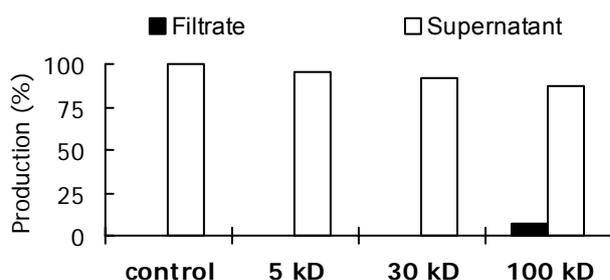


FIGURE 3-9. Enzymatic activity in the filtrate and supernatant after ultrafiltration with Vivaspin tubes of different size exclusion limits. The mesh sizes are indicated on the X-axis. Enzymatic activities assays were performed in 1 ml buffer. Protein was extracted from 1 ml of 18-h-old culture. Average concentration of linolenoyl glutamine and mean errors were calculated from 3 repetitions.

After removal of the cell debris, the clear supernatant was concentrated by ultrafiltration. Then the extracts were subjected to ammonium sulphate precipitation. Only a single active fraction was found in this fractionation, which precipitated at around 40%

Results

saturation. Fractions collected at 25-65% saturation contained more than 90% of the active protein (Figure 3-8). This fraction was used in the subsequent purification step.

Vivaspin columns of different mesh-size were tested (Figure 3-9). The results suggested that the native protein was larger than 100 kDa. Therefore an additional ultrafiltration with a Vivaspin column of 100 kDa exclusion limit was employed after ammonium sulphate precipitation to remove the salt and the smaller molecules which co-precipitated with the active protein. Vivaspin columns of 50 kDa were used to concentrate and partially desalt the protein after anion exchange and size exclusion chromatography.

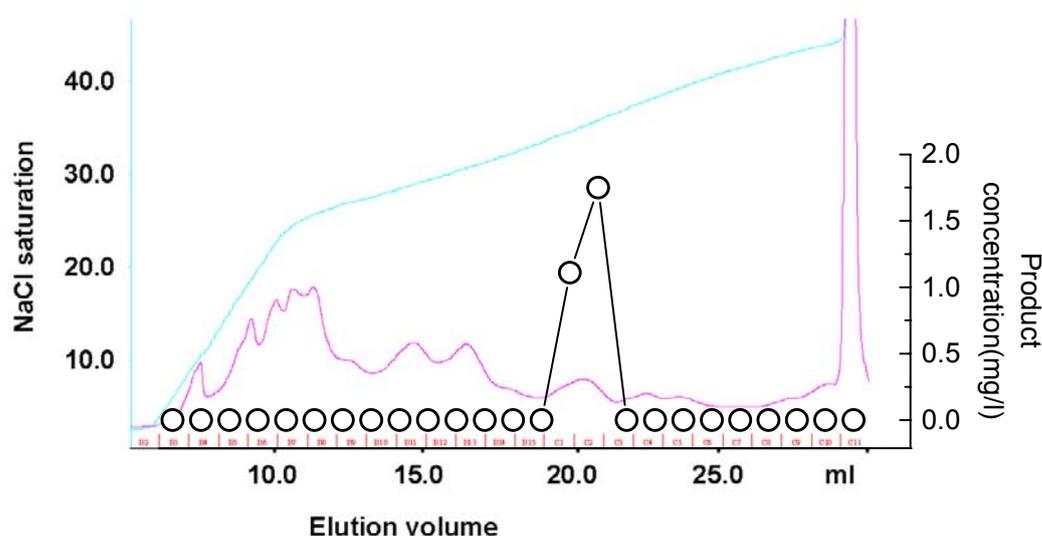


FIGURE 3-10. Elution of the active protein from anion-exchange column resource Q. Activity (circle) is measured with 25 μ l-aliquots of the eluant resuspended in 1 ml buffer after desalination. Pink line is the eluant detected by UV absorption at 254 nm.

The Amersham ÄKTA fast protein liquid chromatograph (FPLC) system is widely used for protein purification. ÄKTA FPLC system equipped with different columns was used for further purifications. The results from ammonium sulphate precipitation indicated that the surface of the active protein is highly charged, because that it was salted out at about 40 % ammonium sulphate concentration. Therefore an anion exchange column was used for further separation of the proteins. This is the critical step in the purification. Most of the contaminating proteins were separated by the ResourceQ column (6ml, Pharmacia Amersham) from the single active peak that eluted at 0.35 M NaCl (Figure 3-10). While a shoulder on the peak indicated that this column could not remove the contaminating fraction from active protein.

The active peak was collected and concentrated. Next, the protein was loaded on a Superdex 200 HR 10/30 size-exclusion column (Amersham Biosciences), and a single active protein, eluted as the major peak at 10.85 ml (Figure 3-11). This suggests that the native molecular weight (M_r) is rather high, coincident with the Vivaspin results (Figure 3-9). Because the active peak eluting from gel filtration column was very symmetric, the protein was assumed to be rather pure.

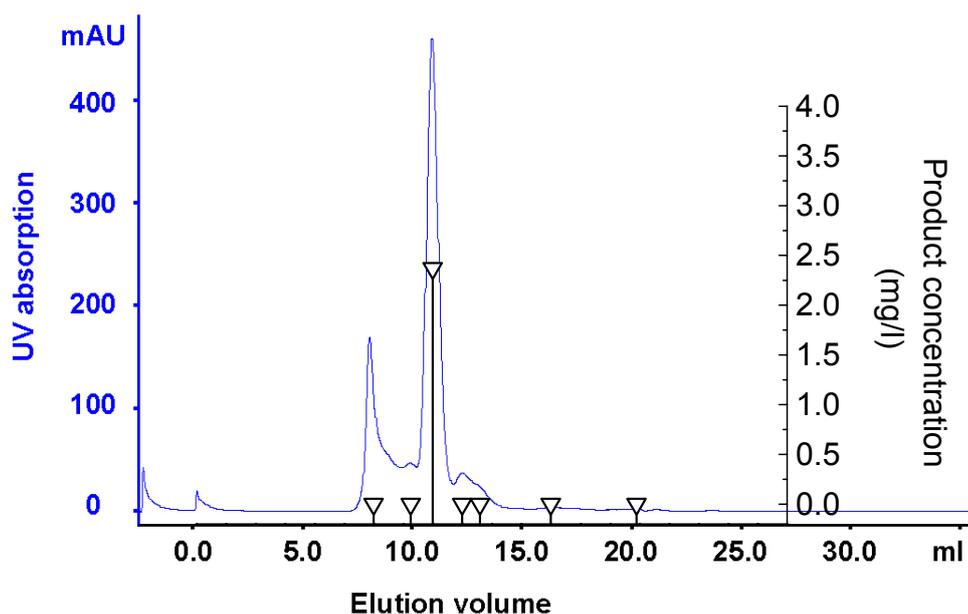


FIGURE 3-11. Size exclusion chromatography of the active fraction collected from anion exchange column. Activity (triangle) was measured with 5 μ l-aliquots of eluant. The protein solution was adjusted after desalination to 1 ml with buffer and activities were measured.

3.4 Characterization of the Afp protein

3.4.1 Determination of the molecular mass

To estimate the native M_r , the High Molecular Weight Calibration Kit for Electrophoresis (Amersham Biosciences) was employed as standard in gel filtration. The purified protein was mixed with the protein standards and loaded onto the size exclusion column. The protein eluted between Ferritin (440 kDa) and Catalase (232 kDa), suggesting an apparent molecular weight around 310 kDa (Figure 3-12).

The apparent M_r estimated by this method is a little higher than anticipated because that such a large protein should be completely retarded by the Vivaspin 100 kDa membrane, but *de facto* about 10% percent of the protein had passed through the membrane. On the other hand, the apparent M_r estimated by gel filtration could be higher if the protein is highly asymmetric or a strong interaction between the protein surface and column matrix takes place.

In order to determine the native M_r , the freshly prepared protein was subjected to flow field-flow fractionation (FFF) coupled to a multi angle light scattering (MALS) photometer (Figure 3-13). This relatively novel combination of FFF and MALS reduces the influence of protein shape and surface properties on M_r determination. Therefore it is suitable for studying the conformational changes and aggregation states of proteins

Results

(Wittgren *et al.* 1998). When 6 μg , 9 μg and 30 μg of the freshly prepared protein were loaded on the FFF-MALS system, the protein eluted as a single symmetric peak in all loads. There was no indication for different aggregates or contamination. The native protein mass was estimated to be 203 kDa based on the MALS data. The retention time delayed a little when the injection volume increased, but this is a well-known phenomenon due to sample overload (suggested by the producer). The calculated molecular mass did not change, however.

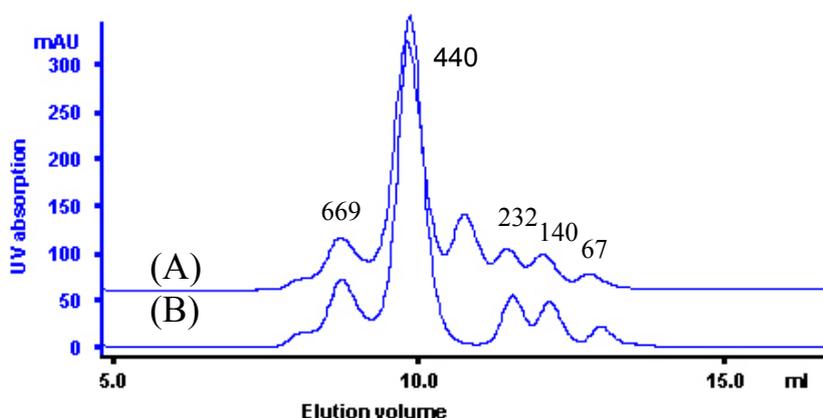


FIGURE 3-12. Estimation of the apparent M_r by gel filtration. (A) Chromatography of a mixture of purified protein and the protein standard. The M_r of corresponding peaks was shown. (B) shows the M_r of protein ladder only.

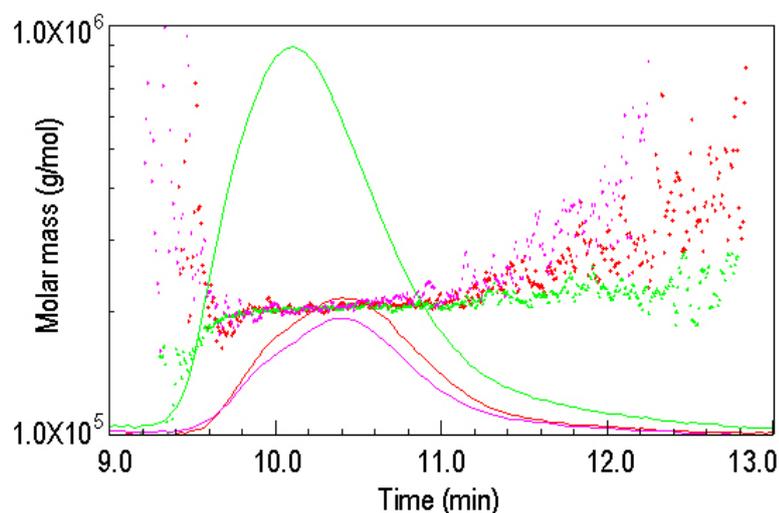


FIGURE 3-13. The molar mass versus time curves of three loadings of purified protein on a flow field-flow fraction coupled to a multi angle light scattering photometer. From top to bottom are protein amounts of 30 μg , 9 μg , and 6 μg , respectively.

When the purified protein was lyophilized and measured by matrix-assisted laser desorption ionization (MALDI) combined with time-of-flight mass spectrometry (TOF)

Results

only a single polypeptide of 17.181 kDa was detected (Figure 3-14). A series of ions with double-, triple- and tetra-fold molecular mass were detected besides the molecular ion and an ion with two charges. In agreement with the FFF-MALS analysis, no other significant amount of contaminating molecules was detected by this method.

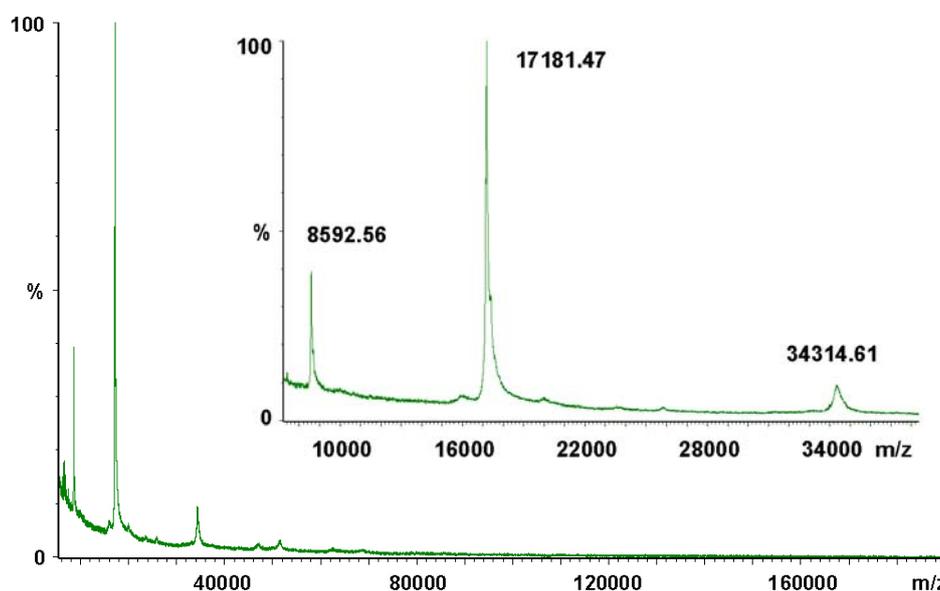


FIGURE 3-14. MALDI-TOF determination of the protein mass. Insert: section of the profile with the three major ions was enlarged. In addition to a prominent ion at $m/z = 17,181$, ions of a dimer with a single charge ($m/z = 34,310$) and protomer with two charges ($m/z = 5,895$) were also detected.

The combination of the above results suggests that the native protein forms a homooligomer consisting of 12 subunits. This is fully coincident with the results of the cloned gene (see 3.7.4). The oligomeric structure has been additionally confirmed by quadrupole-hexapole-quadrupole mass spectrometry (Quattro, Micromass) coupled with electrospray ionization (ESI). A series of ions with regularly increased mass were detected by the Quattro-ESI (Figure 3-15). All of the m/z of these ions in the low-mass region of the MS profile perfectly matched the following equation:

$$M = \frac{17177 \times 12(n + m) + 12n}{12n}$$

(M is the m/z detected by Quattro-ESI; n is the number of charged dodecamers; m is the number of uncharged dodecamers.)

Probably due to the deprotonation in the solution in Quattro-ESI (pH 8.0), the M_r of the protein monomer detected by this method is 17.177 kDa, which is 4 Daltons smaller than that detected by MALDI-TOF (17.181 kDa, pH 3.5). Each major m/z peak separated

Results

from others with exactly the same interval of 1431 Daltons. This difference might also be due to the lower accuracy of MALDI-TOF measurements.

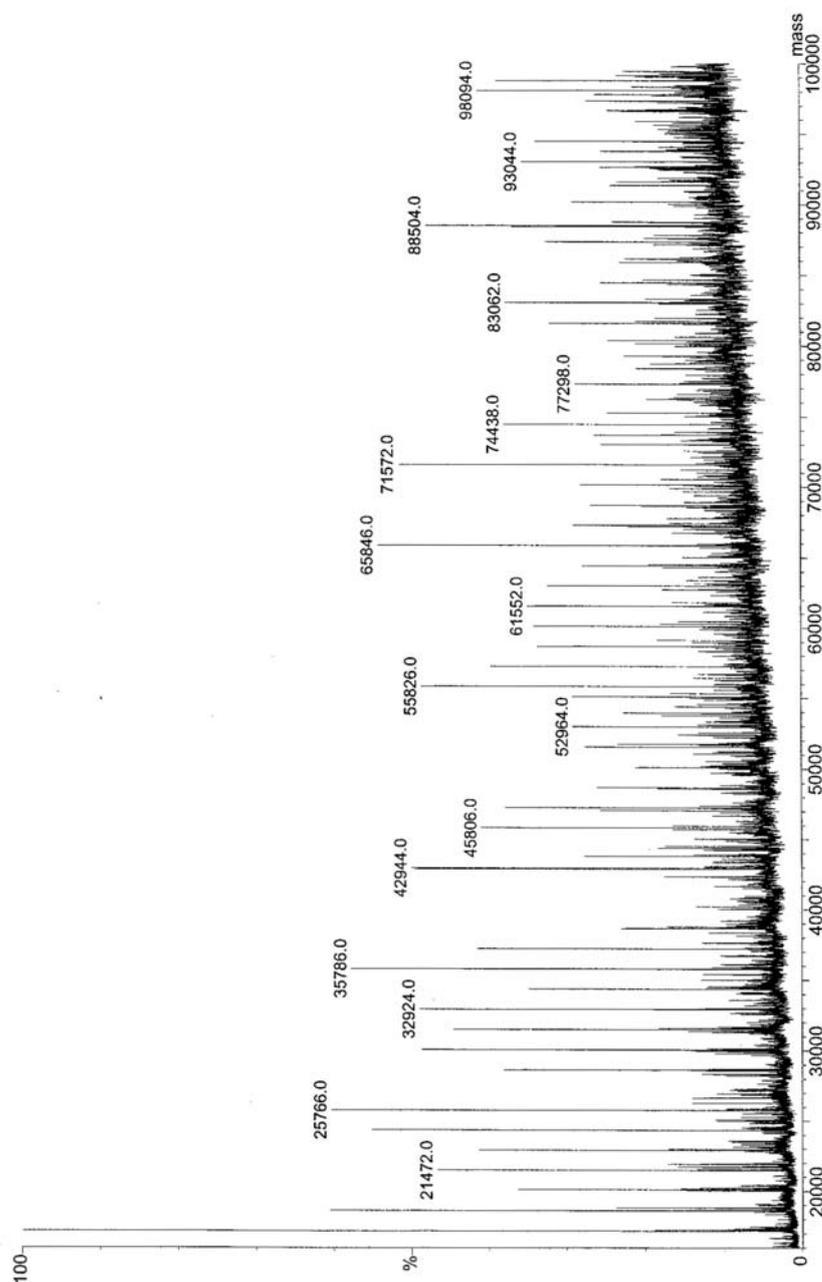


FIGURE 3-15. The mass spectrum of the native protein isolated from *M. arborescens* SE14 measured by Quattra-ESI. Some mass/charge ratios were shown on corresponding peaks.

The “n” value in above equation has been tested up to $n = 60$. All of values of the major peaks in the spectrum have been proven by this calculation. Interestingly, the formation of certain kinds of ions seems to be less likely, the ions of $n=12, m=7$; $n=24, m=34$; and $n=36, m=78$ were almost undetectable. The mechanism underlying this phenomenon is not clear yet. In the high m/z region, the ions did not fit by this calculation. They might be signals generated by some very large aggregates therefore need higher “n” and “m” values or the signals from some uncertain ions. This spectrum gives strong evidence that the protein monomers aggregate forming a stable dodecamer, which in turn could serve as a building block for even larger aggregates.

Results

However, when the protein was subjected to Polyacrylamide Gel Electrophoresis (PAGE), its behaviour didn't look like a normal protein. On native PAGE, a single band of about 100 kDa was detected (Figure 3-16 A). Instead of a dodecamer, it formed an hexamer upon native gel electrophoresis. The hexameric protein is very stable. The protomers did not completely separate from each other even when the protein was loaded on a gel supplemented with the detergent sodium dodecyl sulphate (SDS) (Figure 3-18 A).

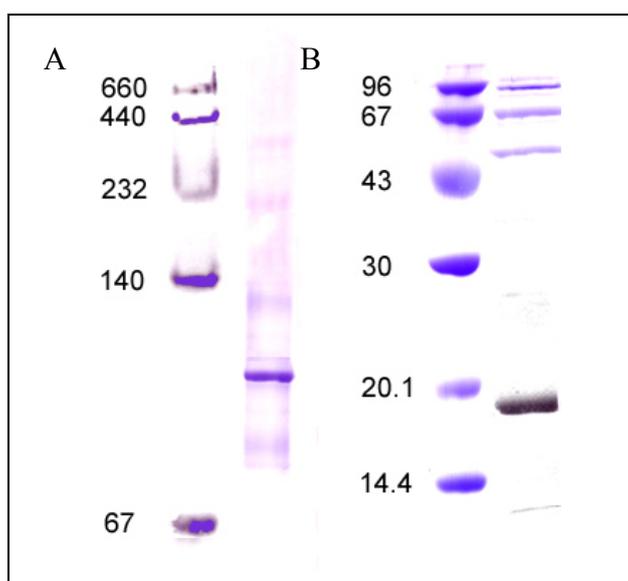


FIGURE 3-16. Protein separation by Native PAGE (A) and SDS-PAGE (B). The Mrs of ladders are shown on the left. The presentations represents 3 separate measurements. Some of them are stained either with Ponceau S or with Coomassie brilliant blue. They have been modified to unify the colour and to increase the resolution.

On the denaturing SDS-PAGE (Laemmli 1970), a native hexameric structure was still visible even after boiling of the samples in the presence of the detergent SDS. The denatured protein fraction on the gel consisted predominantly of a 17 kDa monomer along with several higher Mr bands coincident with oligomers, such as trimers, tetramers, and pentamers (Figure 3-16 B). The relative abundance of these oligomers varied between the preparations, suggesting that the number of the protomers in these oligomers were random.

3.4.2 Determination of the isoelectric point

Both salt precipitation and anion exchange column results indicate that the protein is highly charged under neutral and slightly alkaline conditions. Thus, the isoelectric point was measured by using the Immobiline DryStrip gel (pH 4-7, 24 cm, Amersham Biosciences, determined by Dr. Buechler). On the Immobiline DryStrip, the protein focused as a single band corresponding to pH 4.1 (Figure 3-17). The strip had been stained with acidic violet as suggested by the producer. Although by this staining method

Results

it is not possible to remove the background dye totally from the strip, it is clear that there was no other visible band on the strip.



FIGURE 3-17. Isoelectric focusing of Afp (arrow) on an Immobiline strip stained by acidic violet. The pH distribution is shown under the strip according to the technical data of the manual. Picture is from Dr. Rita Buechler.

3.4.3 Iron content

Accumulating data from parallel genetic work (see 3.7.4) suggested that the Afp is a homolog of bacterial DNA-binding proteins in starved cells (Dps). All known Dps-like proteins contain iron (Ceci *et al.* 2003). The protein fraction was subjected to potassium ferrocyanide staining, an iron detecting method also known as Berlin Blue staining (Papinutto *et al.* 2002). The protein hexamer on native gel was heavily stained by $K_3[Fe(CN)_6]$. Even if, the native protein was loaded on SDS gel, it still maintained the hexameric structure (Figure 3-18 A), which could be stained by the iron dye (Figure 3-18 B). But when the protein was denatured, neither the monomer nor the oligomers was stained on SDS-PAGE page (Figure 3-18 C).

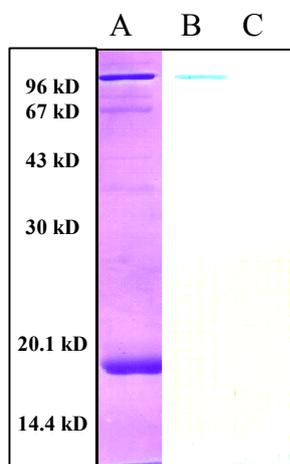


FIGURE 3-18. Potassium ferrocyanide staining of the purified protein separated by SDS-PAGE. (A) is the control, in which the protein was denatured by boiling with SDS and then mixed with same amount of native protein and stained by Coomassie brilliant blue; (B) is the native protein stained with potassium ferrocyanide; (C) is the denatured protein stained with potassium ferrocyanide. Picture is from Dr. Rita Buechler.

In fact, when the purified Afp protein was concentrated during purification, a light brown color had been noticed by naked eyes, suggesting that the iron content of Afp protein is quite high. Inductive coupled plasma optical emission spectroscopy (ICP/OES) (Optima 3300 DV, Perkin-Elmer) with an axially arranged torch and cross-flow nebulizer was used to determine the exact iron content of the native protein. The average result is 115 ± 21 iron atoms in the dodecameric protein. The iron content per subunit was therefore calculated to be 9.6 ± 1.75 .

3.4.4 DNA-binding assay

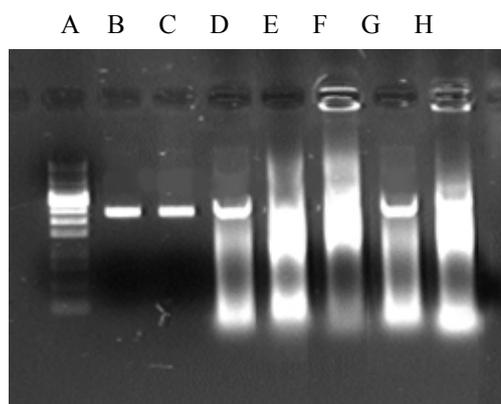


FIGURE 3-19. DNA-binding assay of the purified protein. (A), DNA marker; (B), plasmid without protein; (C), DNA mixed with 0.1 μg Afp; (D), DNA with 0.4 μg Afp; (E), DNA with 1.6 μg Afp; (F), DNA with 0.4 μg *E. coli* Dps; (G), DNA with 0.4 μg denatured Afp; (H), DNA with 0.4 μg denatured *E. coli* Dps.

The DNA binding activity of the Afp protein was tested by comparison with the *E. coli* Dps protein (a kind gift of Prof. R. Kolter, Harvard University). The proteins were mixed with plasmid pET11a DNA (Stratagene) before running on a 1% agarose gel as described (Almirón *et al.* 1992). Consistent with the previous report (Ceci *et al.* 2003) describing that DNA-binding depends only on the primary sequence, the *E. coli* Dps protein, even after denaturation at 70°C for 30 min, retarded the DNA in agarose gel slot as did the non-denatured protein. Using the same assay, neither native nor denatured Afp could form a complex with DNA (Figure 3-19). Even if the Afp amount used was 4 times higher than that of Dps, on complex could be detected.

3.5 The catalytic properties of Afp

3.5.1 A rapid reversible reaction

In our experiments, the Afp protein concentration was 1 international unit (I.U., defined as the amount of enzyme that produces 1 μmol of linolenoyl glutamine per

minute) per litre. *N*-linolenoyl glutamine was formed with an equilibrium constant (K_{eq}) of 0.1863 in 2 min. When the reaction reaches its equilibrium, only a small amount of the substrates was converted to the corresponding conjugate. However, when protein was denatured at 70°C for 20 min, no product was detected.

In order to measure the reaction velocity, samples were taken at 0.5, 1, 2, 4, 10, and 20 min, respectively, with different substrate concentrations. The Afp's turnover number (k_{cat} , the number of moles of glutamine converted to *N*-linolenoyl glutamine each second per mole of enzyme) was calculated with the Mr of 17.18 kDa determined by MALDI-TOF. The reaction k_{cat} (s^{-1}) at pH 8.0 is 0.6378. The K_m of the forward reaction is not detectable, while the K_m of hydrolysis when supplied with synthesized *N*-linolenoyl glutamine was measured as 36 $\mu\text{mol/l}$. This indicates this enzyme prefers **hydrolysis** instead of conjugation.



$$K_{eq} = \frac{[\text{LeGln}]}{[\text{Lea}][\text{Gln}]} = 0.1863$$

$$K_{cat} = 0.6378$$

$$K_{m(\text{hydrolysis})} = 36 \mu\text{mol/l}$$

(Lea, linolenic acid; Gln, glutamine; LeGln, linolenoylglutamine)

3.5.2 The temperature optimum

The thermostability of the Afp protein is very high. It was active in a large temperature range around 40°C *in vitro* (Figure 3-20). The standard assays therefore were always carried out at 37°C. Although the catalysis could be handicapped by partial denaturation at temperatures higher than 40°C, the lost activity at temperatures up to 48°C could be recovered by transferring protein on ice. If the protein had been heated to temperatures higher than 60°C for 20 min, no activity could be restored (Figure 3-20 dashed line). In the research, the protein was usually heated at 70°C for 20 min, if denaturation was required.

The Afp protein is active at relatively high temperatures. The activity curve showed a normal distribution when plotted versus temperature. The maximum activity was observed at 40 °C. As temperature decreased, the activity also decreased rapidly. At 20°C, only a trace amount of products could be detected (Figure 3-20 solid line).

3.5.3 The pH optima of Afp

Although the pI of this protein is as low as 4.2 as determined by isoelectric focusing, the formation of linolenoyl glutamine was detected only at alkaline pHs (Figure 3-21 B).

Two pH optima were observed. These pH ranges are coincident with the pH values of gut segments (Figure 3-21 A) as reported before (Lait *et al.* 2003). The protein is highly active at pH 8.0. A shift of only 0.4 pH units to both directions decreased the protein activity dramatically. No activity was detected at neutral and acidic pH. Interestingly, the activity at pH around 8.5 was extremely low, only trace amount of products was detected. Another broad pH range is from 9.2 to 11.2. The average activity of the enzyme in this region was two times higher than that at pH 8.0. The catalytic activity was totally lost when the protein was suspended in buffers with pH value more than 12.5. This could be also due to the high ion strength in these buffers.

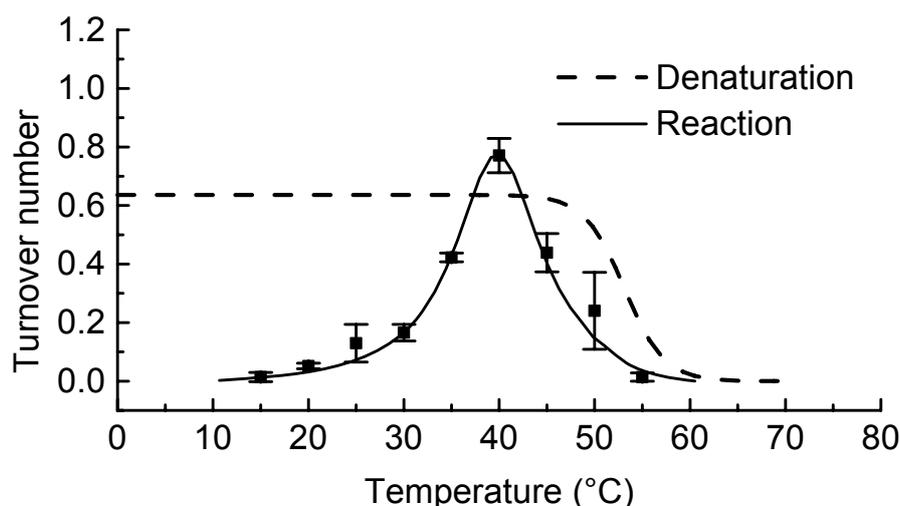


FIGURE 3-20. Effect of temperature on catalytic activity. Solid line represents reaction velocity at different temperatures. Experiments were performed with standard mixtures incubated at different temperatures. Dashed curve represents the activity after heat treatment of the protein for 20 min. The heated protein was returned on ice for 5 min before standard assays. Each data point represents an average result with SEM ($n = 7$).

3.5.4 Characterization of the active centre

Since the reaction was enhanced at alkaline pH range with the highest activity starting at pH 9.2, it is reasonable to assume that the deprotonation of a critical residue in the active center could be necessary for catalysis. The pK for deprotonation of the phenolic hydroxyl group of tyrosine is 10.1 (Stanier *et al.* 1987). Therefore tyrosinate is the most likely candidate for this function. The ultraviolet (UV) absorption was used to detect a shift from tyrosine to deprotonated tyrosine in the enzyme active centre, since the formation of a phenolate ion is usually associated with a large shift to longer wavelength.

The UV absorbance of Afp was therefore monitored at pH 8.0 and pH 10.0 (Figure 3-22 A) by spectrofluorometry (V-550, Jasco). An increase of the absorbance from 245 nm to 295 nm was observed. The maximum increase was found at 255 nm (Figure 3-22 B) when the two curves were subtracted by the Spectra Analysis software. Although we still do not fully understand the physiological significance of the anticipated deprotonation of

Results

tyrosine here, the UV-absorption did change simultaneously with the activity enhancement.

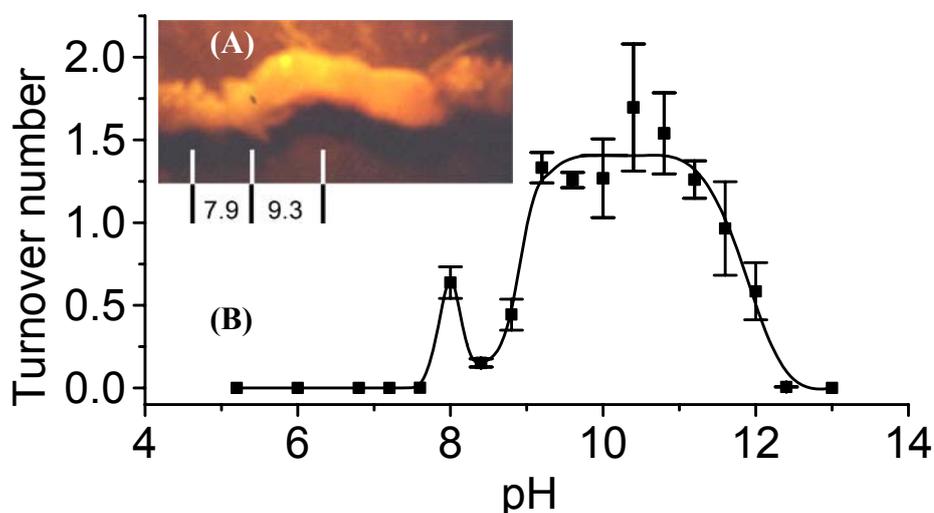


FIGURE 3-21. The pH optima of Afp protein. (A): Gut of a beet armyworm larva with the average pH in different areas. (B): Protein exhibits activity only at alkaline pH values. Enzyme activity was tested with standard reaction mixture in different buffers, which were made by sodium acetate/acetic acid (pH 5.2-6.8), Tris/HCl (pH 7.2-8.8), and Tris/NaOH (pH 9.2-13.0). All of the buffers are 10 mM. Points on the graph are mean \pm SEM (n = 10)

The detected formation of conjugates could also be a side reaction of a synthase with an undiscovered activation mechanism. ATP is the driving force of many different kinds of acylation processes (see 1.3.1). Therefore, 0.03 mM ATP was supplied, but it didn't cause any change on the enzymatic reaction (Table 3-2).

Table 3-2. The influence of some chemicals on the Afp catalytic activity.

Chemical	Target	Concentration	Relative activity %
Standard reaction			100
ATP		0.03mM	100 \pm 35
EDTA	Metal ion	1mM	118 \pm 55
Hydroxylamine	Carbonyl group	1mM	100 \pm 13
Phenylhydrazine	Carbonyl group	1mM	95 \pm 20
Iodoacetate	Cysteine SH group	1mM	93 \pm 34
N-Ethylmaleimide	Cysteine SH group	1mM	111 \pm 30
DTNB	Cysteine SH group	1mM	96 \pm 23
PMSF	Serine OH group	1mM	90 \pm 29

The relative activity was calculated by comparison with the standard reaction at pH 8.0 and pH 9.2, respectively. Then the relative values were averaged (n \geq 3).

The UV absorption suggests a potential involvement of a tyrosine residue in the enzyme active centre, but it is not an absolute proof. The protein only converted a small amount of substrates to *N*-acyl amino acid, which can be considered as a reversal of the

hydrolysis reaction. Some hydrolyases need ions for catalysis (see 1.3.3). The aminopeptidase A is such a kind of metallo-hydrolases. A tyrosine residue had been demonstrated to be involved in the catalysis (Vazeux *et al.* 1997). We had also found that increase the ionic strength of the solution decreased the enzymatic activity (see 3.3). Therefore the ion chelator EDTA (Table 3-2) was then added into the standard reaction and activities were measured both at pH 8.0 and pH 9.2. However, no influence on the enzymatic activity was found at both pHs.

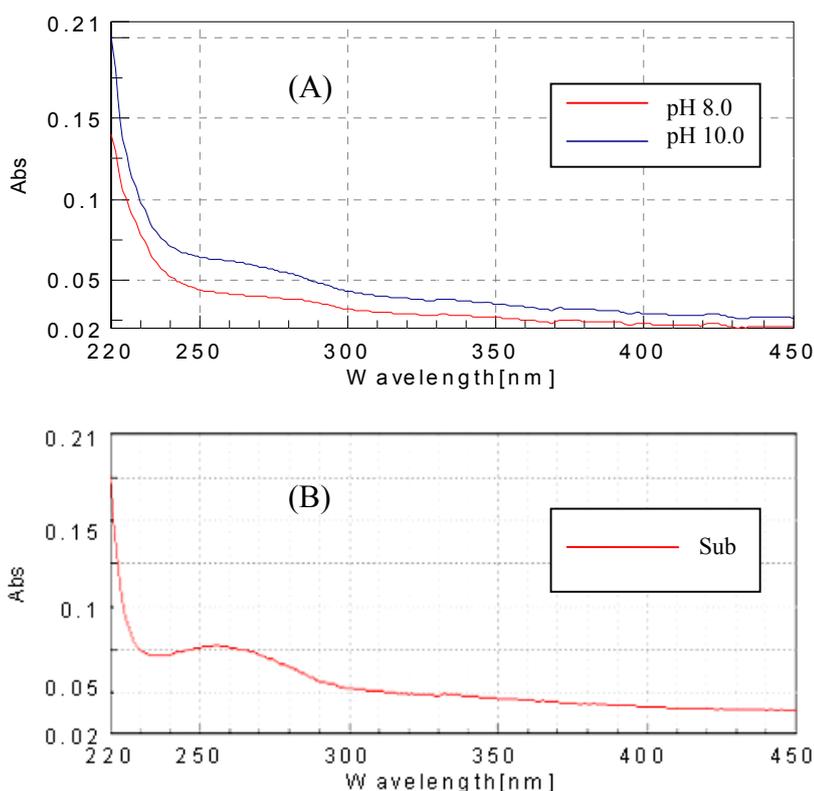


FIGURE 3-22. UV absorption of purified protein at pH 8.0 and 10.0. The experimental curves (A) and subtraction of the two curves indicates a maximum absorbance at 255 nm (B). Abs, absorbance; Sub, subtraction.

Furthermore, some well documented proteinase inhibitors, hydroxylamine, phenylhydrazine, iodoacetate, *N*-ethylmaleimide, phenyl methyl sulfonyl fluoride (PMSF), and 5,5'-Dithio-bis(2-nitrobenzoic acid) (DTNB) were used to block the catalytic centre. They can modify different residues in reactive center of amidases and proteases (see 1.3.3). None of them inhibited the reaction significantly in the assays (Table 3-2). The possibility of Afp being a known type of hydrolyase was thus excluded.

3.6 The substrate specificity

The *N*-acyl amino acids in insect oral secretion (see 1.1.1) only contain conjugates of glutamine and in some cases glutamate. On the other hand, there seems to be no selection on the fatty acid substrate. Therefore the substrate specificity of Afp was determined.

3.6.1 Specificity on amino acid

The Afp catalyses an enantioselective reaction. Only L-amino acids were incorporated in the corresponding conjugates. The Afp protein did not accept the D-amino acids as substrates in these experiments.

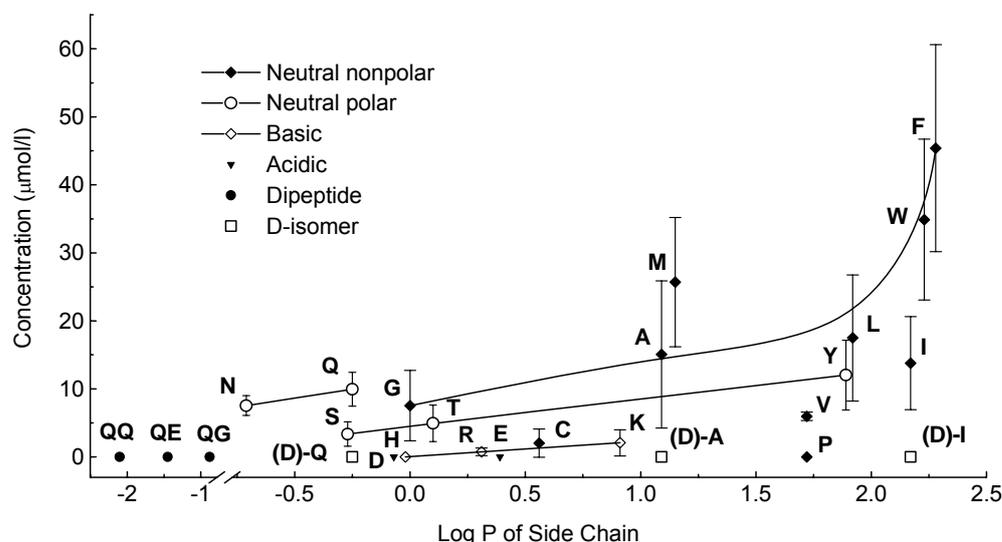


FIGURE 3-23. Substrate selectivities of the amide formation. Concentration of amino acids and derivatives are 20 mM. The fatty acid substrate is 3.92 mM linolenic acid; protein concentration was 1 I.U./l. Product concentration is measured in equilibrium. Log P value is calculated based on side chains only (not including the α carbon atom, carboxyl and amino group). Experiments were repeated 3 times. Mean value and SEM ($n = 3$) were present. One-letter abbreviations of the amino acids were given next to the data points. The three D-isomers tested were indicated by a 'D' in brackets. Dipeptides were labelled by tandem abbreviations of the two residues as following, QQ (Gln-Gln), QE (Gln-Glu), QG (Gln-Gly).

Except proline, all other proteinogenic amino acids were converted into corresponding *N*-acyl-conjugates (Figure 3-23). The acylation clearly could not happen with the imido group of proline. The efficiency of incorporation of other amino acids depended on the hydrophobicity of the side chain. Nonpolar amino acids were better substrates than neutral and polar amino acids. Basic and acidic amino acids conjugates could only be detected at low concentrations. Due to their large hydrophobic side chains, phenylalanine, tryptophan, and methionine were the best amino acid substrates among those tested in this experiment. Neutral polar amino acids were not as good substrates as nonpolar amino acids with side chains of same size.

Glutamine is the best substrate among those with an octanol/water partition coefficient (LogP) less than 1.0. Surprisingly, when the carboxyl group of glutamine was linked to another amino acid, whatever the hydrophobicity of the substitute residues, there was no conversion of those dipeptides into an *N*-acyl-peptide observed. The

blocking of the free α -carboxyl group seems to have a stronger effect than variation on the sidechains.

3.6.2 Fatty acid selectivity

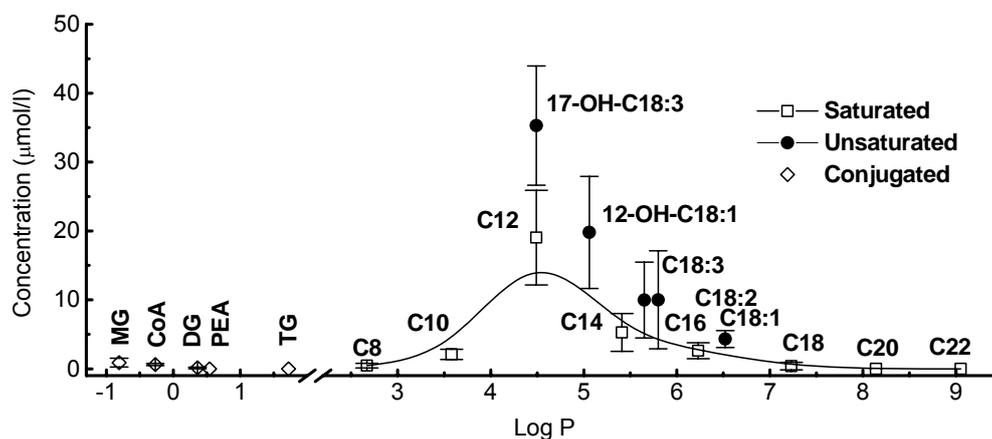


FIGURE 3-24. Conversion of fatty acids into amides by Afp. Fatty acid concentrations in the tests are 10 mM, with 13.69 mM glutamine as the second substrate. Enzyme concentration is 1 I.U./l. Experiments were repeated 9 times, and average values were calculated. Log P values of the fatty acids are calculated with the whole molecules. Fatty acids are represented by their carbon atom number followed by the number of double bonds. Hydroxy group positions of hydroxylated fatty acids are also specified. Some esterified fatty acids had been also tested. Log P of these fatty acid derivatives are those of the conjugated moieties, which do not include the ester bond. Their abbreviations are as follows: TG (trilinoleoyl glycerol), PEA (dioleoyl glycerol phosphoethanolamine), DG (1, 3-dilinoleoyl glycerol), CoA (linoleoyl CoA), MG (monolinolenyl glycerol). Mean \pm SEM (n = 9).

In contrary to the water-soluble amino acids, fatty acids are highly hydrophobic. Nevertheless their incorporation still increased parallel to their hydrophobicity until the acids became insoluble (Figure 3-24). For improving the solubility of the fatty acid, fatty acids were dissolved in ethanol before mixing with other components. The final concentration of ethanol in the reaction mixture was 10%. This concentration had no statistically significant influence on Afp activity. This is another proof that the Afp protein is highly resistant to denaturation.

The incorporation curve of saturated fatty acids shows a normal distribution. Dodecanoic acid was the best fatty acid substrate among those tested. The presence of hydroxyl groups and double bonds increases fatty acid solubility. Such modified 18-carbon fatty acids of a Log P 4.0–6.0 were effectively conjugated, much more efficiently than the saturated octadecanoic acid. The ratio for the incorporation of 18-carbon fatty acids into *N*-acyl glutamines catalyzed by Afp, matched perfectly with the ratio of the

corresponding conjugates in Beet Armyworm caterpillar gut content (Figure 3-25), and caterpillar regurgitant reported (see 1.1.1).

To learn more about the catalytic reaction, some glycerides, phospholipids, and acyl-CoA had also been tested as substrates (Figure 3-24). But only a trace amount of conjugate was detected in mono-acyl glycerol and acyl-CoA reaction. Probably, the conjugation happened after the release of the free fatty acid from them by autohydrolysis.

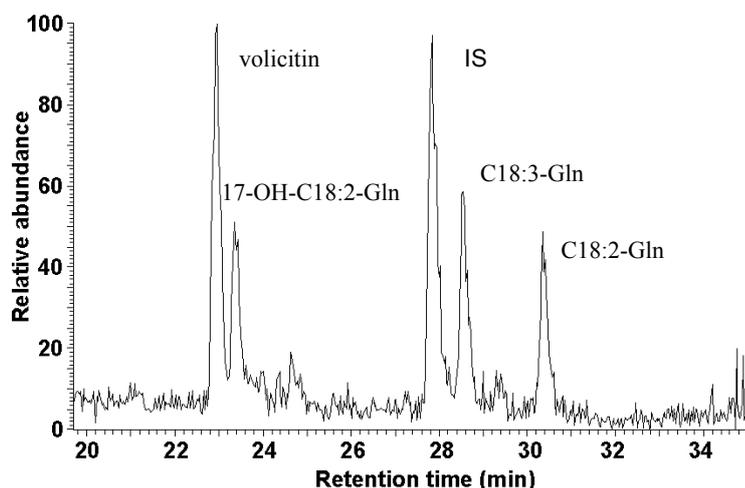


FIGURE 3-25. A typical *N*-acyl glutamine profile in Beet Armyworm gut content measured by HPLC. IS, internal standard. The fatty acid moiety of conjugates were labeled by their carbon atom number followed by the number of double bonds. Hydroxy group positions of hydroxylated fatty acids are specified. The ions used for monitoring are $m/z = 404.5-408.5$; 431.0 ; 445.0 ; 447.0 .

Two oxidized fatty acids, with different numbers of double bonds and positions of the hydroxyl groups, were accepted more efficiently than non-hydroxyl fatty acids. The 17-hydroxyl linolenic acid, the moiety of volicitin in insect secretion, was the best substrate among all fatty acids tested. In BAW caterpillar foregut lumen, the concentration of *N*-acyl amino acids varies significantly between individuals. By HPLC measurements ($n = 6$), the average concentration of volicitin was found to be *ca.* 200 μM , that of *N*-Linolenoyl-L-glutamine was *ca.* 100 μM .

3.7 Cloning the Afp-coding gene

3.7.1 Protein sequencing

The protein, separated by SDS-PAGE, was transferred on to a PVDF membrane (Bio-Rad). The dominant band of 17 kDa was then sequenced by Edman degradation (amino acid sequencer 494A, Applied Biosystem model, USA). A sequence of 17 residues was identified and named *eda-1*. The *N*-terminal sequence (below) did not contain the initial residue methionine.



Results

To get more information on the protein sequence, protein from the gel was digested by trypsin according to published procedures (Shevchenko *et al.* 1996). Electrospray ionisation (ESI)-tandem mass spectrometry (MS/MS) was chosen to sequence the peptides after tryptic digestion. ESI-MS/MS was performed on a Q-TOF 1.5 hybrid mass spectrometer (Micromass, UK) using “medium” nano ESI-capillaries (Protona, Denmark). Five additional peptide sequences were obtained by this approach:

msa-1 NH₂-QA (IL) VVNGK-COOH

msa-2 NH₂-HWNVR-COOH

msa-3 NH₂-TSTAVPAGFAQWQDE (IL) K-COOH

msa-4 NH₂- (IL) VA (IL) G (IL) P (IL) DSR-COOH

msa-5 NH₂-N (IL) TTPA (IL) TADPEVAAAAAQF (IL) T-COOH

The peptide fragment msa-5 did not contain the *N*-terminal three amino acids of *eda-1*, due to the removal of the tripeptide by trypsin, but was longer than *eda-1* at the C-terminus. MS/MS-ESI can't distinguish isoleucine and leucine due to their identical molecular weights. The *N*-terminal peptide was thus proposed by alignment of the two sequences:

NH₂-NITTPALTADPEVAAAAAQF (IL) T-COOH

This sequence is confidently located at the protein amino-terminus, while the other peptides are somewhere in the middle of the protein. PCR upper primers were designed from this sequence and lower primers from the other MS/MS-ESI sequences to amplify the corresponding gene from isolated genomic DNA (Figure 3-26)

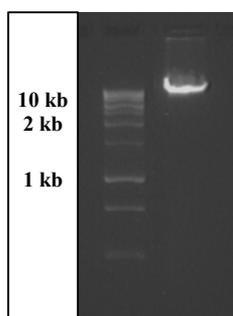


FIGURE 3-26. DNA purified from *M. arborescens* SE14 on a 1.2 % agarose gel page. Left is size maker. The largest band is 10 kb.

3.7.2 Nested degenerate primer PCR

The CTAB (cetyltrimethylammonium bromide, hexadecyltrimethylammonium bromide) assisted procedure for genomic DNA isolation from streptomycetes (Kieser *et al.* 2000) had been modified to isolate the total DNA of *M. arborescens* SE14 (see 5.7). High-quality DNA was isolated by this method (Figure 3-26), which was used for the following PCR reaction and for cloning of the gene.

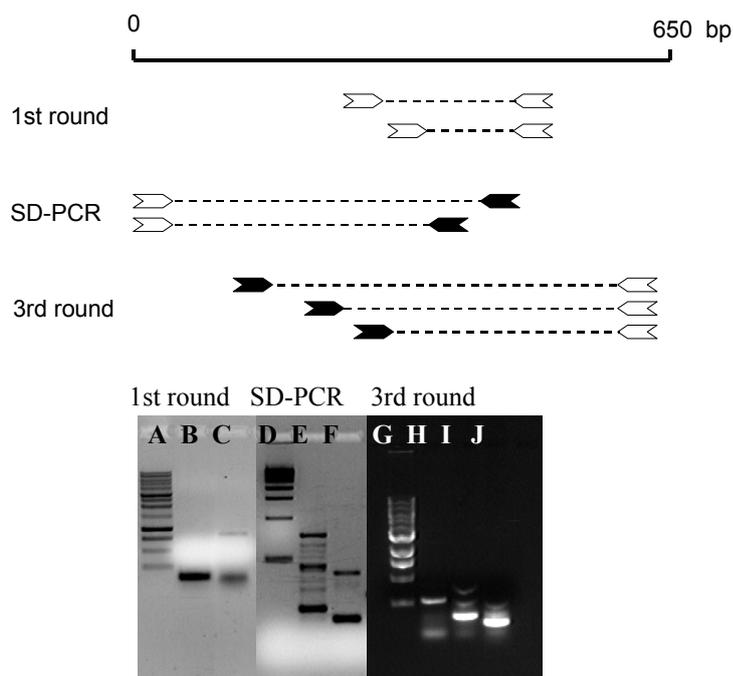


FIGURE 3-27. An overall strategy of the nested priming PCR cloning. In the diagram degenerate primers are shown as open arrows, normal primers are shown as solid arrows. The PCR results of three rounds of amplification are also shown. The 1st round PCR result: A, 1 kb marker, the smallest band is 250 bp; B, *asmf1* vs. *asmr2*; C, *asmf2* vs. *asmr2*. The SD-PCR result: D, the DNA ladder; E, PCR result of *sdu* vs. *asmr3*, notice there is a mismatching band on the top; F, result of *sdu* vs. *asmr4*, the mismatching rate is lower. The two major bands and the corresponding bands in E are specific amplificants. The 3rd PCR: 3 nested primer pairs, *asmf4* (H), *asmf5* (I), and *asmf6* (J) paired with *asmr5*. G, the DNA ladder, smallest band corresponds to 500 bp.

Three rounds of nested degenerate priming PCR (Figure 3-27) were performed to amplify the upstream region and 5'-terminal region of the corresponding gene. Because deoxyinosine (I) can base-pair with all four naturally occurring deoxyribonucleotides (Shen *et al.* 1993), it has been used in this experiment to reduce the number of primer sequences (Table 2-4). According to our observation, the primer degeneracy doesn't influence the annealing temperature, but decreases the matching primer number. Therefore the concentrations of primers in PCR reactions varied according to the degeneracy as shown in table 2-4. To figure out the right amplificant from the relatively high mismatching background produced by degenerate primers, nested primers were used in each round of PCR. Promising PCR products were cloned into a pCR2.1-TOPO vector (Invitrogen) and sequenced (by GATC Biotech AG).

3.7.2.1 Cloning the 5'-sequence of the *afp* gene

Two degenerate primers, *asmf1* and *asmf2* were designed from the *Afp* N-terminal sequence (Table 2-4). They are paired with another degenerate primer, *asmr2* designed from peptide sequence *msa-2*. The PCR product of *asmf1* vs. *asmr2* would be 27 bp longer than that of *asmf2* vs. *asmr2* according to the peptide sequences. When the PCR products were separated on agarose gel, there were two bands produced by *asmf2* vs. *asmr2* reaction. The smaller one was promising because it has the expected size (Figure 3-27). The larger band was shown by a separate experiment to be a mismatching product amplified by *asmr2* alone.

The product of *asmf1* vs. *asmr2* reaction was then cloned and sequenced. It was proven to be a 137 bp fragment. It encodes a peptide sequence containing 45 amino acids. No other MS/MS-ESI peptide sequences were found in this peptide besides *msa-5* and *msa-2* on both ends. A blast search on different protein databases with this peptide sequence in NCBI, EMBL-EBI, SMART, MyHits, ExPASy, PDBj did not yield any matching result. Normally a peptide of 15 amino acids long would give out confident result on protein identity. Therefore the peptide sequenced indicates that this could be a new protein, or this protein has a unique amino-terminus, which might be important for its physiological function.

3.7.2.2 SD-PCR

The upstream sequence of the cloned fragment from the first round was more interesting because there was no starting-codon residue found in the protein sequences. The protein might be truncated during purification or had been posttranslationally cleaved *in vivo*. A new PCR approach was developed to clone the upstream sequence. Prokaryotic mRNA contains a specific ribosome binding site called *Shine-Dalgarno* sequence (SD) locating usually 7 to 9 nucleotides upstream of the initiator "AUG" of a structure gene (Gold *et al.* 1981). Its consensus sequence is:



A hexameric primer, with the same sequence as the SD sequence at its 3'-end, was designed to clone the 5'-end of the *afp* gene. The 5'-end of this primer was highly degenerate (Table 2-4). This primer was assumed to anneal to the corresponding sequence on the template chain of genomic DNA.

When the 137-bp first round PCR result was sequenced, two nested reverse specific primers were designed from it. One of the reverse primers was *asmr3* annealing at a site close to the 3'-end of the cloned fragment. The SD-sequence primer was labelled as *sdu* for Shine-Dalgarno sequence-like ubiquitous primer. With *sdu* pairing with *asmr3*, three strong bands were amplified (Figure 3-27). But the largest one was demonstrated to be a mismatch by PCR using *sdu* vs. *asmr4*, and led to amplification of a smaller fragment inside the former pair. Both primer pairs produced two specifically amplified strong bands indicating that there are two SD sequence analogs in the neighbour upstream sequence of the *afp* gene. All of the major bands were purified and sequenced. The largest sequence was 453 bp. Two *sdu* annealing sites were shown to be the two major

Results

bands in each PCR reaction. One is at –280, another, generating the small fragments, was at –7 (The “A” of the starting triplet is dictated as position “1” for numbering the nucleotides of the sequence). The latter was supposed to be the actual SD-sequence of this gene (Figure 3-28).

```
GGAGTGAGGAACTGCGCCTGCTTGCTGCGATCGCCGATCTCGATGCGCGACCGATCGC 57
  sdu
CGCGGACGACGGCGTTGTACCAGACGCTCGACAGCTCGCCGATGGTCACATCGCCCAC 114
CAGCCGCGCTCCGTCCGCGACGAAAACCGTCGGATGCAGCCGCGGCCGGACTCGCCC 171
AGCGGCAGGATCGTTGCGCTCGATGAGACGACCATGGCCTCACGTTACCCCTGGTGAG 228
CCCAGCCTTCTCTTGCTTGCAGATGTCCTCCGGCGTGAGTAGCGTTTGGACAGATGCAG 285
GCCACGACCTGCCAGAAGGAGCGCAGACATGACCGACACCAACATCACCACCCCGCC 342
  sdu
CTCACCGCCGACCCCGAGGTCGCCGCCGCTGCCGCGCAGTTCCTCACTCCGGTTCGTGC 399
  asmr4
ACAAGATGCAGGCGCTCGTCGTCAACGGCAAGCAGGCGCACTGGAAC 453
  asmr3
```

FIGURE 3-28. Sequence cloned by SD-PCR. The annealing sites of primers are underlined with the name of primer indicated. The starting triplet of the gene is shown in bold and underlined. Putative transcription factors are enclosed in rectangular boxes.

A putative starting triplet was found next to the codon encoding the first threonine of *eda-1*. The corresponding methylated methionine had apparently been removed after translation. Besides of these translation factors, an additional start codon as well as a “TATA” box and a –35 box, which compose the putative transcription machinery has been proposed from this sequence (Figure 3-28). Whether these factors are components of the promoter *per se* needs to be tested by a promoter detection vector.

3.7.2.3 3rd round PCR with degenerate primer

Three gene-specific nested primers annealing to sites close to the 5'-terminus of the SD-PCR cloned DNA fragment were coupled with degenerate primers designed from the other peptide sequences. Only one of them, namely *asmr5* from the C-terminus of the peptide *msa-4* gave strong signals (Figure 3-27). All of these bands were purified from the gel and sequenced. A 557-bp fragment was amplified by these primer pairs. The *asmf4* primer locates at a position, which is 94-bp inside the 5'-end of the 2nd round DNA sequence. When the sequences were aligned, a 650-bp fragment was then identified. The PCR product of *msf4* vs. *msr5* were then used as a probe in the following Southern blotting to clone the entire gene.

All of the ESI-MS/MS sequences were located within the translated peptide from this sequence. But the deduced peptide contained only 108 amino acids. Apparently its C-terminus part was still missing.

3.7.3 Cloning the entire *afp* gene

Genomic DNA of *M. arborescens* SE14 was digested with different restriction enzymes. Fragments separated by electrophoresis were then transferred on to Hybond-N+ membrane (Amersham Biosciences). The 557-bp PCR product was labelled with $\alpha^{32}\text{P}$ -dCTP and hybridized to the blot for 5 days. Only the *Pst*I, *Not*I, and *Sac*I digested fragment showed hybridised signals. The sizes of the corresponding fragments were 19.0 kb, 2.2 kb, and 2.3 kb, respectively.

The digested fragments of *Not*I and *Sac*I corresponding to 2.2-kb and 2.3-kb fragments were sliced and purified from the gel. They were ligated into pBluescript II SK phagemid vector (Stratagene) digested with the same enzyme. 96 colonies of the *Sac*I fragment-cloning library and 288 colonies of the *Not*I fragment-cloning library were screened. 5 μl of the overnight cultures were dipped on nylon membrane and lysed. The membranes were hybridised with the probe of labelled PCR product and detected with the ECL labelling kit thereafter.

Four colonies were positive, namely 1E12, 1G12, 3F11, and 4C5. The plasmids were extracted from these positive clones and sequenced. All of them contain the probe sequence. The *Sac*I cutting sites on both ends are very close to the *Not*I cutting sites. On the 5'-direction of the DNA sequences, the *Not*I site locates 19 bp outside of *Sac*I site, while on its 3'-end, the *Not*I site locates 118 bp inside of the *Sac*I site.

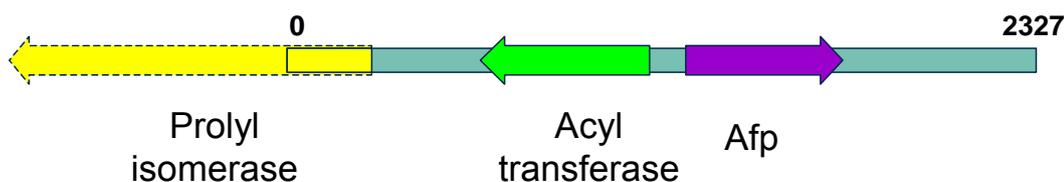


FIGURE 3-29. Presentation of the 1G12 sequence. The dashed arrow represents a proposed ORF based on sequence similarity.

The whole insert of plasmid 1G12 was then sequenced. The insert was 2.327 kb. There were two complete open reading frames (ORF) and a truncated ORF on this DNA fragment. The *afp* gene was in the middle and close to its 3'-terminus. Another gene encoded a putative acyltransferase gene on which two conserved acyltransferase modules were located. The first one started at position 77 and ended at position 94, the E-value for the hexapep domain (Pfam) is 0.034. The second one is at position 100 to position 117, the E-value for the hexapep domain (Pfam) is 0.027. The truncated gene is the N-terminus part of a prolyl isomerase gene (Figure 3-29). The *afp* gene is separated from the putative acyltransferase gene with only 111 bp. They are oppositely oriented.

3.7.4 Analysis of the *afp* gene

The cloned *afp* gene was analysed with the software DNASTAR (DNASTAR, Inc). Since the starting methionine residue was removed after translation, the mature protein consisted of 160 amino acids. All of the peptide sequences generated from purified protein had been located on the protein sequence translated from this gene (Figure 3-30 underlined). Surprisingly, 48% of the amino acids composing the Afp protein are hydrophobic. The Mr and pI values predicted by the DNASTAR program are 17.136 kD

Results

The comparison of Afp with some Dps family representatives shows that the ferroxidase center is also conserved in Afp (Figure 3-30 asterisks). Except of these critical residues, Afp is less conserved throughout the whole sequence, in particular, in the unique *N*-terminal region (Figure 3-30). These results suggest that the ferroxidase center of the Afp is either not involved in the enzymatic active center or, at least, not enough to catalyse the hydrolysis/formation of the amide bond.

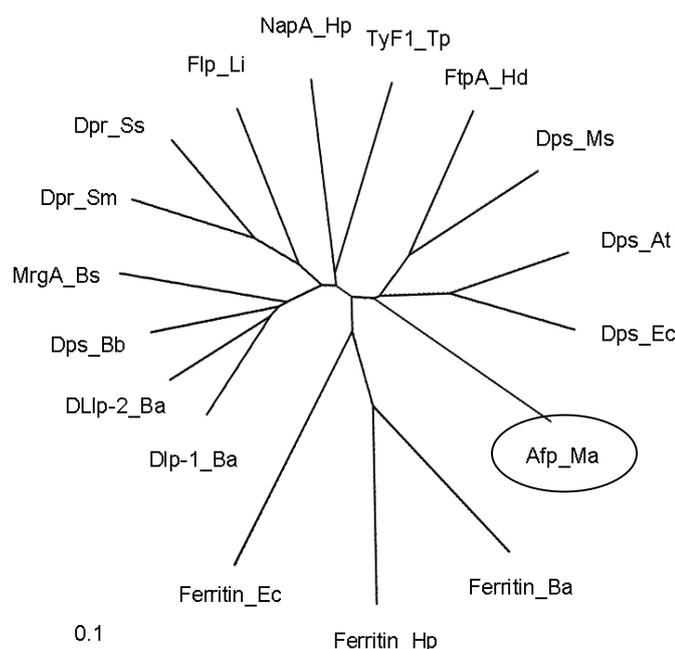


FIGURE 3-31. The unrooted phylogram based on maximum likelihood methods in which branch lengths indicate the relative evolutionary distance between isolates. The proteins is labelled with the first letter of the bacterial genus and species names following (clockwise): Dps-like proteins from *Bacillus anthracis* (Dlp-1_Ba and Dlp-2_Ba), Dps from *Bacillus brevis* (Dps_Bb), MrgA from *B. subtilis* (MrgA_Bs), Dps-like peroxide resistance protein from *Streptococcus mutans* (Dpr_Sm) and Dpr from *S. suis* (Dpr_Ss), Flp from *L. innocua* (Flp_Li), NapA from *H. pylori* (NapA_Hp), TypF1 (4D) from *Treponema pallidum* (TypF1_Tp), Fine tangle pili from *Haemophilus ducryi* (Ftp_Hd), Dps from *M. smegmatis* (Dps_Ms), Dps from *Agrobacterium tumefaciens* (Dps_At), and Dps from *E. coli* (Dps_Ec), the Afp protein purified in this research was circled. Three ferritins from some of the species where the Dps-like protein had been identified were chosen as rooting out-group for comparison.

The *E. coli* Dps protein was therefore employed to test whether or not the reversible amide formation is a common capability of this kind of protein. Two kinds of Dps extract were used, one was the wild type Dps, a protein fraction from *E. coli* BL21; another was purified from *E. coli* ZK1100 carrying a Dps overexpression plasmid pJE100 (Almirón *et al.* 1992). When the Dps protein isolated from *E. coli* was supplied with free linolenic acid and glutamine, no conjugation could be detected in both Dps preparations.

The Dps-like proteins can be grouped into two different classes according to their sequence similarity. Members isolated from Gram-positive bacteria show more homology, likewise the two from Gram-negative bacteria and the two from spirochaetes (Figure 3-31). Afp was isolated from a Gram-positive bacterium, but showed a little higher homology to the Gram-negative group, although this similarity is not very significant. Even the Ms-Dps isolated from mycobacteria, which belong to the actinomycetes like *Microbacterium* spp., was not grouped into same cluster with Afp. This result indicates that Afp has evolved through a distinct line from other Dps-like proteins, coincident with its unique biotransformation activity.

3.8 Expression of the *afp* gene

The *afp* gene was expressed heterologously and homologously to finally proof its identity and function.

3.8.1 Heterologous expression in *E. coli*

As shown before, neither the wild type *E. coli* nor the over-expressed Dps protein has activity on *N*-acyl amino acid synthesis (see 3.7.4). Taking the advantage of this clear biochemical background, the *afp* gene was cloned into pET11a vector under the control of a T7 promoter and terminator (see appendix I). The plasmid was named pEM1i2. Plasmid pEM1i2 was transformed into *E. coli* BL21-CodonPlus (DE3)-RP competent cell (Stratagene).

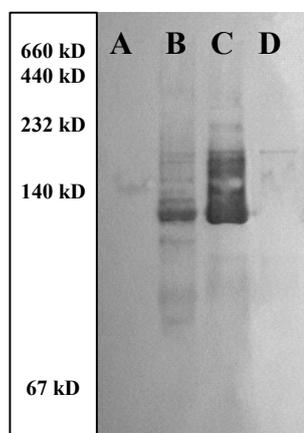


FIGURE 3-32. Detection of native and *E. coli* over-expressed Afp on native gel with a polyclonal antibody. (A), protein molecular marker, only the band of 140 kD is shown; (B), partially purified Afp protein fractions from *M. arborescens* SE14; (C), partially enriched overexpressed Afp protein fraction from *E. coli*; (D), partially purified overexpressed Dps protein from *E. coli*.

After IPTG induction, a large amount of soluble Afp protein was synthesized. This protein formed hexameric aggregates on native electrophoresis gel with the same apparent Mr of 100 kD as the native protein purified from *M. arborescens* SE14. Both the overexpressed and the native proteins were selectively stained by polyclonal

Results

antibodies (Figure 3-32). The polyclonal antibodies are highly selective, although there is certain homology between Afp and Dps, but no significant cross-reaction had been detected in these assays.

Surprisingly, the over expressed Afp protein could not catalyse the conjugation of linolenic acid and glutamine. Only a trace amount of *N*-linolenoyl glutamine was detected when the *E. coli* expressed protein was supplied with ferric ions *in vitro*. In the iron supplemented assay, the metal ion was added to the protein solution and incubated on ice for 5 min. The unbound iron was removed by ultrafiltration, because high ion strength inhibits the enzyme activity (see 3.3).

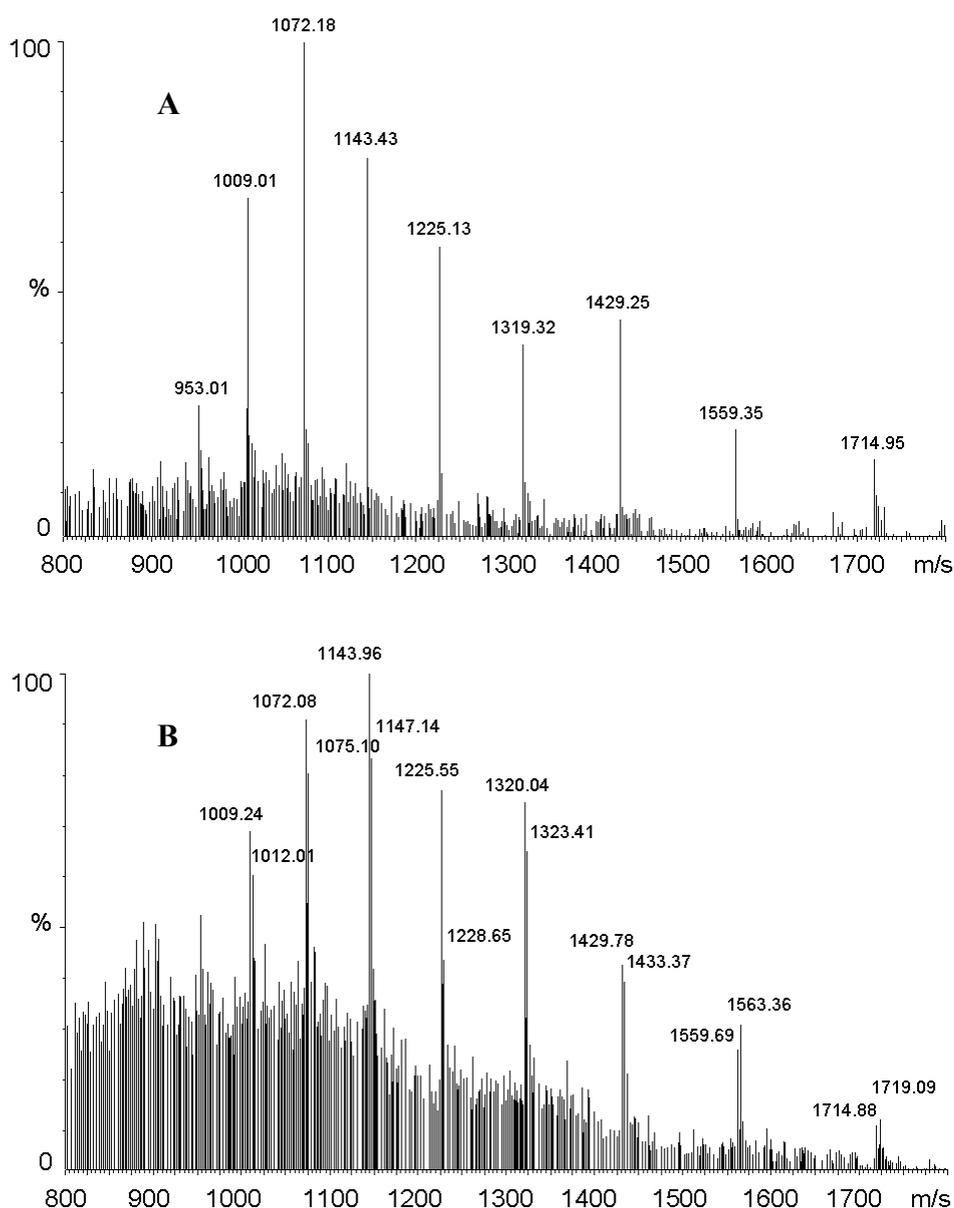


FIGURE 3-33. The MS profile of Afp protein overexpressed in *E. coli* (A) and native Afp purified from *M. arborescens* SE14 (B). Mass/charge ratio of the corresponding ions are shown.

To elucidate the difference between the protein expressed in *E. coli* and the protein purified from *M. arborescens* SE14, both samples were denatured in 50% methanol and subjected into Quattro-ESI-MS to measure their exact Mr. The MS results indicated that in the *E. coli* fraction, there is only one kind of protomer. The Mr calculated from the m/z was 17138.22 ± 1.85 Da (Figure 3-33 A), fully coincident with that deduced from the cloned gene (17136.34 Da).

On the other hand, there were always double peaks in the protein from *M. arborescens* SE14; the two Mrs calculated from the two parallel ions were 17142.84 ± 3.66 and 17187.33 ± 3.51 (Figure 3-33 B), respectively. The latter is the major component, which had been detected by MALDI-TOF. The larger subunit is 42.84 ± 3.98 Da bigger than the small one. The smaller one is identical to *E. coli* over expressed protein. The difference is exactly the molecular weight of an acetyl group (43.02 Da). Whether or not the difference on molecular mass is due to an acetylation needs further proof.

Even the larger subunit is much smaller than the pro-Afp directly transcribed from the gene. With the starting methionine, the pro-protein mass should be 17267.54 Da. In bacteria the methionine is usually methylated, the protein mass would then as high as 17281.57. The Quattro-ESI-MS data suggested that in *E. coli* the methionine had been removed after translation like in the original host. It is known that in bacteria, the *N*-terminal residue of the newly-synthesized protein is modified to remove the formyl group. The *N*-terminal methionine may also be removed (Stanier *et al.* 1987).

3.8.2 Expression in *M. arborescens* DSM20754

The low activity of *M. arborescens* DSM20754 (see 3.2) left a potential for additional expression of the *afp* gene in this organism. An expression plasmid was constructed and successfully transformed in *M. arborescens* DSM20754 protoplasts. Although recombination occurred in each transformants, the truncation in most cases did not influence the activity enhancement.

3.8.2.1 The *afp* gene ortholog in *M. arborescens* DSM20754

Because *Microbacterium* is a rarely studied genus, the genetic background is not very clear. The *afp* gene specific nested primers, *asmf4*, *Afc1*, and *Arc3* were used to detect this gene in *M. arborescens* DSM20754 with PCR. Coincident with the very low amide forming activity in *M. arborescens* DSM20754, we didn't obtain any signal in PCR under the same conditions that were used for cloning *afp* from *M. arborescens* SE14 in which the buffer and Taq polymerase from Invitrogen were used, and elongation was performed at 72°C. When the PCR was performed with buffer and Taq enzyme from NEB (New England Biolab) and with elongation at 75°C, an orthologous gene of *afp* was amplified in *M. arborescens* DSM20754 (Figure 3-34).

The gene was labelled as Ma-DSM-*afp*. The NEB enzyme could tolerate to certain extent mismatching of the primers. Amplification from *M. arborescens* DSM20754 DNA template favours the higher annealing temperature, coincident with the sequencing result that its GC content is 0.1% higher than that of the *afp* gene from *M. arborescens* SE14. The Ma-DSM-*afp* gene is different from *afp* on DNA and protein sequence. We still do

not know whether the altered residues in Ma-DSM-*afp* are responsible for the low activity detected in *M. arborescens* DSM20754, or there is no difference on the activity of these orthologs, and it is the result of different regulatory factors which enhanced the expression of Ma-DSM-*afp*. These hypothesis will be tested in future experiments. However, the absence of identical *afp* in *M. arborescens* DSM20754 makes the expression and confirmation of *afp* in this host easier.

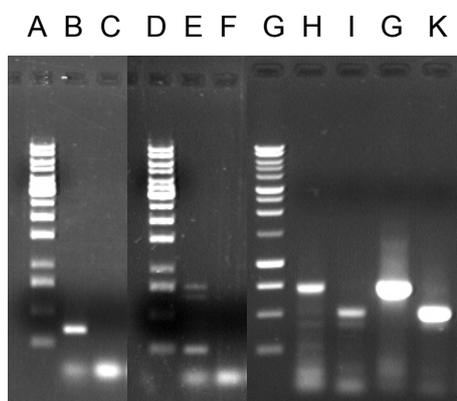


FIGURE 3-34. The PCR results of *afp* gene detection with nested primers. A, D, and G are DNA markers; B, C, E, and F were elongated at 72°C with Invitrogen enzyme and primers, Afc1 vs. Arc3 (B and C) and *asmf4* vs. Arc3 (E and F); H, I, G, K were elongated at 75°C with NEB Taq and primers, Afc1 vs. Arc3 (I and K) and *asmf1* vs. Arc3 (H and G). The DNA templates are from *M. arborescens* SE14 (B, E, H, and I) and from *M. arborescens* DSM20754 (C, F, G. and K).

3.8.2.2 Plasmid construction

Because there is no expression system in *M. arborescens* available, a new plasmid was constructed. In the expression plasmid, pRMM, a *Bacillus subtilis veg* promoter and a T7 terminator were engineered into an expression cassette to facilitate the *afp* gene expression (see appendix I). The open reading frame of the *afp* gene and the following T7 terminator has been subcloned from pEM1i2 and inserted into the downstream of a *B. subtilis veg* promoter. Therefore, the expressed protein would be the same as in *E. coli* BL21 CodonPlus, provided that the *veg* promoter would function properly in *M. arborescens*. The replication origin of the Streptomyces low copy plasmid SCP2* supplied by pSEK4 (a derivative of pRM5) ensures stable replication in *Microbacterium*.

A control plasmid pRMC had been constructed in this work. The only difference between expression and control plasmids is, that the *afp* gene is missing in control plasmid. The *veg* promoter and T7 terminator were still on the control plasmid. This secured that any activity increase observed in pRMM transformants can be considered as the result of *afp* gene expression. The thiostrepton resistant gene on both pRMM and pRMC rendered the transformants able to grow on plates supplemented with thiostrepton, while the wild type has been shown to be very susceptible to this and other antibiotics (see appendix II). The correct arrangement of the plasmid sequences was confirmed by restriction enzyme digestion (Figure 3-35) and subsequent sequencing with the universal primer T7-term (see 2.2.8).

3.8.2.3 Transformation of *M. arborescens* DSM20754

To bypass the obstacle of the methyl-specific restriction system widely distributed in streptomycetes (Kieser *et al.* 2000), the plasmids have been demethylated by transformation into methylation deficient *E. coli* INV110 (*dam*⁻ *dcm*⁻, Invitrogen) prior to introduction into *M. arborescens* DSM20754. When the plasmid was prepared directly from *E. coli* INVαF', the *dam*⁺ *dcm*⁺ host, no transformation occurred.

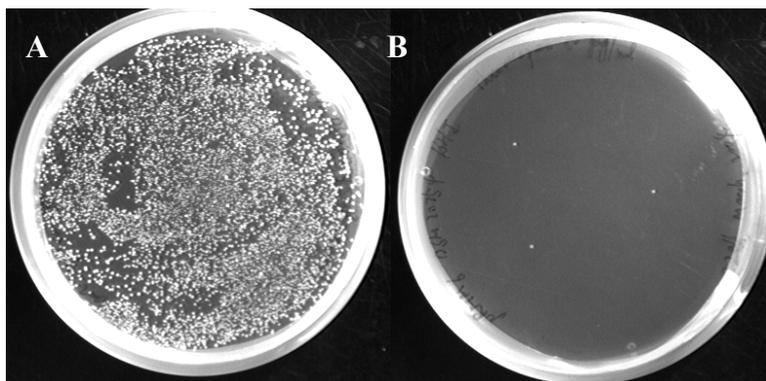


FIGURE 3-35. Colonies of *M. arborescens* DSM20754 developed from protoplasts after transformation. (A): Colonies developed on a normal BHI plate after protoplast transformation. (B): Three colonies developed on a plate supplemented with 50 µg/ml thiostrepton.

After two days of growth on selective plates, colonies emerged on the medium surface. Thousands of colonies developed from transformed protoplast if the bacterium was spread on a non-selective BHI plate control. In contrast, on antibiotic selective BHI plates, usually only three or four colonies grew on each plate (Figure 3-35). The correct introduction of *afp* into *M. arborescens* host was confirmed by PCR with the primers Ande1 and Abam3 (see 2.2.8) under conditions at which only *afp* could be amplified. The *afp* specific band of 500 bp was strongly amplified in most transformants. The presence of the whole *afp*-expression cassette in the four selected colonies was confirmed by nested primer PCR using Afc1, Arc3, t7-term, and vegprom.

3.8.2.4 Increase of *N*-acyltransferase activity

The pRMM transformed *M. arborescens* strains showed an enhanced *N*-acyl amino acid production, whereas the strains transformed with the control plasmid, pRMC did not (Figure 3-36). Two kinds of colonies were detected in pRMC transformed *M. arborescens* DSM20754. Some of the regenerated colonies still retained the low level of catalytic activity of the wild strains, while a small population lost the catalytic activity for unknown reason. In four tested pRMM plasmid transformants, the activities of three colonies were 2.5 times higher than the pRMC transformants when the strains were cultivated for 20 hour. Only in one sample, statistically insignificant difference to the activity of pRMC strains had been occasionally observed.

Results

Because the amide forming activity in wild *M. arborescens* strains was dependent on growth stage, the enzymatic activities of both pRMM and pRMC transformed strains were monitored throughout their growth. For unknown reason, the growth rate of all these thiostrepton resistant strains was much faster in antibiotic containing medium than that of the wild strain in normal medium. The resistant strains needed only 12-hours growth to reach an OD₆₀₀ of 2.0 (Figure 3-36). This might be partially due to the good air access in this assay. In the latter, the bacteria were grown in 40 ml Falckon tubes, while all of the other measurements were done in 3 ml batches. Not only the bacteria grew faster, but also their cell density reached a much higher level and could not be measured by the photometry in the later exponential growth and thereafter. Therefore the bacterial culture was 2 times diluted prior to measurements at this time.

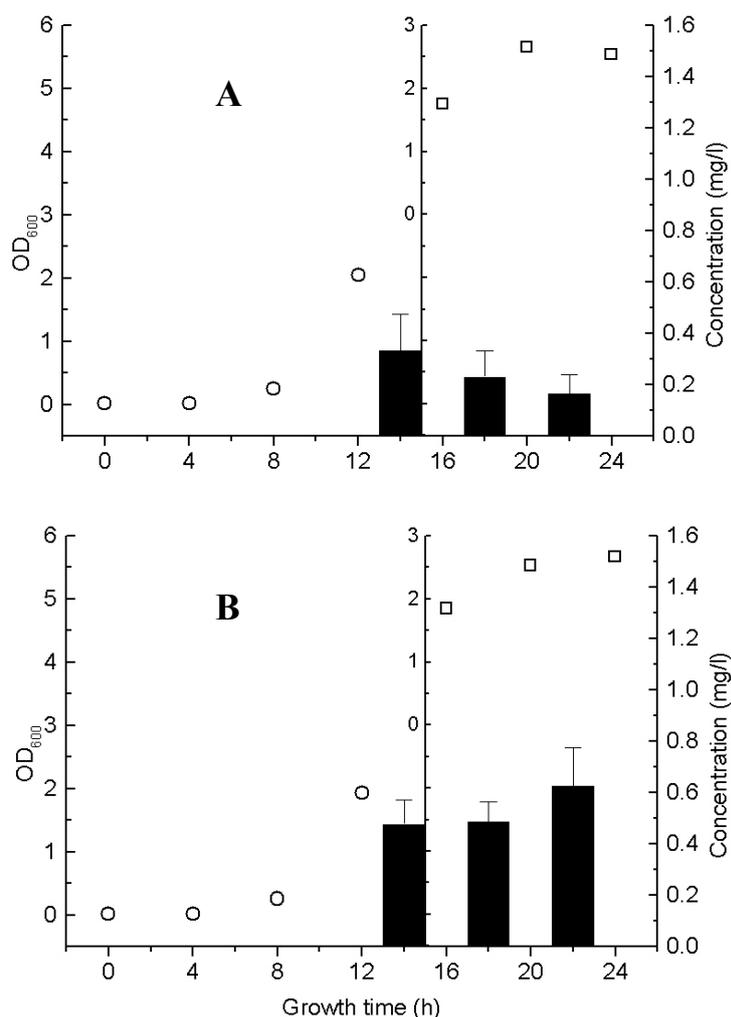


FIGURE 3-36. Time-dependent enzyme activity of a control plasmid pRMC transformant (A) and an expression pRMM transformant (B). The cell density of cultures was measured with OD₆₀₀ (hollow circles). Because the cell density was too high to be measured after 16 hours growth, the OD₆₀₀ was determined after dilution by a factor of 2 (hollow squares), which was shown using second axis in the middle of the diagram. Average product concentrations and deviations were calculated from more than 5 separate repetitions using linolenic acid and glutamine as substrates (solid bars).

Results

Nevertheless, the maximum enzymatic activity of pRMC transformed bacteria was detected exactly when OD_{600} was around 2.0 like in the wild type. If the bacteria were allowed to grow beyond that point, the enzyme activity decreased quickly (Figure 3-36). It seems to be that it is not the ongoing of stationary phase, but the cell density of OD_{600} of 2.0 serving as a checkpoint for the corresponding gene activation. On the contrary, the enzyme activity of pRMM transformed bacteria maintained at relatively high level throughout the measurement. When the OD_{600} of the bacterial cultures was around 2.0, where the catalytic activity of pRMC was highest, the difference between the activity of pRMC transformants and that of pRMM transformants wasn't statistically significant. But during the later growth, the catalytic activity of pRMM strains didn't decrease like pRMC strains, but increased a little. Unexpectedly, we didn't detect any enzymatic activity in both pRMC and pRMM strains before their cell density reached OD_{600} of 2.0. Probably due to their extraordinary high duplication rate in exponential growth, the substrates were metabolized as nutrient no matter the biotransformation occurred or not.

Interestingly, the *afp* gene was found not only to increase linolenoyl glutamine production, but also promoted bacterial cell aggregation. The overnight cell culture of pRMM transformed strains became flocculent after addition of the substrates. This structure was not as thick as the clag formed by *M. arborescens* SE14. Normally the phenomenon could not be observed in wild type *M. arborescens* DSM20754 and the strains transformed with pRMC. Only when they were cultivated for more than 28 hours, they formed tiny patches, which were easily homogenized by shaking.

Furthermore, transformation with pRMM made the bacterial cell wall more fragile than that of the wild type or those transformed with pRMC. It took 6 minutes with sonication at 60 % power to totally break the cell of *M. arborescens* SE14 and all pRMM transformants, while 16 min were required to break the cell of wild type *M. arborescens* DSM20754 and pRMC transformed strains. Whether this is the consequence of the produced *N*-acyl amino acid or the Afp protein itself deserves further investigation.

Finally the plasmid is instable in *M. arborescens* DSM20754. The increased activity in the transformants as well as the thiostrepton resistance was easily lost when the bacterial culture became old, e. g. stored at 4°C for more than 5 days or cultivated at 37°C for more than 48 hours. This could account for the difference of activity between control and expression plasmids transformed bacteria was always observed between repetitions.

4 Discussion

N-Acyl amino acids have never been detected in eukaryotic digestion systems. Some Lepidopteran species, like *Heliothis virescens* and *Helicoverpa zea*, have different regurgitant composition from others, in part due to degradation of *N*-acyl amino acid in midgut lumen by an extracellular enzyme (Mori *et al.* 2001). The hydrolyzed products are no longer elicitor-active (Alborn *et al.* 2000). The different *N*-acyl amino acid profile might be also due to the different fatty acid profiles in the insect food, or different substrate selectivity of the synthetic enzyme(s) (see 1.1).

In any case, the different *N*-acyl amino acid composition leads to a different volatile signature which influences the attraction of parasitic wasps (De Moraes *et al.* 1998). The presence of synthetic/hydrolytic enzymes from microorganisms indicates that the *N*-acyl amino acids pool in insect gut is controlled precisely and maintained by a complex system.

4.1 Screening of the commensal gut bacteria

Strains representing diverse bacterial groups were selected in this study to investigate their capability to produce *N*-acyl glutamines. Two of them, namely *P. rettgeri* and *S. aureus* are well known mammalian pathogens. The gram-negative *P. rettgeri* is a pathogen causing nosocomial infections. *S. aureus* is an important pathogen of warm-blooded animal species. Both have been isolated simultaneously from oil fly (*Helaeomyia petrolei*) larval gut (Kadavy *et al.* 2000) and from the gut of Mexican fruit flies (*Anastrepha ludens*) (Kuzina *et al.* 2001). *S. aureus* has been reported to be capable of killing the nematode *Caenorhabditis elegans* when accumulating in its digestive tract (Sifri *et al.* 2003). It is possibly also an insect pathogen. *P. rettgeri* has been isolated from house fly's (*Musca domestica*) intestinal tract, and it has been demonstrated to be necessary for the larvae to support its development (Zurek *et al.* 2000).

Ochrobactrum sp. is an endophytic bacterium (Verma *et al.* 2004), and also a free living soil inhabitant (Lebuhn *et al.* 2000). *B. pumilus* is a common soil bacterium (Hallmann *et al.* 1999). It may lodge within plant tissues as a result of physical damage during growth in the field (Isenegger *et al.* 2003). *Microbacterium* spp. had been isolated from various niches, including the phyllosphere (Behrendt *et al.* 2001), rhizosphere (Abou-Shanab *et al.* 2003), sponges (Lang *et al.* 2004), and termites (Wenzel *et al.* 2002) *etc.* Owing to their wide distribution, they could easily, by feeding or other processes, opportunistically colonize the insect gut.

4.1.1 Different catalytic properties

According to the HPLC-MS screening results, some of the gut bacteria were able to produce the typical *N*-linolenoyl glutamine when supplied with free linolenic acid and glutamine, coincident with previous results (Table 4-1). Some of them additionally produce several structurally uncharacterized, probably related compounds (Figure 3-1). For example, in addition to the typical *N*-linolenoyl glutamine, *M. arborescens* SE14 produces a small amount of a compound with same m/z, but higher polarity. The

Ochrobactrum sp. A13 produces three peaks with identical MS profiles. Besides the two fractions found in the preparation of *M. arborescens* SE14, a third peak was detected. But the chemical nature of these probably isomeric peaks has to be confirmed by further studies.

Table 4-1. Some bacterial strains isolated from *S. exigua*.

Species	16S rDNA identity ^a	Origin of standard stain	<i>N</i> -acyl-glutamine production ^b
<i>Bacillus pumilus</i>	100 %	Soil	+
<i>Rhodococcus erythropolis</i>	100 %	Soil, insect gut	+
<i>Enterococcus rottae</i>	99 %	Digestion tract	-
<i>Microbacterium arborescens</i>	99 %	Environment	+
<i>Ochrobactrum</i> sp. PR17	97 %	Rhyzosphere	+
<i>Escherichia coli</i>	100 %	Digestion tract	-

^a The data are based on the alignment of cloned sequences and that of the standard strains on NCBI; ^b The product was measured after addition of substrates 4 h. '+' indicates production of AAAs in stationary-phase growth, whereas '-' indicates no AAAs production had been detected. Reproduced according to Spiteller 2002.

Neither *P. rettgeri* nor *S. aureus* SE8 produce the putative isomers. The activities of *P. rettgeri* and some other bacteria is associated with the bacterial cell, either in the cytosol or associated with the cell wall. Their activities were detected mainly in crude extracts (Figure 3-2). The active enzyme seems to be secreted only by *S. aureus* SE8 into the medium. It is known that Staphylococci, especially *S. aureus*, produce a wide variety of exoproteins. Some are toxins, others contribute to pathogenicity by attacking the intercellular matrix or attack host cells directly. They are accessory proteins which are synthesized at the end of exponential growth and in early stationary phase (Novick 1993).

4.1.2 Growth stage-dependent activities

The enzymatic activities of all of the tested gut bacterial strains increased dramatically within a quite short time window at the early stationary phase, although they all displayed different growth kinetics (Figure 3-3). At the on-set of the stationary phase, nutrient depletion occurs. Such stringent conditions appear to be general for bacteria colonizing higher organisms. A number of bacterial genes, especially those responsible for colonization and pathogenicity, are known to be regulated by quorum sensing systems (Parsek and Greenberg 2000), which control the bacterial genes and allow them to be expressed only at the stationary growth phase. Carbon starvation conditions thereby have been used as a model to study the physiology of mycobacteria persistence in hosts (Gupta *et al.* 2002). Whether the gene responsible for *N*-acyl amino acids biosynthesis is also controlled by quorum sensing system is unknown and requires further research.

4.1.3 Selection of the most active strain

The production of *N*-acyl glutamine varies among species. *M. arborescens* SE14 showed the highest activity among all tested bacterial isolates (Figure 3-2 and 3-3). The

culture of *Ochrobactrum* sp. A13 showed lowest activity. *M. arborescens* SE14 (Figure 4-1) was therefore chosen for further investigations.



FIGURE 4-1. Electronmicroscopic image of *M. arborescens* SE14. Photo: Dr. Rita Büchler, MPI for Chemical Ecology and Dr. Katrin Buder, Institute of Molecular Biotechnology, Jena, Germany.

4.2 Evidence of selectively enhanced activity

M. arborescens was originally described by Frankland in 1889 as *Bacillus aborescens* and published in Bergey's Manual of Determinative Bacteriology as *Flavobacterium aborescens* in 1923. In 1984 the strain was reclassified as *Microbacterium aborescens*. From the name it can be assumed that the strain was isolated from water or soil environment (personal communication with Dr. R. Pukall, DSMZ, Braunschweig, Germany). The reference strain of this genus is *M. arborescens* DSM 20754.

The 16S rDNA sequence of the strain isolated from beet armyworm gut, *M. arborescens* SE14 (NCBI, AY649756) is 99% identical to that of *M. arborescens* DSM 20754 (EMBL, X77443). The degree of relatedness between the *Microbacterium* species is between 94.5% and 98.4%, which is significantly lower than for other B-peptidoglycan containing genera (Rainey *et al.* 1994). According to the 99% similarity of 16S rDNA sequence between *M. arobescens* SE14 and *M. arborescens* DSM20754, it is reasonable to group them into same species.

The amount of *N*-linolenoyl glutamine produced by *M. arborescens* SE14 is about 20 times higher than that of the reference strain (Figure 3-4). Interestingly, *M. arborescens* DSM20754 produced significant amounts of the unidentified structurally related byproduct. The difference between the two strains might be either due to different catalytically active proteins, or due to an unknown control mechanism. In fact, we found an orthologous gene in the type strain genome, which is highly homologous to the gene identified in *M. arborescens* SE14 (see 3.8.2.1). It is known that insect gut microbiota well adapt to their special niches. The gut bacterial population of the American cockroach (*Periplaneta americana*) is closely linked to the insect diet (Dugas *et al.* 2001). No matter what is the underlying mechanism of the observed activity difference, it

could be a kind of adaptation to live in the insect gut. It could be also due to that the *M. arborescens* DSM20754 bacteria decreased this ability during the long-time cultivation under laboratory condition, which exerts less selective stress on this activity.

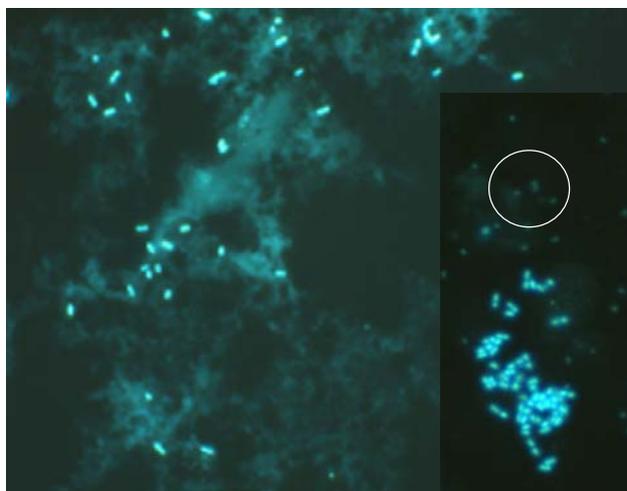


FIGURE 4-2. The insect gut is a niche for bacteria populations. The sample was prepared according to the FISH protocol (see 5.1). The insect gut lumen was washed out and fixed by paraformaldehyde. After staining by DAPI, slides were observed under the fluorescence microscope (X 800). The blue spots are bacteria. There are bacilli (majority), gram-negative bacteria (circled), and a cluster of cocci (lower right corner). The background is the undigested food.

The genus *Microbacterium* comprises irregularly rod-shaped bacteria (Collins and Bradbury 1991). No morphological difference between the two strains has been found under the microscope (Figure 3-5). On the other hand, the cell density of stationary phase-growing *M. arborescens* SE14 is higher than that of *M. arborescens* DSM20754. In addition, *M. arborescens* SE14 grows much faster than *M. arborescens* DSM20754. There are no obvious special structures in the gut of lepidopteran larvae that are usually associated with microorganisms. A large bacterial population lives commensally in the gut lumen (Figure 4-2). Because of the rapid food throughput, the bacteria must grow fast enough to maintain a stable population (Dillon and Dillon 2004).

It has been proposed that the alkaline pH and other factors in lepidopteran larval gut exert stress on the indigenous gut consortia making the gut a “hot spot” for gene transfer (Dillon and Dillon 2004). No matter what’s the mechanism contributing to the difference of the enzymatic activity is, these data suggests that *M. arborescens* SE14 had acquired an ability not only to grow faster but also to produce much more *N*-acyl amino acids than its free-living cousin *M. arborescens* DSM20754. This may also indicate that *N*-acyl amino acids are somehow involved in the interaction between the gut bacteria and their insect hosts.

4.3 Purification of the amide-forming protein

The catalytic activity of *M. arborescens* SE14 was found to raise at the early stationary growth and was mainly found in extracts (Figure 3-7). An active protein

fraction was purified from the bacteria harvested at this time. The proteins specific activity was increased 186 fold in an overall yield of about 26% of total activity after purification (Table 3-1).

Different purification methods, ammonium sulphate precipitation, anion exchange chromatography, and gel filtration coincidentally indicated that there was only a single active fraction in the cell extract. Furthermore, both ultrafiltration and size exclusion chromatography suggested that the protein mass was very large. This unusual character as well as other features made the purification easier than expected, e.g. (i) the protein was very abundant at the early stationary growth; (ii) the protein surface is highly charged under the experimental conditions. Because the purified protein catalyses the reversible amide formation through an unknown mechanism, it was henceforth named Afp.

4.4 Characterization of the Afp protein

In addition to gel filtration, FFF-MALS, MALDI-TOF, isoelectric focusing, and native PAGE all showed that the purified protein fraction consisted only of a single component. Each method has its own limitation, e.g. the MALDI-TOF can only measure protein mass between 0.5 and 200 kDa (personal communication with Dr. Svatoš). If there would be a contaminating protein with a Mr of 300 kDa, it would have not been detected by this method; the background of PAGE depends on the loading amount as well as destaining strength. However, small amounts of contaminating protein in this fraction can not be ruled out. In any case, the Afp protein is the dominant component in the purified protein fraction.

The protein has an interesting ability to form aggregates of different size. The molecular masses determined by FFF-MALS (203 kDa) and by MALDI-TOF (17.181 kDa) (Figure 3-13, 3-14) suggest that the Afp protein forms a homododecamer in solution. This is coincident with the result of the cloned gene (see 3.7.4). Further more, the dodecamer could aggregate randomly as suggested by the Quattro-ESI-MS profiles (see discussion in 3.4.1). The protein appeared to be hexameric on native PAGE (Figure 3-16). Other kinds of oligomers, such as trimers, tetramers, and pentamers were detected on SDS-PAGE (Figure 3-16). The relative abundance of these oligomers varied between the preparations. As will be shown later, these oligomers were probably due to random inter-subunit crosslinking produced by free radicals from iron-oxidation which is a common phenomenon in this kind of proteins (see 4.7.4.2).

The native protein Mr, estimated by gel filtration, is 310 kDa (Figure 3-12). This apparent Mr is higher than that measured by FFF-MALS and other methods. Later results indicated that Afp is a hollow cage-like protein (see 4.7.4), whose actual size is larger than a compact globular structure. Coincident with the previous observation that the protein was highly charged under neutral pH condition (the protein is very acidic), the measured pI is only 4.2 (Figure 3-17). Afp is homologous to bacterial Dps (see 3.7.4). It contains iron like the other Dps-like proteins (Ceci *et al.* 2003), but the iron content is very high, about 10 iron atoms per monomer. Only the native protein binds iron, the denatured protein is not able to bind iron (Figure 3-18). On the other hand, Afp does not bind DNA (Figure 3-19), coincident with its lack of the positively charged terminus, which is the functional motif required for DNA-binding (Ceci *et al.* 2003).

4.5 The catalytic properties of Afp

Afp catalyses a reversible formation of *N*-acyl amino acid without the need for activation. This reaction looks like a reverse amidase reaction, in which only a small amount of substrates could be converted into conjugates.

4.5.1 A rapid reversible reaction

As discussed above, the amide bond formation between a fatty acid and glutamine is an endergonic process that doesn't occur spontaneously (see 1.1.3). Afp catalyses the reversible conjugation of linolenic acid and glutamine with a very low transformation rate. The reaction K_{eq} is 0.1863. Denaturation made the protein totally inactive, suggesting that the secondary structure is important for the catalysis. On the other hand, the formation of *N*-linolenoyl glutamine is very fast, the k_{cat} (s^{-1}) at pH 8.0 is 0.6378. It is a little bit better than that of lysozyme (0.5) that usually works extracellularly.

The rate constant (K_m) for the amide synthesis was not detectable. While that of the hydrolysis was 36 $\mu\text{mol/l}$, which corresponds to a normal enzyme indicating that this Afp is a hydrolytic enzyme. Nevertheless, the enzyme could operate in the synthetic direction, if the products are removed by the insect or exported by the bacteria. The same strategy is employed by living organisms, *esp.* by bacteria, e.g. in the initiation of dissimilatory sulfate reduction. Sulfate is activated by esterification to an adenylyl group from ATP to form adenosylphosphosulfate and pyrophosphate. The equilibrium of this reaction lies strongly towards ATP and sulfate, the isolated reaction does not proceed to any significant extent. However, the pyrophosphate is rapidly hydrolysed *in vivo*, thereby allowing the reaction to continue (Stanier *et al.* 1987).

4.5.2 The temperature optimum

Afp was active in a broad temperature range around 40°C *in vitro*. The loss of activity at temperatures up to 48 °C is at least partly reversible (Figure 3-20). This matches with the poikilothermic environment of insect bodies, which can become very hot on sun-exposed leaves. The activity clearly depends on its secondary structure. If the protein was fully denatured at a temperature higher than 60 °C for 20 min, no activity could be restored. The protein is very stable on that it is resistant to a temperature as high as 48°C. It also sustains the presence of the detergent SDS, and even 10% ethanol in the solution.

4.5.3 The pH influence on Afp activity

The protein showed two pH optima, one at about pH 8.0 and another in the range from 9.2 to 11.2, respectively (Figure 3-21). The average enzyme activity at the more alkaline pH was two times higher than that of pH 8.0. The pH optima of catalytic activity are coincident with the pH values of the insect crop (7.9 ± 0.1) and midgut (9.3 ± 0.1) (Figure 3-21), from where the content could be regurgitated by the larvae (Lait *et al.*, 2003).

4.5.4 Characterization of the active centre

Because the reaction is faster at alkaline pH, the highest activity starts at around pH 9.2. The deprotonation of a tyrosine residue in active center could facilitate the activity enhancement. The absorbance of UV of wavelength around 255 nm increased when the pH of Afp suspension shifted from 8.0 to 10.0 (Figure 3-22). The UV wavelength shift caused by protonation/deprotonation of tyrosine may vary according to the molecular environment where tyrosine locates, from 245 nm (Clarkson and Smith 2001) to 295 nm (Rami and Udgaonkar 2001). It has also been demonstrated that metal binding of ferritin in animals results in the appearance of peaks with maximum absorbance at 242 and 295 nm in UV spectra. These spectra shifts are generally considered to be characteristic for a deprotonated tyrosine (Kumar and Prasad 1999).

The protein seems to be a hydrolytic enzyme that catalyses a reaction with a relatively high reverse reaction rate. This is more likely since it does not require ATP for substrate activation. It is known that a tyrosine residue is essential for the catalytic activity of metallo proteases such as aminopeptidase A and N (Vazeux *et al.* 1997). The protease inhibitor PMSF, which blocks tyrosine residues in active centres, did not show any effect on the catalytic activity of Afp when supplemented to the protein solution (Table 3-2). Therefore the tyrosine might not be an essential residue in the catalytic centre. Its function might be similar to the tyrosine residue in aminopeptidase A. The tyrosine residue stabilizes the transition state complex in catalytic centre. When it was replaced by other amino acids, the K_m value of reaction did not change, whereas K_{cat} value decreased (Vazeux *et al.* 1997). Surprisingly, the ion chelator EDTA also did not influence the enzymatic activity (Table 3-2).

The amide forming enzymes detected in cannabinoid formation in mammalian neurosystems and insect alimentary tissues do not require ATP for substrate activation (see 1.1.3). ATP is not required in this reaction (Table 3-2). DTNB was reported to be able to block the ATP- and CoA-independent *N*-acyl ethanolamine formation (Kruszka and Gross 1994) as mentioned before. In that reaction, a putative sulphhydryl group of a cysteine residue in the catalytic site is essential for the reaction. But DTNB does not inhibit the *N*-acyl amino acid formation by Afp. Besides DTNB, none of the other tested protein modifications, which have been used to modify the active center of hydrolases (Honda *et al.* 2002), blocked this reaction (Table 3-2). Therefore the catalytic mechanism of Afp deserves further investigations.

4.6 The substrate specificity

4.6.1 Specificity on amino acid

Of all 20 proteinogenic amino acids only proline was not incorporated into the corresponding products (Figure 3-23). This result seems to be not consistent to the *N*-acyl amino acid profile in insect oral secretion (see 1.1.1), where only conjugates of glutamine and in some cases glutamate were detected. Glutamine is the best substrate among those whose LogP was lower than 1.0. It is also the most abundant amino acid in plant tissue, because plants use this compound for nitrogen storage (Lamber *et al.* 1998).

The data indicate that Afp could be an enzyme that contributes to *N*-acyl amino acid formation in insect gut.

The incorporation efficiency of other amino acids depended on the hydrophobicity. The amino acids with large hydrophobic side chains were better substrates. In previously known bacterial lipoamino acids, the amino acid moieties had been identified to be serine, lysine, glycine, leucine and isoleucine, ornithine *etc.* But in certain species there were only one or a few amino acid conjugates detected (see 1.2.1). There might be enzyme(s) with different selectivity or other unknown mechanism(s) involved in their biosynthesis. The Afp only catalysed reaction with free L-amino acid as substrates. Dipeptides with glutamine on its *N*-terminus could not be used. Likewise, *D*-amino acids were not incorporated into corresponding amides.

4.6.2 Fatty acid selectivity

The preference of Afp for fatty acids is also very broad, coincident with the spectrum of fatty acid moieties in insect regurgitant. The incorporation efficiency of fatty acids depends on their hydrophobicity. A limit is reached, if the fatty acids become insoluble in the reaction medium (Figure 3-24). Dodecanoic acid was the best substrate among the tested saturated fatty acids. But this fatty acid is known to be rare in bacteria (Schweizer 1989). This result might only indicate the bacterial preference for short chain fatty acids.

The presence of hydroxyl groups and double bonds increases the solubility of fatty acids. The 18-carbon fatty acids with a Log P between 4.0 and 6.0 were transformed much more efficiently than the saturated octadecanoic acid. The incorporation ratio for the tested 18-carbon fatty acids reminiscent of the ratio of the conjugates in BAW gut content (Figure 3-25), and caterpillar regurgitant reported (see 1.1.1). The 17-hydroxyl linolenic acid present in volicitin from insect secretion, was the best substrate among all substrates tested. The incorporation efficiencies also reflected the solubility of the fatty acids, which might be the true underlying discipline. In BAW caterpillar foregut lumen, the average concentration of volicitin is *ca.* 200 μ M, that of *N*-Linolenoyl-L-glutamine is *ca.* 100 μ M. They are the two most abundant *N*-acyl amino acids. Glycerides, phospholipids, and acyl-CoA were not accepted as substrates. The reverse amidase, the apolipoprotein *N*-acyltransferase mentioned above (1.2.3.3), uses a phospholipid as an acyl donor, while standard acyl transferase use CoA derivatives (1.3.2). None of these could serve as the substrate in this reaction, indicating that the acyl transfer mechanism used by Afp is different from these systems.

4.7 Cloning the corresponding gene

4.7.1 Protein sequencing

The protein had been sequenced by Edman degradation and ESI-MS/MS sequencing after trypsin digestion according to published procedures (Shevchenko *et al.* 1996). A complete 22-amino acids peptide at the amino terminus of the protein was found, while the 4 other sequences were located somewhere in the middle of the protein. Degenerate PCR primers were then designed from these sequences.

4.7.2 Nested degenerate primer PCR

The GC content of *Microbacterium* DNA is as high as 69 to 75 mol% (Collins and Bradbury 1991), and the bacterium has a thick polysaccharide layer on its cell wall (Figure 4-1). This makes the isolation of DNA from this species difficult. To overcome these obstacles, a modified CTAB assistant genomic DNA isolation procedure according to (Kieser *et al.* 2000) was developed to isolate genomic DNA from *M. arborescens* SE14. The modified procedure (see 2.2.7) is very effective.

Degenerate primers designed from protein sequences, or from conserved regions of related genes have been widely used for amplification of novel genes (Shen *et al.* 1993), and for screening orthologous genes in diverse phylogenetic populations (Hamann *et al.* 1999). Because deoxyinosine (I) can base-pair with all four naturally occurring deoxyribonucleotides (Shen *et al.* 1993), it has been used to reduce the number of primer sequences (Table 2-3). A 650-bp DNA fragment was amplified through three rounds of nested degenerate priming PCR (Figure 3-27). The primer degeneracy doesn't influence the annealing temperature according to our observation, but decreases the number of matching primers.

Strategic variations have been developed all the time (Cheung and Nelson 1996). Novel PCR methods established in eukaryotic systems have been commonly adopted by microbiologists after certain modifications, e.g. random arbitrary primed PCR (Welsh *et al.* 1992) for fingerprinting of the eukaryotic mRNA had been applied to study bacterial gene expression (Shepard and Gilmore 1999). A new strategy, SD-PCR was developed to amplify the 5'-terminal region of the *afp* gene. In eukaryotic gene cloning, a protocol called rapid amplification of cDNA ends (RACE) had been developed to facilitate the cloning of the ends of a cDNA using information obtained from an incomplete cDNA clone (Frohman *et al.* 1988). This strategy takes the advantage of the polyA tail at the end of eukaryotic mRNA, the 3'-end of cDNA was amplified just by PCR with a gene-specific primer in a known region and a polyT tailed primer (3'-RACE). The 5'-end of the cDNA was duplicated by reverse transcription. When the 5'-end of mRNA was reached, a stretch of adenosine was added by terminal transferase (Frohman *et al.* 1988). The full sequence could then be amplified by a complementary polyT tailed primer (5'-RACE). There is an advanced approach of 5'-RACE, called CapSelect, in which people take the advantage of the CAP structure (7-methylG-5'-ppp-5'-G), specifically at the 5'-end of eukaryotic mRNA. The reverse transcriptase had been found to be able to add 3 or 4 dCMP residue, when reverstranscription reaches the CAP structure. A polyG tailed primer was used to anneal at this sequence in order to amplify the subsequent region (Schmidt and Mueller 1999).

Prokaryotic mRNA contains neither a polyA tail nor CAP structure, but a specific ribosome binding site called Shine-Dalgarno sequence (SD) locating usually 7 to 9 nucleotides upstream of the initiator "AUG" of a structure gene (Gold *et al.* 1981). Furthermore, there is no posttranscriptional recombination happening on prokaryotic mRNA, the relative position of SD sequence to the downstream gene on mRNA is the same as on genomic DNA. With a primer using same sequence as SD sequence coupling with a gene-specific reverse prime, the 5'-end of the interested gene can be amplified directly from genomic DNA template. Therefore I designed a hexameric primer, which has the same sequence as SD sequence at its 3'-end to clone the 5'-end of *afp* gene. The 5'-end of this primer is highly degenerate (Table 2-4). This primer is assumed to anneal on the corresponding sequence on the template chain of genomic DNA.

Due to its functional importance, the SD sequence is highly conserved in prokaryotes. The alignment of 120 translation initiation sites of bacterial genes showed most of them contain the typical “GGA” or “AGG” sequence in positions from 0 to –25 (Sazuka and Ohara 1996). This method could be used in studies where the *N*-terminus of protein is missing due to either artificial truncation or posttranslational modification, and where the protein sequence is known, but the upstream regulator sequence is of interest. Furthermore, SD sequence does not exist on eukaryotic genes. In consequence this method could be very useful for cloning prokaryotic genes from a DNA sample contaminated with eukaryotic DNA or cDNA.

4.7.3 Cloning the entire *afp* gene

The *afp* gene was located in a 2.3 kb fragment of genomic DNA from *M. arborescens* SE14 digested by restriction enzymes. It was adjacent to a putative acyltransferase gene. The primitive transcript contains 161 residues. The first methionine residue is removed from the primary sequence. Therefore the mature protein is composed of 160 amino acids, 48% of which are hydrophobic. The M_r and pI values calculated from the gene sequence are 17.136 kD and 4.32, respectively, coincident with the results from the native protein. According to the sequence data, the protein might be folded in a way that some parts of the protein are highly hydrophilic and others are hydrophobic. It carries 12 negative charges at PH 7.0. The polar surface would be exposed to the medium, while the hydrophobic surface would be involved in the intersubunits interaction as well as the protein central cavity formation (see 4.7.4.2).

4.7.4 Analysis of *afp* gene

The protein translated from the *afp* gene is a member of a Dps-like protein superfamily. All peptide sequences generated from protein sequencing have been located in the full protein (underlined in Figure 3-30). Although Afp shares the critical residues of a ferroxidase centre (asterisks in Figure 3-30) with other Dps family proteins, it is overall less conserved, in particular with respect to the unique *N*-terminus region.

4.7.4.1 DNA binding activity

The Dps protein family is a group of proteins from distantly related bacteria (Table 4-2). The name is somewhat misleading because only a small part of them are able to bind DNA by unspecific interaction of a positively charged *N*-terminus or *C*-terminus with the negatively charged DNA backbone (Ceci *et al.* 2003); a striking number of homologs can't form complexes with DNA (Evans Jr. *et al.* 1995; Bozzi *et al.* 1997; Tonello *et al.* 1999; Pulliainen *et al.* 2003). Our data confirmed that DNA-binding activity of Dps from *E. coli* does not require its secondary structure, as the denatured Dps protein still retarded DNA from entering the agarose gel (see 3.4.4). In contrast, neither native nor denatured Afp protein formed a complex with DNA, which is coincident with its lack of positively charged termini.

4.7.4.2 Stable structure

All of the known Dps-like proteins are composed of small subunits (Table 4-2). They usually form very large aggregates. The native structures of most Dps members comprise 12 subunits arranged in two hexameric rings with a central hollow cavity (Figure 4-3 A). Two ferrous ions bind to each of the 12 dinuclear ferroxidase sites at the dimer interface inside the dodecamer (Grant *et al.* 1998; Ilari *et al.* 2000; Papinutto *et al.* 2002; Ren *et al.* 2003). The Afp protein from *M. arborescens* is a dodecamer of 12 identical subunits. The presence of intersubunit ferroxidase centres will be revealed with X-ray crystallography in the near future. The apparent Mr of Afp on size exclusion column is much higher than actual (see 3.4.1), probably due to the cage-like structure that is larger than a compact structure of a protein of same mass.

Table 4-2. Some reported Dps-like proteins from eubacteria.

Name	Species	Location	MW (kd)	DNA binding	Iron binding	Other suggested function	Reference
Dps	<i>Escherichia coli</i>	Cytoplasm	19	Yes	Yes	Expression regulation	Almirón <i>et al.</i> 1992
Dps	<i>Campylobacter jejuni</i>	Cytoplasm	17		Yes		Ishikawa T. <i>et al.</i> 2003
AtDps	<i>Agrobacterium tumefaciens</i>	Cytoplasm	18.6		Yes		Ceci P. <i>et al.</i> 2003
TpF1 (4D)	<i>Treponema pallidum</i>	Cell surface	19			Antigen	Noordhoek <i>et al.</i> 1989
NapA	<i>Helicobacter pylori</i>	Cell surface	17	No		Neutrophil activation Mucin binding	Namavar <i>et al.</i> 1998
FtpA	<i>Haemophilus ducreyi</i>	Pili	24			Laminin binding	Brentjens <i>et al.</i> 1996
MrgA	<i>Bacillus subtilis</i>	Cytoplasm	16	Yes	Yes		Chen L. & Helmann. 1995
Flp	<i>Listeria innocua</i>	Cytoplasm	18		Yes		Bozzi M. <i>et al.</i> 1997
Dpr	<i>Streptococcus mutans</i>	Cytoplasm	19.7		Yes		Yamamoto Y. <i>et al.</i> 2000
Dlp-1	<i>Bacillus anthracis</i>	Cytoplasm	16.9	No	Yes		Papinutto E. <i>et al.</i> 2001
Dlpi-2	<i>Bacillus anthracis</i>	Cytoplasm	16.7	No	Yes		Papinutto E. <i>et al.</i> 2001
Dpr	<i>Streptococcus suis</i>	Cell surface		No	Yes	Galactose adhesion	Pulliainen A. T. <i>et al.</i> 2002
BbDps	<i>Bacillus brevis</i>	Cytoplasm	16.3	Yes	Yes		Ren B. <i>et al.</i> 2003
Ms-Dps	<i>Mycobacterium smegmatis</i>	Cytoplasm	21.6	Yes	Yes		Gupta S. & Chatterji D. 2003

Blank in this table means no report

The Dps-like four-helix bundle proteins (Figure 4-3 B) are highly resistant to denaturation (Tonello *et al.* 1999). The neutrophil-activating protein (HP-NAP) from *Helicobacter pylori* is resistant to temperatures up to 80°C. As much as 4 M guanidinium chloride is required to denature it. We observed that the Afp protein is also highly stable. The enzymatic activity of Afp did not change in the presence of SDS (see 3.4.3) and 10 % ethanol (3.6.2), which is destructive to most of the known enzymes. It was also resistant to denaturation by heating at a temperature as high as 48°C.

4.7.4.3 Different aggregation states

Besides the dodecameric structure mentioned above, the fine tangled pili protein (Ftp) from *Heamophilus ducreyi* exists only as a hexamer. It forms thin filaments of 3 nm in diameter (Brentjens *et al.* 1996). Trimeric structures have also been observed in Ms-Dps from *Mycobacterium smegmatis*. But interestingly, the protein lost its DNA-binding activity when its structure changed from dodecamer to trimer (Gupta and Chatterji 2003). When some of these proteins were measured *in vitro*, hexameric aggregates had frequently been detected on native gel (Peña and Bullerjahn 1995; Brentjens *et al.* 1996; Hébraud and Guzzo 2000). According to our results, the dodecameric aggregate of Afp in solution separates into two hexamers on native PAGE (see 3.4.1).

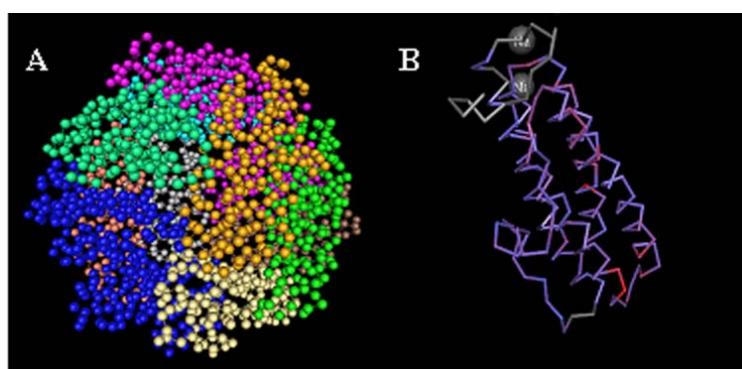


FIGURE 4-3. The dodecameric aggregate of Afp (A) and the four-helix bundle structure of Afp protomer (B) predicted by the software Cn3D 4.1 (NCBI) according to the sequence similarity of Dps-like proteins.

Furthermore, even smaller oligomers of Afp were revealed after denaturing on SDS-PAGE. Some of the oligomers could result from denatured monomers which were crosslinked *in vivo* or during protein preparation. Such high-Mr oligomers had also been detected when Ftp were separated on SDS-PAGE page (Brentjens *et al.* 1996), and when *E. coli* Dps were sedimented by ultracentrifugation in solution (Zhao *et al.* 2002). They probably arise from iron-mediated radical chemistry. Such intersubunit crosslinking is common for the ferritins, which contain similar ferroxidase sites (Zhao *et al.* 2002). The iron mediated radicals (see below 4.7.4.4) could randomly link adjacent peptide chains of separate subunits together (Yang *et al.* 2000).

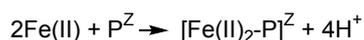
Dps-like proteins are abundant when starvation occurs. It had been estimated that a single cell of stationary-phase *E. coli* may contain *ca.* 20,000 Dps monomers (Almirón *et al.* 1992). This large amount of proteins had been observed to be sequestered into intracellular crystals. This so-called biocrystallization was supposed to protect DNA from oxidative damage (Wolf *et al.* 1999). It might be that like the dodecameric structure, the ability of further aggregation of the dodecamers is just an evolution relic. We had observed random aggregation of the Afp dodecamers when the native protein was measured by Quattro-ESI (2.4.1). Such a process needs not only the symmetric structure, but also a properly charged surface. The Afp protein is also very abundant under starved conditions and well charged under neutral pH as predicted by DNASTar. This might contribute to hydrogen-bond interaction inter and intra the subunits.

4.7.4.4 Iron-binding activity

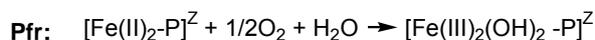
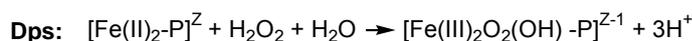
The so-called DNA-binding signature, from position 43 to position 82 of the Afp sequence (Figure 3-30), is conserved in all the Dps-like proteins. It is misnamed because these residues are not involved in the reaction with DNA, but in the formation of the ferroxidase centre (Bozzi *et al.* 1997; Ceci *et al.* 2003). This ferroxidase centre is conserved in all Dps family proteins so far investigated. It is known that Dps-like proteins oxidize iron through a mechanism similar to ferritin (Yang *et al.* 2000) although most of them do not contain heme at the active center (Zhao *et al.* 2002).

Bacterial ferritin and Dps oxidize iron to form a mineral core of hydrated ferric oxide within their internal cavities. In some species, such as *E. coli* (Ilari *et al.* 2002) and *H. pylori* (Dundon *et al.* 2001), both Dps-like protein and ferritin (Pfr) have been well investigated. The dinuclear ferroxidase sites in typical 24-mer ferritins locates in the four-helix bundle of individual monomers, while in 12-mer Dps-like proteins ligands are provided by two symmetry-related subunits. The mineralization of both kinds of proteins was proposed to have three phases in common:

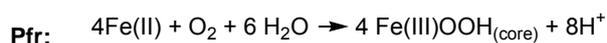
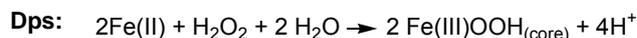
Binding:



Oxidation:



Mineralization:



The first steps in both reactions are the binding of iron to the ferroxidase center. In the following reactions, unlike ferritin, which oxidizes iron with molecular oxygen, Dps can use hydroxyl radicals for iron oxidation. Therefore its physiological function was supposed to remove hydrogen peroxide in addition to form the ferric oxide mineral core.

The ferrous ion salt bridge in the ferroxidase centre is not necessary for dodecameric aggregation. When the two Dps-like proteins from *Bacillus anthracis* were expressed in *E. coli*, no iron was detected (Papinutto *et al.* 2002). Similarly, when the HP-NAP of *H. pylori* was expressed in *B. subtilis*, the iron content was only 0.2-0.3 atoms per monomer (Tonello *et al.* 1999). The dodecameric structure of these proteins indicated that iron is not involved in the dodecamer formation. We had found that the iron binding activity of Afp depends on secondary structure. The native Afp proteins in hexameric and dodecameric structures were able to bind iron, whereas the denatured protein in SDS-PAGE lost this ability. This suggests that the oligomer on SDS-PAGE pages was probably denatured protein.

Some Dps-like proteins have been found to be capable to store up to 500 Fe atoms in the central cavity of a single protein *in vitro* under ferrous ion overloading conditions (Bozzi *et al.* 1997; Tonello *et al.* 1999; Ceci *et al.* 2003). Only 24 (Zhao *et al.* 2002; Ceci *et al.* 2003) and 12 (Pulliainen *et al.* 2003) iron atoms, or even less, (Bozzi *et al.* 1997) have been detected in the native protein. The Afp protein may contain as many as 120

iron atoms in a native dodecamer. This is the highest value among the tested Dps family proteins. When the *afp* gene was expressed in *E. coli*, no catalytic activity has been detected. This might be the consequence of the improperly incorporation of iron in the overexpressed protein. The localization and function of this large amount of iron in Afp is an interesting point for future investigations.

4.7.4.5 An unknown catalytic mechanism

X-ray crystallography of Dps-like protein revealed the existence of hydrophobic pores in its shell-like structure, but the function is unknown (Ren *et al.* 2003). The Afp protein is composed of 48% hydrophobic amino acids, and catalyzes a reaction, in which the substrate selectivity clearly depends on hydrophobicity (see 3.6). It neither can use acyl-CoA as a substrate nor need ATP to drive its reaction. It is clear that the Afp catalyses *N*-acyl amino acids formation through reversal of a hydrolytic reaction with a K_{eq} of 0.1863 (see 2.5.1). The Afp protein very likely uses the same kind of mechanism to produce *N*-acyl amino acids as the protein extract found in homogenates of *M. sexta* digestion tract (Lait *et al.* 2003). Both of them catalyze reversible condensation of free fatty and amino acids. The pH optima of the catalytic activity are coincident with the pH values of insect crop and midgut (see 3.5.3) from where the content could be regurgitated by the larvae.

The conserved ferridoxase centre is apparently not essential for this biotransformation. When the Dps protein isolated from *E. coli* was supplied with free linolenic acid and glutamine, no conjugation had been detected, even if a very large amount of Dps protein purified from the Dps overexpressed strains was used. To reveal the involvement of other groups in this reaction, successful heterologous expression with intact and modified gene will be necessary.

4.7.4.6 Evolution from a distinct line

The homology between Dps proteins is most likely due to a common ancestry rather than common function (Evans Jr. *et al.* 1995). Besides the oxygen radicals scavenging and iron mineralization mediated by the ferroxidase centre, many of them have been reported to be involved in carbohydrate binding for adhesion (Table 4-2), as well as in pili formation (Brentjens *et al.* 1996), and cold shock tolerance (Hébraud and Guzzo 2000). Despite of sharing the few highly conserved amino acid residues, even the putative ferroxidase centres vary considerably and might lead to different functional properties (Pulliainen *et al.* 2003). Afp was isolated from the gram-positive *M. arborescens* strain, but showed a little higher homology to the Dsp-like proteins isolated from gram-negative bacteria. The calculation showed the largest evolutionary distance among Dps family proteins (Figure 3-31).

Protein phylogenetic parsimony analysis revealed that the Dps were clustered in a branch separated prior to the separation of anaerobes from aerobic species (Rocha *et al.* 2000). Although Afp was isolated from a Gram-positive bacterium, the alignment showed low homology to other Dps-like proteins and it could not be grouped into either Gram-positive or Gram-negative subgroups. It might have been evolved from a different line from other Dps family members and gained a new function.

Furthermore, Dps-like proteins share a common evolutionary origin with bacterioferritins and eukaryotic ferritins (Peña and Bullerjahn 1995). A heme-containing

dodecameric Dps-like protein had been isolated from a cyanobacterium *Synechococcus* sp. (Peña and Bullerjahn 1995). A more fascinating Dps homolog was isolated from the archaeon *Halobacterium salinarum* (Reindel *et al.* 2002). Afp showed some homology to bacterioferritin, as well as insect ferritins. This also contributes to their common iron-binding activities. It is possible that in some other living organisms proteins using a similar catalytic mechanism as Afp might exist.

4.8 Expression of the *afp* gene

Because the capability of amide hydrolysis had never been detected in Dps-like proteins, the *afp* gene was expressed in *E. coli* and in the *M. arborescens* DSM20754 strain to finally proof its function. The successful gene expression using *Microbacterium* spp. as host provides a new useful tool for cloning and expression of streptotomycetes gene.

4.8.1 Overexpression in *E. coli*

Escherichia coli cells are most commonly used host for production of recombinant proteins, since they are easy to manipulate and well characterized (Nakashima and Tamura 2004). The *afp* gene was first expressed in *E. coli*. This translated protein formed hexameric aggregates on native electrophoresis gel like the native *M. arborescens* SE14 protein did. The protein expression was demonstrated by antibody detection (Figure 3-32). But surprisingly, the *E. coli* overexpressed Afp protein was not able to produce *N*-acyl amino acids. It is known that in *E. coli* some proteins are difficult to be expressed and purified due to inhibitory effects on the growth of the host cells, or for other unknown reasons (Nakashima and Tamura 2004). Since the extremely low activity was not a reliable functional proof, a new expression strategy was required.

The protein expressed in *E. coli* is different from the protein expressed in *M. arborescens* SE14 since there is only one type of protomer in the *E. coli* fraction. The Mr is exactly same as that deduced from the cloned gene. However, two kinds of subunits were detected in the *M. arborescens* SE14 fraction. The larger one is dominant (see 3.4.1). Further more, the larger protomer is 43 Da larger than the smaller one, suggesting a modification. According to the difference on Mrs, this modification could be an acetylation. It is known that some proteins require posttranslational modification for activity, Mycobacteria had been reported to acetylate proteins (Connell 2001). Probably the active form of Afp was either composed by only modified subunits, or by a mixture of the two kinds of subunits. In any case, this posttranslational modification seems to be required for the protein's catalytic activity. In fact, there is a putative acyltransferase gene in adjacent to *afp* gene in *M. arborescens* SE14 genome. They share a putative promoter region (see 3.7.3).

Probably, *E. coli* lacks the molecular apparatus for this posttranslational modification. On the other hand, the *N*-terminal methionine of Afp was removed in mature protein as determined by Edman degradation (see 3.7.2.2). The Ms data suggested that both *E. coli* and *M. arborescens* DSM20754 could recognize the signal peptide in this protein, and perform this cleavage on the protein *N*-terminus.

4.8.2 Expression in *M. arborescens* DSM20754

Microbacterium spp. belong to the high GC actinomycetes, which includes many commercially and clinically important species, such as Streptomyces, Mycobacteria, and Corynebacteria. They comprise not only causative agents of tuberculosis, leprosy, and diphtheria *etc.*, but also many important pharmaceutical producers. More than 7000 polyketides have been identified up to date (Xue *et al.* 1999), more than two-thirds of them are produced by actinomycetes (Hu *et al.* 2000). Furthermore, *Corynebacterium* spp. and *Brevibacterium* spp. had also been employed for industrial production of amino acids since 1957 (Krämer 1994). Over the last decades, explosive advances in molecular biology had improved the investigation and manipulation of critical genes to meet scientific and industrial requirement.

Accordingly, a robust system for the heterologous expression of actinomycetes genes is in great demand. *Escherichia coli* is the most widely used host, but results are often unsatisfactory due to the taxonomical distance (Pfeifer and Khosla 2001). To reconstitute the expression of these high GC content genes in genetically and genomically friendly heterologous host, many efforts had been invested on looking for an appropriate host in the genus of Streptomyces. *Streptomyces* spp. are the most studied and developed hosts in current use. The complete genome of *S. coelicolor* A3(2) even has been sequenced (Bentley *et al.* 2002).

Although *Streptomyces* spp. are the most widely investigated actinomycetes gene expression hosts (Bentley *et al.* 2002), there are many unsolved problems, for example proteolytic degradation of expressed protein (Tremblay *et al.* 2002), deficiency of biosynthetic substrate (Pfeifer and Khosla 2001), and so on. Moreover, Streptomyces grow as multicellular, multinucleoid, branching hyphae, which form a mycelial mass. They have a very complex life cycle, produce spores by differentiation of the aerial hyphae into uninucleoid compartments (Kieser *et al.* 2000). These characters make the cultivation and transformation complicated and time-consuming. It requires many different mediums for the different growth stages as well as sporulation and transformation (Kieser *et al.* 2000).

Therefore, looking for easy-handling gene cloning and expression host has never stopped. Some other successful cases of gene expression in this high GC bacterial group have been reported for the genera *Mycobacterium* (Connell 2001), *Corynebacterium* (Kirchner and Tauch 2003), *Brevibacterium* (Adham *et al.* 2001), *Saccharopolyspora* (Wilkinson *et al.* 2002), *Micromonospora* (Li *et al.* 2003), and *Rhodococcus* (Nakashima and Tamura 2004), but never for *Microbacterium*.

4.8.2.1 The *afp* gene ortholog in *M. arborescens* DSM20754

Because the activity of the protein expressed in *E. coli* was not satisfactory, the *afp* gene was expressed in *M. arborescens* DSM20754, which is a low activity strain (see 3.2). Surprisingly, it does contain an orthologous gene of *afp*. The ortholog can be easily discriminated from *afp* by using different PCR conditions due to their difference on the DNA sequence. The *afp* orthologous gene, Ma-DSM-*afp* is different from *afp* on DNA and protein sequences. We still do not know whether or not the different residues in Afp and the ortholog are responsible for the low activity detected in *M. arborescens*

DSM20754. There could be no difference between the catalytic activity of the two orthologs, but different regulatory factors enhancing the *afp* expression in *M. arborescens* SE14. Anyhow, the absence of identical *afp* in *M. arborescens* DSM20754 makes the expression and detection of *afp* in this host easier.

4.8.2.2 Construction of expression plasmid

The entire *afp* gene and the following T7 terminator have been engineered downstream of a *B. subtilis* *veg* promoter. The *afp* is a small gene, it encodes a small protein of only 160 amino acids. The *B. subtilis* *veg* promoter initiates transcription of an 86-amino acid polypeptide gene (Moran *et al.* 1982) during all growth phases (Fukushima *et al.* 2003). This promoter can be recognized by a wide range of heterologous host such as *E. coli* (Fukushima *et al.* 2003), *Streptococcus pyogenes* (Opdyke *et al.* 2003), and *Streptomyces coelicolor* (Kang *et al.* 1997). In *Streptomyces* spp. the *veg* promoter was recognized by σ^{66} , the major vegetative sigma factor (Kang *et al.* 1997). The successful expression of *afp* in *M. arborescens*, as described in this work, confirmed the conservation of this kind of sigma factor in distantly related species. Furthermore, the high activity in pRMM transformants confirmed that the *afp* gene was transcribed by a σ^{66} -like sigma factor different from that transcribed the same gene in *M. arborescens* SE14, which seems to be controlled by cell density.

Integrative vectors are widely used in actinomycetes expression systems, sustainable plasmid had only been reported occasionally on *Streptomyces* (Kieser *et al.* 2000), *Corynebacterium* (Kirchner and Tauch 2003), *Rhodococcus* (Nakashima and Tamura 2004), and *Mycobacterium* (Connell 2001). In this work, the truncated replicon of a low copy number *Streptomyces* plasmid, SCP2* on pRM5 (McDaniel *et al.* 1993) was employed to facilitate the duplication of the alien plasmids in *M. arborescens* DSM20754. But this plasmid origin is isolated from streptomyses, not directly from *Microbacterium* spp., plasmid instability had been observed frequently.

4.8.2.3 Transformation of *M. arborescens* DSM20754

To bypass the obstacle of methyl-specific restriction systems which are common in Streptomycetes (Kieser *et al.* 2000), the plasmids must be demethylated prior to transformation. Methylated plasmids can not be introduced into the *M. arborescens* DSM20754 protoplast and may suggest the existence of the methyl-specific restriction system in *Microbacterium* spp. Although the transformation is always successful, the transformation efficiency is not very high. Similar to most of the streptomycetes transformation cases, only three or four colonies developed on a selective plate surface (Figure 3-35). Regeneration of Streptomycetes protoplasts was also problematic and its efficiency varied from one strain to another (Kirchner and Tauch 2003). Improvement of the transformation process is required.

Plasmid instability is always the problem of expression gene in Streptomycetes. When the gene cluster from lactic acid bacteria were introduced into *Corynebacterium glutamicum*, the modified corynebacteria exhibited also plasmid instability (Barrett *et al.* 2004). The plasmid instability was confirmed by plasmid recovery from these transformants. The plasmids had been rearranged in all of the transformants. But because the *afp* gene is very small in comparison to the whole plasmid, the chance of truncation in the expression cassette is very low.

4.8.2.4 Increase of *N*-acyltransferase activity

The *afp* gene transformed cells of *M. arborescens* DSM20754 produce more *N*-acyl amino acids than the wild type. The enzyme activity of the strains transformed with the control plasmid stayed at the same level as the wild type *M. arborescens* DSM20754. On the other hand, the enzyme activity of strains carrying the *afp* gene maintained a relatively high level throughout the measurement. A small increase has even been observed at later growth stage (Figure 3-36). It was known, that with different vectors, expression levels varied widely in different actinomycetes strains, indicating that different host factors are needed for optimal expression (Wilkinson *et al.* 2002). In this experiment, although the enzyme activity had been increased 2.5 times after the introduction of expression plasmid, it was still quite low when compared to *M. arborescens* SE14. More effort is required to optimize the gene expression in this bacterium.

Interestingly, the *afp* gene was found not only to increase linolenoyl glutamine production, but also to promote cell aggregation. Furthermore, transformation with the *afp* gene made the bacterial cell wall more fragile than that of wild type. Whether these changes are caused by the *N*-acyl amino acids or by the Afp protein itself needs further investigation.

Finally, our results had demonstrated that the *afp* gene increased the *N*-acyl amino acid production in *M. arborescens* DSM20754, and provided a conceptual framework for using this species as a GC-rich gene expression host. By comparison to streptomyces, Microbacteria have several advantages: They are unicellular bacteria. In young cultures, the cell is an irregular short slender rod. Primary branching is uncommon. In older cultures the cells become shorter, a portion may be coccoid (Collins and Bradbury 1991). The preparation of protoplasts is easy and efficient. The growth rate of *M. arborescens* DSM20754 is as fast as that of *E. coli* after transformation. It is a saprophyte, and grows well on simple BHI medium. It can also grow on LB medium, but with a slower growth rate. These advantages should enable the development of an amendable host for production and analysis of actinomycetes gene. For this purpose further research on the cell structure, metabolism, and genome is required.

4.9 The ecological role of Afp

Afp has been demonstrated to be homologous to Dps-like proteins, which are highly expressed in starved cells (Almirón *et al.* 1992; Gupta *et al.* 2002). The *N*-acyl amino acid-forming activity of *M. arborescens* SE14 also dramatically increased at the onset of the stationary-phase growth when nutrient depletion occurs. Such stringent conditions appear to be similar to bacteria colonizing higher organisms with respect to gene expression and physiology (Parsek and Greenberg 2000; Gupta *et al.* 2002). When the bacteria lives in insect gut, they must compete intensely with the insect digestion system for available nutrient. Hence, the precisely controlled time window of this bio-transformation suggests a defined interaction between the bacteria and their insect hosts.

The amide forming enzymatic activities in both *M. arborescens* SE14 and *M. arborescens* DSM20754 were highest when the bacterial cell density reached an OD₆₀₀ of 2.0, which seems a critical checkpoint controlling the *afp* gene expression (see 3.2 and 3.3). Likewise, the tested gut bacterial strains increased their *N*-acyl amino acids forming

activity all at the early stationary phase (see 3.1). These data suggests that the expression of the genes corresponding to amide formation in gut bacteria, *esp.* the *afp* gene from *M. arborescens* may be controlled by Quorum Sensing. This system enables a given species to sense a critical population density. Quorum Sensing is involved in successful pathogenic or symbiotic interactions of a variety of Gram-negative bacteria (Parsek and Greenberg 2000) and Gram-positive bacteria (Dunny and Leonard 1997) with plant and animal hosts. In gram-positive bacteria the signaling molecules are peptides or modified peptides (Dunny and Leonard 1997) instead of Acyl-homoserine lactone, which only works in gram-negative bacteria (Parsek and Greenberg 2000).

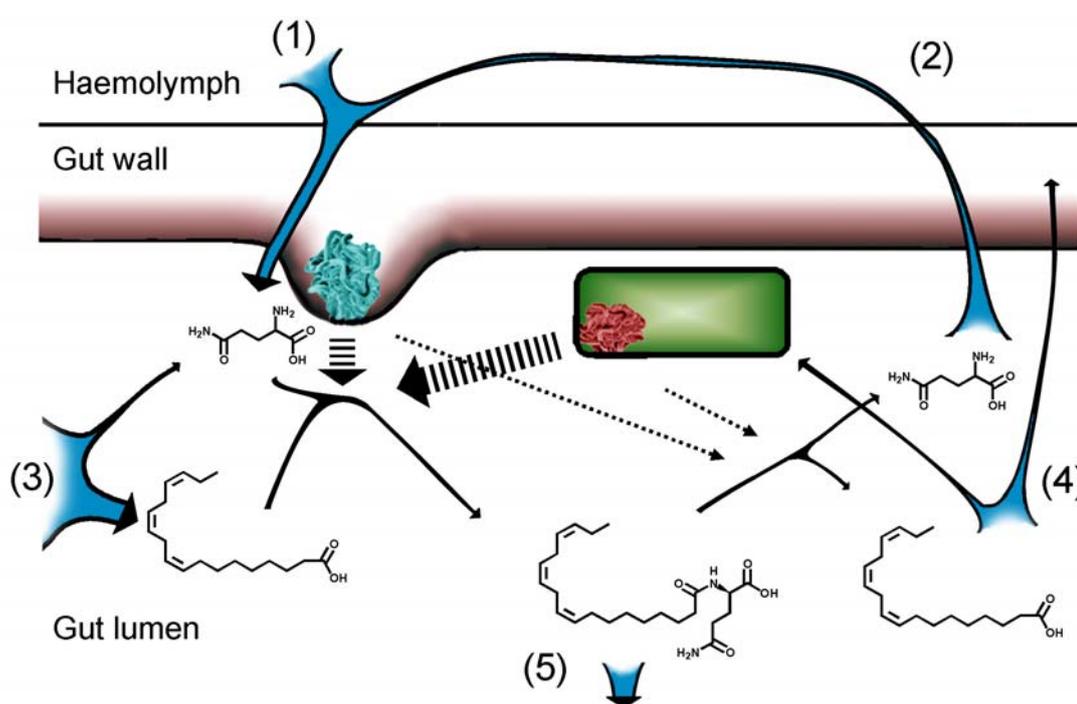


FIGURE 4-4. A proposed mechanism of biosynthesis of *N*-linolenoyl glutamine in insect gut. The glutamine could be from three sources: insect biosynthesis (1), recycled from hydrolysed *N*-linolenoyl glutamine (2), and insect diet (3). The fatty acid linolenic acid is from insect diet (3). The fatty acid could be used as nutrient after hydrolysis of the conjugates (4). The insect protein located in the epithelial cells of the gut membrane is supposed to catalyses synthesis. The bacteria protein might catalyses hydrolysis. At meantime, both could catalyse reverse reactions. Their relative importance in the synthesis or hydrolysis of conjugate was shown as dashed arrows. The *N*-linolenoyl glutamine produced by the cooperation of the two kinds of enzymes is regurgitated by the insect during feeding (5), and induces plant defence reaction.

One of the suggested functions of *N*-acyl amino acids is to defend against harmful bacteria (Alborn *et al.* 2000). Another is to solubilize nutrients in the insect food for either the insect or gut biota (Spiteller *et al.* 2002). According our observation, at least *N*-acyl amino acids are important for the interaction between gut bacteria and their insect

host, because so many gut bacterial isolates are able to produce this kind of compounds (see 3.1.1).

Interestingly, insect could also produce these compounds. There is an active protein fraction isolated from the insect *M. sexta* digestion tract (Lait *et al.* 2003). This insect protein shares many common catalytic properties and is present simultaneously at the same place with the bacterial Afp. According to its catalytic property, Afp might contribute mostly to hydrolysis of *N*-acyl amino acids. At high local concentration of amino acid and fatty acids, Afp could also catalyse to a certain level of biosynthesis.. Although it is still not clear which one responsible for hydrolysis, and which one responsible for synthesis, their cooperation would maintain a stable pool of *N*-acyl amino acids in insect gut (Figure 4-4).

The dynamic equilibrium of this pool is maintained by a complex cooperation. The fatty acids are taken from the food plant, while most of the amino acid glutamine is derived from the insect (Paré *et al.* 1998). We assume some of the glutamine could be plant derived because this amino acid is the most abundant amino acid in plant tissue (Lamber *et al.* 1998). The amino acid could also be recycled in the insect, because when labeled amino acid was injected into insect haemolymph it had been detected in the secreted conjugates (Spiteller *et al.* 2000). Not only Afp can catalyse the hydrolysis of *N*-acyl amino acids, but a protein fraction from midgut lumen of insects (Mori *et al.* 2001). The hydrolysed products could be reabsorbed by the insect gut for either reincorporation or use as nutrients.

The *N*-acyl amino acids are regurgitated from this dynamic pool by the insect during feeding, and may subsequently act in the damaged plant as a signal of coming attack. A volicitin-binding activity has been detected in the enriched plasma membrane fraction in *Z. mays* (Truitt *et al.* 2004). But how and when plant evolves this receptor system deserves further studies. In the mammalian central nervous system the receptor of the fatty acid amide hormone cannabinoid has been characterized (Cravatt *et al.* 2004). Another structurally related hormone Oleamide uses totally different receptor system at the same place (Lees and Dougalis 2004). These mammalian neural transmitters have similar amphiphilic structures to *N*-acyl amino acids (see 1.1.3). Whether these systems share similarities with the plant plasma membrane fraction is an interesting question. If so, probably there is a signaling system *in planta* using messenger(s), which is structurally related to *N*-acyl amino acids. This could finally explain why both insect and its gut biota evolve systems to control the level of *N*-acyl amino acids. Successful interference with the plant signaling pathways would benefit the herbivore, as well as the gut biota, which share the food with the insects.

4.10 Perspectives

The ability to hydrolyse or synthesize *N*-acyl amino acids is a common property of many phylogenetically unrelated gut bacteria. This suggests that *N*-acyl amino acids play a key role in the interaction between the bacteria and the insect host. Furthermore, the catalytic activity of the reference strain *M. arborescens* DSM20754 is much lower than that of the strain isolated from insect gut. This could be either due to an enhanced ability of *M. arborescens* SE14 during colonization of the gut, or a decreased ability of *M. arborescens* DSM20754 during cultivation under laboratory conditions. In any case, the difference between the two strains supports the assumption that the insect gut may exert a selective stress to increase the bacterial ability to affect *N*-acyl amino acids.

The pH and temperature optima of Afp are coincident with the environment of insect's digestion tract. The catalytic properties of purified Afp share many features with a protein fraction isolated from the homogenate of alimentary tissue of the tobacco hornworm (Lait *et al.* 2003). It is surprising that enzymes of very distantly related organisms catalyse the same type of conjugation of free fatty acids and amino acids without the need for activation and in the same broad pH range (Lait *et al.* 2003). The identification of the active enzyme(s) from the insect is therefore essential to answer the question, which enzymes contribute to which extent to the biosynthesis/hydrolysis of the *N*-acyl glutamines.

The synthetic activity of Afp utilizes almost all proteingenuous amino acids as substrates except proline. On the other hand, only glutamine, and occasionally glutamate were found in insect regurgitants. If Afp is the enzyme responsible for their biosynthesis, a strictly controlled substrate supply and product removal are required. The selectivity of Afp for fatty acids is also broad. Almost all fatty acids present in the food plant seem to be used, coincident with the *N*-acyl amino acid profiles observed in the regurgitant of the caterpillar (see 1.1.1) and the profile of lip amino acid in bacteria (see 1.2.1.1).

Afp is homologous to the bacterial Dps proteins, which are also known as scavengers of reactive oxygen species (ROS). One could speculate on whether or not Afp may have a dual function. Not only could it catalyse the hydrolysis/synthesis of *N*-acyl amino acids, but also protect the gut bacteria by scavenging the ROS from the plant food (Ceci *et al.* 2003). Superoxide dismutase assays (Flohe and Oetting 1984) would be helpful to detect the a potential free radical scavenging activity of Afp. The Afp protein has the highest iron content among tested Dps family proteins. Elucidation of its 3D structure by X-ray diffraction of the protein crystal will finally reveal the function and localization of the iron atoms as well as the ferroxidase center. The purified protein from *M. arborescens* SE14 has been already successfully crystallized.

Biosurfactants have been proposed to be involved in biofilm formation (Ron and Rosenberg 2001). We observed the formation of similar clag structures by *M. arborescens* SE14, and *M. arborescens* DSM20754 after transformation with the *afp* gene. After successful identification of the gene, the *afp* gene in the original host, *M. arborescens* SE14, should be silenced by gene disruption to investigate its role in bacterial life and biofilm formation. Colonization of germ-free insects with the genetically modified bacteria would be an especially promising approach.

Dsp-like proteins are widely distributed in prokaryotes. Degenerate primers should be designed from the conserved region of Dsp-like proteins and used in PCR approaches to screen other gut bacteria and to, finally, construct the phylogeny of this interesting and widespread enzyme. This approach is also applicable for the amplification of homologous genes from total DNA extract of the insect gut. This result will reveal homologous genes from uncultivable bacterial symbionts. With the SD-PCR method developed in this study (see 3.7.2.2 and 4.7.2) the 5'-end of the interesting DNA fragments can be amplified directly from gut DNA extract. Therefore the homologous genes from uncultivable bacterial symbionts should be likewise available.

The localization of the Afp protein in the bacteria and that of the bacteria in insect alimentary tissue will be approached by an Afp-protein fused with the green fluorescent protein (GFP). The fusion gene will be reintroduced into *M. arborescens* to replace the prototype gene by homologous recombination. This would create a transgenic microorganism that will be used to "monitor" environmental conditions turning on the hydrolysis/biosynthesis reaction. Such a "probe" could be used potentially to understand why some insects (e.g. *Pieris rapae*) produce distinct conjugate profiles.

Discussion

The insect gut is a favourite place to facilitate horizontal gene transfer (Dillon and Dillon 2004). To evaluate the role of horizontal gene transfer in the wide distribution of this biotransformation, the gene transfer efficiency in insect gut will be measured by using a genetically marked strain of *Bacillus subtilis* strain, in which an antibiotic resistance marker or GFP will be permanently introduced into its genome by recombination. This bacterium could then be fed to insects, and reisolated after some time.

To understand the evolutionary origin of Afp cloning of the flanking region of the *afp* gene might also provide useful information. A cosmid library has already been successfully constructed. Some of the colonies have been sequenced.

Since most symbionts are not yet cultivable (Amann *et al.* 1995), the isolated strains represent only a small fraction of the total microorganisms of the insect's digestive tract. The complete bacterial population in caterpillar gut will be estimated by using *Fluorescence In Situ Hybridization* (FISH). Preliminary experiments have been done with larvae from *Heliothis virescens* and *Spodoptera littoralis*, and we know already that the bacterial gut community is highly variable and strongly depends on the nature of the food plant. This interesting dynamic allocation of microbial populations has to be analysed and functionally understood for generalist insects.

Acknowledgements

I am grateful to my major advisor, Professor Wilhelm Boland, for his guidance and support, for giving me the opportunity to study in this wonderful institute. His expertise in chemistry and ecology improved my research skills and prepared me for future challenges. My special appreciation also goes to the other members of my advisory committee, Professor Erika Kothe, for her helpful suggestions and comments during my study, and Prof. Dr. Jörn Piel for his friendship, encouragement, and numerous fruitful discussions. He has also contributed much to these experiments. He sequenced the bacterial 16S ribosomal DNAs of gut bacterial isolates. He also provided me some useful plasmids.

I feel privileged to have been able to work with Dr. Rita Büchler and Dr. Dieter Spiteller. Dr. Büchler has contributed much to and also given me help in the protein purification and determination. She had determined the protein pI by isoelectric focusing and its Mr by FFF-MALS. Dr. Spiteller did fundamental work for this research especially the isolation of the bacterial strains, which I worked on, and the development of the HPLC methods. I also thank Dr. Aleš Svatoš for his excellent technical help on different mass spectrometry measurements. Dr. Axel Mithöfer for his assistance in protein sequencing with mass spectrometry and fruitful discussions. I thank Dr. Nicolas Delaroque for providing me the cloning vectors and helpful technical discussion. I thank Miss Katrin Zimmermann, Dr. Dequan Hui, Dr. Aiyang Li, and all the people working in the molecular biology lab for assistance in experiments. I thank PD. Dr. Georg Pohnert, Miss Janine Rattke, and Mr. Sven Adolph for their assistance on HPLC measurements, and Dr. Antje Burse for her discussion.

I thank Prof. Konrad Dettner (University of Bayreuth, Bayreuth, Germany) for providing insects and discussion, and Dr. Michael Rässler (Max-Planck Institute for Biogeochemistry, Jena, Germany) for the iron determination, Dr. Bernhard Schlott (Institute for Molecular Biotechnology, Jena, Germany) for sequencing the protein with Edman degradation. I also thank Prof. Rudolf I. Amann (Max Planck Institute of Marine Microbiology, Bremen, Germany) for teaching me the FISH technique, and thank Prof. Roberto Kolter (Harvard Medical School, Boston, USA) for his kind gift of the *E. coli dps* gene.

I thank our secretary, Mrs. Grit Winnefeld, and all the group members of the Bioorganic Chemistry for their general support. I learnt not only a lot of new knowledge in past years, but also the way of being a scientist, in other words, how to observe, to deduce, to think, and to cooperate with colleagues. I will miss working in the Bioorganic Chemistry group and keep the memory and encourage around wherever I go in my life.

This work was supported by the Max-Planck-Gesellschaft, which deserves my unending gratitude.

Finally I am grateful to my parents and family for their love and thrust. They re the best family one could wish for.

References

- Abou-Shanab, R. A., J. S. Angle, T. A. Delorme, R. L. Chaney, P. van Berkum, H. Moawad, K. Ghanem and H. A. Ghazlan (2003). "Rhizobacterial effects on nickel extraction from soil and uptake by *Alyssum murale*." *New Phytol.* **158**(1): 219-224.
- Adham, S. A. I., P. Honrubia, M. Díaz, J. M. Fernández-Abalos, R. I. Santamaría and J. A. Gil (2001). "Expression of the genes coding for the xylanase *xys1* and the cellulase *cel1* from the strew-decomposing *Streptomyces halstedii* jm8 cloned into the amino-acid producer *Brevibacterium lactofermentum* ATCC13869." *Arch. Microbiol.* **177**: 91-97.
- Admiraal, S. J., C. T. Walsh and C. Khosla (2001). "The loading module of rifamycin synthetase is an adenylation-thiolation didomain with tolerance for substituted benzoates." *Biochemistry* **40**: 6116-6123.
- Ahimou, F., P. Jacques and M. Deleu (2000). "Surfactin and iturin A effects on *Bacillus subtilis* surface hydrophobicity." *Enz. Microb. Tech.* **27**: 749-754.
- Alborn, H. T., T. H. Jones, G. S. Stenhagen and J. H. Tumlinson (2000). "Identification and synthesis of volicitin and related components from beet armyworm oral secretions." *J. Chem. Ecol.* **26**(1): 203-220.
- Alborn, H. T., T. C. J. Turlings, T. H. Jones, G. S. Stenhagen and J. H. Loughrin, Tumlinson (1997). "An elicitor of plant volatiles from beet armyworm oral secretion." *Science* **276**: 945-949.
- Almirón, M., A. J. Link, D. Furlong and R. Kolter (1992). "A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*." *Genes Dev.* **6**: 2646-2654.
- Arimura, G.-i., R. Ozawa, T. Shimoda, T. Nishioka, W. Boland and J. Takabayashi (2000). "Herbivory-induced volatiles elicit defence genes in lima bean leaves." *Nature* **406**: 512-515.
- Avrahami, D. and Y. Shai (2004). "A new group of antifungal and antibacterial lipopeptides derived from non membrane active peptides conjugated to palmitic acid." *J. Biol. Chem.* **279**(13): 12277-12285.
- Bar-Ness, R., N. Avrahamy, T. Matsuyama and M. Rosenberg (1988). "Increased cell surface hydrophobicity of a *Serratia marcescens* ns38 mutant lacking wetting activity." *J. Bacteriol.* **170**(9): 4361-4364.
- Barrett, E., C. Stanton, O. Zelder, G. Fitzgerald and R. P. Ross (2004). "Heterologous expression of lactose- and galactose-utilizing pathways from lactic acid bacteria in *Corynebacterium glutamicum* for production of lysine in the whey." *Appl. Environ. Microbiol.* **70**(5): 2861-2866.
- Batrakov, S. G. and L. D. Bergelson (1978). "Lipids of the streptomycetes, structural investigation and biological interrelation." *Chem. Phys. Lipids* **21**: 1-29.
- Batrakov, S. G. and D. I. Nikitin (1996). "Lipid composition of the phosphatidylcholine-producing bacterium *Hyphomicrobium vulgare* np-160." *Biochim. Biophys. Acta* **1302**: 129-137.
- Behrendt, U., A. Ulrich and P. Schumann (2001). "Description of *Microbacterium foliorum* sp. nov. and *Microbacterium phyllosphaerae* sp. nov., isolated from the phyllosphere of grasses and the surface litter after mulching the sward, and reclassification of *Aureobacterium resistens* (Funke et al. 1998) as *Microbacterium resistens* comb. nov." *Int. J. System. Evol. Microbiol.* **51**: 1267-1276.
- Bentley, S. D., K. F. Chater, A.-M. Cerdano-Tárraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser and D. Harper (2002). "Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)." *Nature* **417**: 141-147.

References

- Boulton, C. A. (1989). Extracellular microbial lipids. Microbial lipids. C. Ratledge and S. G. Wilkinson. London, Academic Press. **2**: 669-694.
- Bozzi, M., G. Mignogna, S. Stefanini, D. Barra, C. Longhi, P. Valenti and E. Chiancone (1997). "A novel non-heme iron-binding ferritin related to the DNA-binding proteins of the Dps family in *Listeria innocua*." J. Biol. Chem. **272**: 3259-3265.
- Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal. Biochem. **72**: 248-254.
- Brady, S. F., C. J. Chao and J. Clardy (2002). "New natural product families from an environmental DNA (eDNA) gene cluster." J. Am. Chem. Soc. **124**(34): 9968-9969.
- Brady, S. F. and J. Clardy (2000). "Long-chain *N*-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA." J. Am. Chem. Soc. **122**: 12903-12904.
- Braun, V. and H. C. Wu (1994). Lipoproteins, structure, function, biosynthesis and model for protein export. Bacterial cell wall. J. M. Ghuyssen and R. Hakenbeck. Amsterdam, Elsevier Science B.V.: 319-341.
- Brenner, C. (2002). "Catalysis in the nitrilase superfamily." Curr. Opin. Struct. Biol. **12**: 775-782.
- Brentjens, R. J., M. Ketterer, M. A. Apicella and S. M. Spinola (1996). "Fine tangled pili expressed by *Haemophilus ducreyi* are a novel class of pili." J. Bacteriol. **178**: 808-816.
- Carter, P. and J. A. Wells (1988). "Dissecting the catalytic triad of a serine protease." Nature **332**: 564-568.
- Ceci, P., A. Ilari, E. Falvo and E. Chiancone (2003). "The Dps protein of *Agrobacterium tumefaciens* does not bind to DNA but protects it toward oxidative cleavage." J. Biol. Chem. **278**: 20319-20326.
- Cheung, V. G. and S. F. Nelson (1996). "Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA." Proc. Natl. Acad. Sci. USA **93**: 14676-14679.
- Clarkson, J. and D. A. Smith (2001). "UV raman evidence of a tyrosine in apo-human serum transferrin with a low pK_a that is elevated upon binding of sulphate." FEBS Lett. **503**: 30-34.
- Collins, M. D. and J. F. Bradbury (1991). The genera *Agromyces*, *Aureobacterium*, *Clavibacter*, *Curtobacterium*, and *Microbacterium*. The prokaryotes. A. Balows, H. G. Trüper, M. Dworkin et al. New York, Springer-Verlag. **2**: 1355-1368.
- Connell, N. D. (2001). "Expression systems for use in actinomycetes and related organisms." Curr. Opin. Biotechnol. **12**: 446-449.
- Cravatt, B. F., D. K. Giang, S. P. Mayfield, D. L. Boger, R. A. Lerner and N. B. Gilula (1996). "Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides." Nature **384**: 83-87.
- Cravatt, B. F., A. Saghatelian, E. G. Hawkins, A. B. Clement, M. H. Bracey, and A. H. Lichtman (2004). "Functional disassociation of the central and peripheral fatty acid amide signaling systems" Proc. Natl. Acad. Sci. USA **101**: 10821-10826.
- Cull, M. and C. S. McHenry (1990). Preparation of extracts from prokaryotes. Guide to protein purification. M. P. Deutscher. San Diego, Academic Press. **182**: 147-193.
- da Fonseca, D. P. A. J., D. Joosten, R. van der Zee, D. L. Jue, M. Singh, H. M. Vordermeier, H. Snippe and A. F. M. Verheul (1998). "Identification of new

References

- cytotoxic t-cell epitopes on the 38-kilodalton lipoglycoprotein of *Mycobacterium tuberculosis* by using lipopeptides." *Infect. Immun.* **66**(7): 3190-3197.
- Dardel, F., S. Ragusa, C. Lazennec, S. Blanquet and T. Meinel (1998). "Solution structure of nickel-peptide deformylase." *J. Mol. Biol.* **280**: 501-513.
- De Moraes, C. M., W. J. Lewis, P. W. Pare, H. T. Alborn and J. H. Tumlinson (1998). "Herbivore-infested plants selectively attract parasitoids." *Nature* **393**: 570-573.
- de Souza, J. T., M. de Boer, P. de Waard, T. A. van Beek and J. M. Raaijmakers (2003). "Biochemical, genetic, and zoosporicidal properties of cyclic lipopeptide surfactants produced by *Pseudomonas fluorescens*." *Appl. Environ. Microbiol.* **69**(12): 7161-7172.
- Dees, C. and J. M. Shively (1982). "Localization and quantification of the ornithine lipid of *Thiobacillus thiooxidans*." *J. Bacteriol.* **149**: 798-799.
- DeLong, E. F. (1992). "Archaea in coastal marine environments." *Proc. Natl. Acad. Sci. USA* **89**: 5685-5689.
- Desai, J. D. and I. M. Banat (1997). "Microbial production of surfactants and their commercial potential." *Microbiol. Mol. Biol. Rev.* **61**(1): 47-64.
- Dillon, R. J. and V. M. Dillon (2004). "The gut bacteria of insects: Nonpathogenic interaction." *Annu. Rev. Entomol.* **49**: 71-92.
- Dillon, R. J., C. T. Vennard and A. K. Charnley (2000). "Exploitation of gut bacteria in the locust." *Nature* **403**: 851-851.
- Dugas, J. E., L. Zurek, B. J. Paster, B. A. Keddie and E. R. Leadbetter (2001). "Isolation and characterization of a *Chryseobacterium* strain from the gut of the american cockroach, *periplaneta americana*." *Arch. Microbiol.* **175**: 259-262.
- Dundon, W. G., A. Polenghi, G. D. Guidice, R. Rappuoli and C. Monteucco (2001). "Neutrophil-activating protein (hp-nap) versus ferritin (pfr): Comparison of synthesis in *helicobacter pylori*." *FEMS Microbiol. Lett.* **199**: 143-149.
- Dunny G. M. and B. A. B. Leonard (1997) "Cell-cell communication in gram-positive bacteria" *Annu. Rev. Microbiol.* **51**:527-564.
- Ehmann, D. E., J. W. Trauger, T. Stachelhaus and C. T. Walsh (2000). "Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases." *Chem. Biol.* **7**: 765-772.
- Evans Jr., D. J., D. G. Evans, H. C. Lampert and H. Nakano (1995). "Identification of four new prokaryotic bacterioferritins, from *Helicobacter pylori*, *Anabaena variabilis*, *Bacillus subtilis*, and *Treponema pallidum*, by analysis of gene sequences." *Gene* **153**: 123-127.
- Floss, H. G. and T.-W. Yu (1999). "Lessons from the rifamycin biosynthetic gene cluster." *Curr. Opin. Chem. Biol.* **3**: 592-597.
- Fournand, D. and A. Arnaud (2001). "Aliphatic and enantioselective amidases: From hydrolysis to acyl transfer activity." *J. Appl. Microbiol.* **91**: 381-393.
- Frey, M., D. Spittler, W. Boland and A. Gierl (2004). "Transcriptional activation of *igl*, the gene for indole formation in *Zea mays*: A structure-activity study with elicitor-active *N*-acyl glutamines from insects." *Phytochemistry* **65**: 1047-1055.
- Frohman, M. A., M. K. Dush and G. R. Martin (1988). "Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer." *Proc. Natl. Acad. Sci. USA* **85**: 8998-9002.
- Fukushima, T., S. Ishikawa, H. Yamamoto, N. Ogasawara and J. Sekiguchi (2003). "Transcriptional, functional and cytochemical analyses of the *veg* gene in *Bacillus subtilis*." *J. Biochem* **133**: 475-483.
- Galm, U., M. A. Dessoy, J. Schmidt, L. A. Wessjohann and L. Heide (2004). "*In vitro* and *in vivo* production of new aminocoumarins by a combined biochemical, genetic, and synthetic approach." *Chem. Biol.* **11**: 173-183.

References

- Ghosh, M. and A. K. Mishra (1987). "Occurrence, identification and possible significance of ornithine lipid in *Thiobacillus ferrooxidans*." Biochem. Biophys. Res. Commun. **142**(3): 925-931.
- Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer and G. Stormo (1981). "Translational initiation in prokaryotes." Annu. Rev. Microbiol. **35**: 365-403.
- Grant, R. A., D. J. Filman, S. E. Finkel, R. Kolter and J. M. Hogle (1998). "The crystal structure of Dps, a ferritin homolog that binds and protects DNA." Nature Struct. Biol. **5**: 294-303.
- Gupta, S. and D. Chatterji (2003). "Bimodal protection of DNA by *Mycobacterium smegmatis* DNA-binding protein from stationary phase cells." J. Biol. Chem. **278**: 5235-5241.
- Gupta, S., S. B. Pandit, N. Srinivasan and D. Chatterji (2002). "Proteomics analysis of carbon-starved *Mycobacterium smegmatis*: Induction of Dps-like protein." Protein Eng. **15**: 503-511.
- Hairston, N. G., F. E. Smith and L. B. Slobodkin (1960). "Community structure, population control, and competition." Amer. Nat. **94**: 421-425.
- Halitschke, R., U. Schittko, G. Pohnert, W. Boland and I. T. Baldwin (2001). "Molecular interactions between the specialist herbivore *Maduca sexta* (Lepidoptera, Sphingidae) and its natural conjugates in herbivore oral secretions are necessary and sufficient for herbivore-specific plant responses." Plant Physiol. **125**: 711-717.
- Hallmann, J., R. Rodríguez-Kábana and J. W. Kloepper (1999). "Chitin-mediated changes in bacterial communities of the soil, rhizosphere and within roots of cotton in relation to nematode control." Soil Biol. Biochem. **31**: 551-560.
- Hamann, C., J. Hegemann and A. Hildebrandt (1999). "Detection of polycyclic aromatic hydrocarbon degradation genes in different soil bacteria by polymerase chain reaction and DNA hybridization." FEMS Microbiol. Lett. **173**: 255-263.
- Hébraud, M. and J. Guzzo (2000). "The main cold shock protein of *Listeria monocytogenes* belongs to the family of ferritin-like proteins." FEMS Microbiol. Lett. **190**: 29-34.
- Heerklotz, H. and J. Seedlig (2001). "Detergent-like action of the antibiotic peptide surfactin on lipid membranes." Biophys. J. **81**: 1547-1554.
- Hino, M., A. Fujie, T. Iwamoto, T. Hori, M. Hashimoto, Y. Tsurumi, K. Sakamoto, S. Takase and S. Hashimoto (2001). "Chemical diversity in lipopeptide antifungal antibiotics." J. Ind. Microbiol. Biotech. **27**: 157-162.
- Hoffmann, D., J. Hevel, R. E. Moore and B. S. Moore (2003). "Sequence analysis and biochemical characterization of the nostopeptolide biosynthetic gene cluster from *Nostoc* sp. Gsv224." Gene **311**: 171-180.
- Holt, J. G., Ed. (1993). Bergey's manual of determinative bacteriology. Baltimore, Williams & Wilkins.
- Honda, K., M. Kataoka, H. Ono, K. Sakamoto, S. Kita and S. Shimizu (2002). "Purification and characterization of a novel esterase promising for the production of useful compounds from *Microbacterium* sp. 7-1w." FEMS Microbiol. Lett. **206**: 221-227.
- Horowitz, S. and W. M. Griffin (1991). "Structural analysis of *Bacillus licheniformis* 86 surfactant." J. Ind. Microbiol. Biotechnol. **7**: 45-52.
- Hu, Z., D. A. Hopwood and C. Khosla (2000). "Directed transfer of large DNA fragments between *Streptomyces* species." Appl. Environ. Microbiol. **66**(5): 2274-2277.
- Hwang, P. M., W.-Y. Choy, E. I. Lo, L. Chen, J. D. Forman-Kay, C. R. H. Raetz, G. G. Prive, R. E. Bishop and L. E. Kay (2002). "Solution structure and dynamic of

References

- the outer membrane enzyme PAGP by NMR." Proc. Natl. Acad. Sci. USA **99**: 13560-13565.
- Ilari, A., P. Ceci, D. Ferrari, G. L. Rossi and E. Chiancone (2002). "Iron incorporation into *Escherichia coli* Dps gives rise to a ferritin like microcrystalline core." J. Biol. Chem. **277**: 37619-37623.
- Ilari, A., S. Stefanini, E. Chiancone and D. Tsernoglou (2000). "The dodecameric ferritin from *Listeria innocua* contains a novel intersubunit iron-binding site." Nature Struct. Biol. **7**: 38-43.
- Isenegger, D. A., P. W. J. Taylor, K. Mullins, G. R. McGregor, M. Barlass and J. F. Hutchinson (2003). "Molecular detection of a bacterial contaminant *Bacillus pumilus* in symptomless potato plant tissue cultures." Plant Cell Rep. **21**: 814-820.
- Kadavy, D. R., J. M. Hornby, T. Haverkost and K. W. Nickerson (2000). "Natural antibiotic resistance of bacteria isolated from larvae of the oil fly, *Helaeomyia petrolei*." Appl. Environ. Microbiol. **66**: 4615-4619.
- Kang, J.-G., M.-Y. Hahn, A. Ishihama and J.-H. Roe (1997). "Identification of sigma factor for growth phase-related promoter selectivity of RNA polymerases from *Streptomyces coelicolor* A3(2)." Nucl. Acids. Res. **25**(13): 2566-2573.
- Kawai, Y., Y. Nakgawa, T. Matuyama, K. Akagawa, K. Itagawa, K. Fukase, S. Kusumoto, M. Nishijima and I. Yano (1999). "Atypical bacterial ornithine-containing lipid N^{alfa} -(d)-[3-(hexadecanoyloxy)hexadecanoyl]-ornithine is a strong stimulant for macrophages and a useful adjuvant." FEMS Immun. Med. Microbiol. **23**: 67-73.
- Kawai, Y. and I. Yano (1983). "Ornithine-containing lipid of *Bordetella pertussis*, a new type of hemagglutinin." Eur. J. Biochem. **136**: 531-538.
- Kawai, Y., I. Yano and K. Kaneda (1988). "Various kinds of lipoamino acids including a novel serine-containing lipid in an opportunistic pathogen *Flavobacterium*." Eur. J. Biochem. **171**: 73-80.
- Kawai, Y., I. Yano, K. Kaneda and E. Yabuuchi (1988). "Ornithine-containing lipids of some *Pseudomonas* species." Eur. J. Biochem. **175**: 633-641.
- Kawazoe, R., H. Okuyama, W. Reichardt and S. Sasaki (1991). "Phospholipids and a novel glycine-containing lipoamino acid in *Cytophaga johnsonae* stanier strain c21." J. Bacteriol. **173**(17): 5470-5475.
- Keating, T. A., C. G. Marshall and C. T. Walsh (2000). "Vibriobactin biosynthesis in *Vibrio cholerae*: VibH is an amide synthase homologous to nonribosomal peptide synthetase condensation domains." Biochemistry **39**: 15513-15521.
- Kessler, A. and I. T. Baldwin (2002). "Plant responses to insect herbivory: The emerging molecular analysis." Annu. Rev. Plant. Biol. **53**: 299-328.
- Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood (2000). Practical *Streptomyces* genetics. Colney, Norwich, The John Innes Foundation.
- Kirchner, O. and A. Tauch (2003). "Tools for genetic engineering in the amino acid - producing bacterium *Corynebacterium glutamicum*." J. Biotechnol. **104**: 287-299.
- Kobayashi, M., Y. Fujiwara, M. Goda, H. Komeda and S. Shimizu (1997). "Identification of active sites in amidase: Evolutionary relationship between bond- and peptide bond-cleaving enzyme." Proc. Natl. Acad. Sci. USA **94**: 11986-11991.
- Kobayashia, M., T. Nagasawab and H. Yamadaa (1992). "Enzymatic synthesis of acrylamide: A success story not yet over." Trends Biotechnol. **10**: 402-408.
- Konz, D., S. Doekel and M. A. Marahiel (1999). "Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin." J. Bacteriol. **181**(1): 133-140.

References

- Krämer, R. (1994). "Secretion of amino acids by bacteria: Physiology and mechanism." FEMS Microbiol. Rev. **13**: 75-94.
- Kriakov, J., S. h. Lee and W. R. Jacobs, Jr. (2003). "Identification of a regulated alkaline phosphatase, a cell surface-associated lipoprotein, in *Mycobacterium smegmatis*." J. Bacteriol. **185**(6): 4983-4991.
- Kruszka, K. K. and R. W. Gross (1994). "The ATP- and CoA-independent synthesis of arachidonoyl ethanolamide." J. Biol. Chem. **269**(20): 14345-14348.
- Kuczek, K., K. Pawlik, M. Kotowska and M. Mordarski (1997). "*Streptomyces coelicolor* DNA homologous with acyltransferase domains of type I polyketide synthase gene complex." FEMS Microbiol. Lett. **157**: 195-200.
- Kumar, T. R. and M. N. V. Prasad (1999). "Metal binding properties of ferritin in *Vigna mungo* (L.) hepper (black gram): Possible role in heavy metal detoxification." Bull. Environ. Contam. Toxicol. **62**: 502-507.
- Kuzina, L. V., J. J. Peloquin, D. C. Vacek and T. A. Miller (2001). "Isolation and identification of bacteria associated with adult laboratory *Mexican fruit flies*, *Anastrepha ludens* (Diptera: Tephritidae)." Curr. Microbiol. **42**: 290-294.
- Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**: 680-685.
- Lait, C. G., H. T. Alborn, P. E. A. Teal and J. H. Tumlinson (2003). "Rapid biosynthesis of *N*-linolenoyl-L-glutamine, an elicitor of plant volatiles, by membrane-associated enzyme(s) in *manduca sexta*." Proc. Natl. Acad. Sci. USA **100**(12): 7027-7032.
- Lamber, H., F. S. Chapin III and T. L. Pons (1998). Plant physiological ecology. New York, Springer-Verlag Inc.
- Laneelle, M.-A., D. Prome, G. Laneelle and J.-C. Prome (1990). "Ornithine lipid of *Mycobacterium tuberculosis*: Its distribution in some slow- and fast-growing mycobacteria." J. Gen. Microbiol. **136**: 773-778.
- Lang, S. (2002). "Biological amphiphiles (microbial biosurfactants)." Curr. Opin. Coll. Int. Sci. **7**: 12-20.
- Lang, S., W. Beil, H. Tokuda, W. C. and V. Lurtz (2004). "Improved production of bioactive glucosylmannosyl-glycerolipid by sponge-associated *Microbacterium* species." Mar. Biotechnol. **6**: 152-156.
- Lees G. and A. Dougalis (2004). "Differential effects of the sleep-inducing lipid oleamide and cannabinoids on the induction of long-term potentiation in the CA1 neurons of the rat hippocampus *in vitro*" Brain Res. **997**: 1- 14.
- Lebuhn, M., W. Achouak, M. Schloter, O. Berge, H. Meier, M. Barakat, A. Hartmann and T. Heulin (2000). "Taxonomic characterization of *Ochrobactrum* sp. Isolates from soil samples and wheat roots, and description of *Ochrobactrum tritici* sp. Nov. and *Ochrobactrum grignonense* sp. Nov." Int. J. System. Evol. Microbiol. **50**: 2207-2223.
- Li, X., X. Zhou and Z. Deng (2003). "Vector systems allowing efficient autonomous or integrative gene cloning in *Micromonospora* sp. Strain 40027." Appl. Environ. Microbiol. **69**(6): 3144-3151.
- Linder, M. E. and R. J. Deschenes (2003). "New insights into the mechanisms of protein palmitoylation." Biochemistry **42**: 4311-4320.
- Lindum, P. W., U. Anthoni, C. Christophersen, L. Eberl, S. Molin and M. Givskov (1998). "*N*-acyl-L-lactone autoinducers control production of an extracellular lipopeptide biosurfactant required for swarming motility of *Serratia liquefaciens* MG1." J. Bacteriol. **180**(23): 6384-6388.
- Matsuyama, T., A. Bhasin and R. M. Harshey (1995). "Mutational analysis of flagellum-independent surface spreading of *Serratia marcescens* 274 on a low-agar medium." J. Bacteriol. **177**(4): 987-991.

References

- Matsuyama, T., M. Fujita and I. Yano (1985). "Wetting agent produced by *Serratia marcescens*." FEMS Microbiol. Lett. **28**: 125-129.
- Mattiacci, L., M. Dicke and M. A. Posthumus (1995). "Beta-glucosidase—an elicitor of herbivore-induced plant odor that attracts host-searching parasitic wasps." Proc. Natl. Acad. Sci. USA **92**: 2034-40.
- McDaniel, R., S. Ebert-Khosla, D. A. Hopwood and C. Khosla (1993). "Engineered biosynthesis of novel polyketides." Science **262**: 1546-1550.
- McPartland, J., V. D. Marzo, L. D. Petrocellis, A. Mercer and M. Glass (2001). "Cannabinoid receptors are absent in insects." J. Com. Neurol. **436**: 423-429.
- Menkhaus, M., C. Ullrich, B. Kluge, J. Vater, D. Vollenbroich and R. M. Kamp (1993). "Structural and functional organization of the surfactin synthase multienzyme system." J. Biol. Chem. **268**(11): 7678-7684.
- Merkler, D. J., K. A. Merkle, W. Stern and F. F. Fleming (1996). "Fatty acid amide biosynthesis: A possible new role for peptidylglycine α -amidating enzyme and acyl-coenzyme A:Glycine *N*-acyltransferase." Arch. Biochem. Biophys. **330**: 430-434.
- Minnikin, D. E. and H. Abdolrahimzadeh (1974). "The replacement of phosphatidylethanolamine and acidic phospholipids by ornithine-amide lipid and a minor phosphorus-free lipid in *Pseudomonas fluorescens* ncb129." FEBS Lett. **43**(3): 257-260.
- Moran, C., P., Jr., N. Lang, S. F. J. LeGrice, G. Lee, M. Stephens, A. L. Sonenshein, J. Pero and R. Losick (1982). "Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*." Mol. Gen. Genet. **186**: 339-346.
- Mori, N., H. T. Alborn, P. E. A. Teal and J. H. Tumlinson (2001). "Enzymatic decomposition of elicitors of plant volatiles in *Heliothis virescens* and *helicoverpa zea*." J. Insect Physiol. **47**: 749-757.
- Munford, R. S. and J. P. Hunter (1992). "Acyloxylacyl hydrolase, a leukocyte enzyme that decylates bacterial lipopolysaccharides, has phospholipase, diacylglycerolipase, and acyltransferase activities *in vitro*." J. Biol. Chem. **267**: 10016-10121.
- Nakano, M. M., R. Magnuson, A. Myers, J. Curry, A. D. Grossman and P. Zuber (1991). "*Srfa* is an operon required for surfactin production and competence development, and efficient sporulation in *Bacillus subtilis*." J. Bacteriol. **173**: 1770-1778.
- Nakashima, N. and T. Tamura (2004). "A novel system for expression recombinant proteins over a wide temperature range from 4 to 35 °C." Biotechnol. Bioeng. **86**(2): 136-148.
- Navarre, W. W., S. Daefler and O. Schneewind (1996). "Cell wall sorting of lipoproteins in *Saphylococcus aureus*." J. Bacteriol. **178**(2): 441-446.
- Navarre, W. W. and O. Schneewind (1999). "Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope." Microbiol. Mol. Biol. Rev. **63**(1): 174-229.
- Neumann, S. and M.-R. Kula (2002). "Gene cloning, overexpression and biochemical characterization of the peptide amidase from *Stenotrophomonas maltophilia*." Appl. Microbiol. Biotechnol. **58**: 772-780.
- Novick, R. (1993). Staphylococcus. *Bacillus subtilis* and other gram-positive bacteria. A. L. Sonenshein, J. A. Hoch and R. Losick. Washington, American Society for Microbiology: 17-33.
- Oliver, K. M., J. A. Russell, N. A. Moran and H. Martha S (2003). "Facultative bacterial symbionts in aphids confer resistance to parasitic wasps." Proc. Natl. Acad. Sci. USA **100**(4): 1803-1807.

References

- Opdyke, J. A., J. R. Scott and C. Moran, P., Jr. (2003). "Expression of the secondary sigma factor σ^x in streptococcus pyogenes is restricted at to levels." J. Bacteriol. **185**(15): 4291-4297.
- Osterreicher-Ravid, D., E. Z. Ron and E. Rosenberg (2000). "Horizontal transfer of an exopolymer complex from one bacterial species to another." Environ. Microbiol. **2**: 366-372.
- Otto, H.-H. and T. Schirmeister (1997). "Cysteine protease and their inhibitors." Chem. Rev. **97**: 133-171.
- Pace, H. C. and C. Brenner (2001). "The nitrilase superfamily: Classification, structure and function." Genome Biol. **2**(1): 1-9.
- Papinutto, E., W. G. Dundon, N. Pitulis, R. Battistutta, C. Montecuccio and G. Zanotti (2002). "Structure of two iron-binding proteins from *Bacillus anthracis*." J. Biol. Chem. **277**: 15093-15098.
- Pare, P. W., H. T. Alborn and J. H. Tumlinson (1998). "Concerted biosynthesis of an insect elicitor of plant volatiles." Proc. Natl. Acad. Sci. USA **95**: 13971-13975.
- Parsek, M. R. and E. P. Greenberg (2000). "Acyl-homoserine lactone quorum sensing in gram-negative bacteria: A signaling mechanism involved in associations with higher organisms." Proc. Natl. Acad. Sci. USA **97**: 8789-8793.
- Patricelli, M. P. and B. F. Cravatt (2000). "Clarifying the catalytic roles of conserved residues in the amidase signature family." J. Biol. Chem. **275**(25): 19177-19184.
- Peña, M. M. O. and G. S. Bullerjahn (1995). "The DpsA protein of *Synechococcus* sp. Strain PCC7942 is a DNA-binding hemoprotein." J. Biol. Chem. **270**: 22478-2282.
- Pfeifer, B. A. and C. Khosla (2001). "Biosynthesis of polyketides in heterologous hosts." Microbiol. Mol. Biol. Rev. **65**(1): 106-118.
- Piel, J. (2002). "A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles." Proc. Natl. Acad. Sci. USA **99**(22): 14002-14007.
- Pohnert, G., V. Jung, E. Haukioja, K. Lempa and W. Boland (1999). "New fatty acid amides from regurgitant of lepidopteran (noctuidae, geometridae) caterpillars." Tetrahedron **55**: 11275-11280.
- Pompeo, F., A. Mushtaq and E. Sim (2002). "Expression and purification of the rafamycin amide synthase, rff, an enzyme homologous to the prokaryotic arylamine *N*-acetyltransferases." Protein Express. Purif. **24**: 138-151.
- Powers, J. C., J. L. Asgian, O. D. Ekici and K. E. James (2002). "Irreversible inhibitors of serine, cysteine, and threonine proteases." Chem. Rev. **102**: 4639-4750.
- Pugsley, A. P. and M. G. Kornacker (1991). "Secretion of the cell surface lipoprotein pullulanase in *Escherichia coli*." J. Biol. Chem. **266**(21): 13640-13645.
- Pullainen, A. T., S. Haataja, S. Kähkönen and J. Finne (2003). "Molecular basis of H₂O₂ resistance mediated by Streptococcal dpr." J. Biol. Chem. **278**: 7996-8005.
- Rainey, F., N. Weiss, H. Prauser and E. Stackerbraudt (1994). "Further evidence for the phylogenetic coherence of actinomycetes with group B-peptidoglycan and evidence for the phylogenetic intermixing of the genera *Microbacterium* and *Aureobacterium* as determined by 16S rDNA analysis." FEMS Microbiol. Lett. **118**: 135-140.
- Rami, B. R. and J. B. Udgaonkar (2001). "Ph-jump-induced folding and unfolding studies of barstar: Evidence for multiple folding and unfolding pathways." Biochemistry **40**: 15267-15279.
- Reggio, P. H. (2002). "Endocannabinoid structure-activity relationships fro interaction at the cannabinoid receptors." Prost. Leuk. Ess. Fatty acids **66**(2&3): 143-160.

References

- Reindel, S., S. Anemuller, A. Sawaryn and B. F. Matzanke (2002). "The DpsA-homologue of the archaeon *Halobacterium salinarum* is a ferritin." Biochem. Biophys. Acta **1598**: 140-146.
- Ren, B., G. Tibbelin, T. Kajino, O. Asami and R. Ladenstein (2003). "The multi-layered structure of Dps with a novel di-nuclear ferroxidase center." J. Mol. Biol. **329**: 467-477.
- Rocha, E. R., G. Owens Jr. and C. J. Smith (2000). "The redox-sensitive transcriptional activator OxyR regulates the peroxide response regulation in the obligate anaerobe *Bacteroides fragilis*." J. Bacteriol. **182**: 5059-5069.
- Ron, E. Z. and E. Rosenberg (2001). "Natural roles of biosurfactants." Environ. Microbiol. **3**(4): 229-236.
- Rosenberg, E. and E. Z. Ron (1999). "High- and low-molecular-mass microbial surfactants." Appl. Microbiol. Biotechnol. **52**: 154-162.
- Rosenheim, J. A. (1998). "Higher-order predators and the regulation of insect herbivore populations." Annu. Rev. Entomol. **43**: 421-447.
- Saiz, J. L., C. Lopez-Zumel, B. Monterroso, J. Varea, J. L. R. Arrondo, I. Iloro, J. L. Garcia, J. Laynez and M. Menendez (2002). "Characterization of ejl, the cell-wall amidase coded by the pneumococcal bacteriophage ej-1." Protein Sci. **11**: 1788-1799.
- Sambrook, J., E. F. Fritsch and T. Maniatis (2001). Small-scale preparation of plasmid DNA. Molecular cloning. New York, Cold Spring Harbor Laboratory Press. **2**: 1.25.
- Sankaran, K. and H. C. Wu (1994). "Lipid modification of bacterial prolipoprotein." J. Biol. Chem. **269**: 19701-19706.
- Sazuka, T. and O. Ohara (1996). "Sequence features surrounding the translation initiation sites assigned on the genome sequence of *Synechocystis* sp. Strain PCC6803 by amino-terminal protein sequencing." DNA Res. **3**: 225-232.
- Schmelz, E. A., H. T. Alborn and J. H. Tumlinson (2001). "The influence of intact-plant and excised-leaf bioassay designs on volicitin- and jasmonic acid-induced sesquiterpene volatile release in *Zea mays*." Planta **214**: 171-179.
- Schmidt, W. M. and M. W. Mueller (1999). "Capselect: A highly sensitive method for 5' cap-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs." Nucl. Acids. Res. **27**: e31.
- Schmutz, E., M. Steffensky, J. Schmidt, A. Porzel, S.-M. Li and L. Heide (2003). "An unusual amide synthetase (CouL) from the coumermycin a₁ biosynthetic gene cluster from *Streptomyces rishiriensis* DSM 40489." Eur. J. Biochem. **270**: 4413-4419.
- Schweizer, E. (1989). Biosynthesis of fatty acids and related compounds. Microbial lipids. C. Ratledge and S. G. Wilkinson. London, Academic Press. **2**: 3-50.
- Shen, Z., J. Liu, R. L. Wells and M. M. Elkind (1993). "Cycle sequencing using degenerate primers containing inosines." BioTechniques **15**: 82-89.
- Shepard, B. D. and M. S. Gilmore (1999). "Identification of aerobically and anaerobically induced genes in *Enterococcus faecalis* by random arbitrarily primed PCR." Appl. Environ. Microbiol. **65**: 1470-1476.
- Shevchenko, A., M. Wilm, O. Vorm and M. Mann (1996). "Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gel." Anal. Chem. **68**: 850-858.
- Sifri, C. D., J. Begun, F. M. Ausube and S. B. Calderwood (2003). "*Caenorhabditis elegans* as a model host for *Staphylococcus aureus* pathogenesis." Infect. Immun. **71**: 2208-2217.
- Smedley, S. R., F. C. Schroeder, D. B. Weibel, J. Meinwald, K. A. Lafleur, J. A. Renwick, R. Rutowski and T. Eisner (2002). "Mayolenes: Labile defensive

References

- lipids from the glandular hairs of a caterpillar (*Pieris rapae*)." Proc. Natl. Acad. Sci. USA **99**(10): 6822-6827.
- Spiteller, D., K. Dettner and W. Boland (2000). "Gut bacteria may be involved in interactions between plants, herbivores and their predators: Microbial biosynthesis of *N*-acylglutamine surfactants as elicitors of plant volatiles." Biol. Chem. **381**: 755-762.
- Spiteller, D., G. Pohnert and W. Boland (2001). "Absolute configuration of volictin, an elicitor of plant volatile biosynthesis from lepidopteran larvae." Tetrahedron Lett. **42**: 1483-1485.
- Spiteller, D. (2002). Charakterisierung von *N*-Acyl-Glutaminkonjugaten aus dem Regurgitat von Lepidoptera Larven. Ph. D thesis, Friedrich-Schiller-Universit: p. 240.
- Spiteller, D. and W. Boland (2003a). "*N*-(15,16-dihydroxylinoleoyl)-glutamine and *n*-(15,16-epoxylinoleoyl)-glutamine isolated from oral secretions of lepidopteran larvae." Tetrahedron **59**: 135-139.
- Spiteller, D. and W. Boland (2003b). "*N*-(17-acyloxy-acyl)-glutamines: Novel surfactants from oral secretions of lepidopteran larvae." J. Org. Chem. **68**(23): 8743-8749.
- Stachelhaus, T., H. D. Mootz, V. Bergendahl and M. A. Marahiel (1998). "Peptide bond formation in nonribosomal peptide biosynthesis." J. Biol. Chem. **273**(35): 22773-22781.
- Stanier, R. Y., J. L. Ingraham, M. L. Wheelis and P. R. Painter (1987). General microbiology. Houndmills, Macmillans Press. Ltd.
- Stanley, P., C. Hyland, V. Koronakis and C. Hughes (1999). "An ordered reaction mechanism for nacterial toxin acylation by the specialized acyltransferase HlyC: Formation of a ternary complex with acylACP and protoxin substrates." Mol. Microbiol. **34**: 887-901.
- Stanley, P., V. Koronakis and C. Hughes (1998). "Acylation of *Escherichia coli* hemolysin: A unique protein lipidation mechanism underlying toxin function." Microbiol. Mol. Biol. Rev. **62**: 309-333.
- Steffensky, M., S.-M. Li and L. Heide (2000). "Cloning, overexpression, and purification of novobiocic acid synthetase from *Streptomyces spheroides* NCIMB 11891." J. Biol. Chem. **275**(28): 21754-21760.
- Sutcliffe, I. C. (1998). "Cell envelope composition and organisation in the genus *Rhodococcus*." Ant. von Leeuwenh. **74**: 49-58.
- Thiele, O. W., C. J. Biswas and D. H. Hunneman (1980). "Isolation and characterization of an ornithine-containing lipid from *Paracoccus denitrificans*." Eur. J. Biochem. **105**: 267-274.
- Thiele, O. W., J. Oulevey and D. H. Hunneman (1984). "Ornithine-containing lipids in *Thiobacillus A2* and *Achromobacter* sp." Eur. J. Biochem. **139**: 131-135.
- Tokunaga, M., H. Tokunaga and H. C. Wu (1982). "Post-translational modification and processing of *Escherichia coli* prolipoprotein *in vitro*." Proc. Natl. Acad. Sci. USA **79**: 2255-2259.
- Tonello, F., W. G. Dundon, B. Satin, M. Maurizio, G. Tognon, G. Gramdi, G. D. Giudice, R. Rappuoli and C. Montecucco (1999). "The *Helicobacter pylori* neutrophil-activating protein is an iron-binding protein with dodecameric structure." Mol. Microbiol. **34**: 238-246.
- Tremblay, D., J. Lemay, M. Gilbert, Y. Chapdelaine, C. Dupont and R. Morosoli (2002). "High-level heterologous expression and secretion in *Streptomyces lividans* of two major antigenic proteins from *Mycobacterium tuberculosis*." Can. J. Microbiol. **48**: 43-48.

References

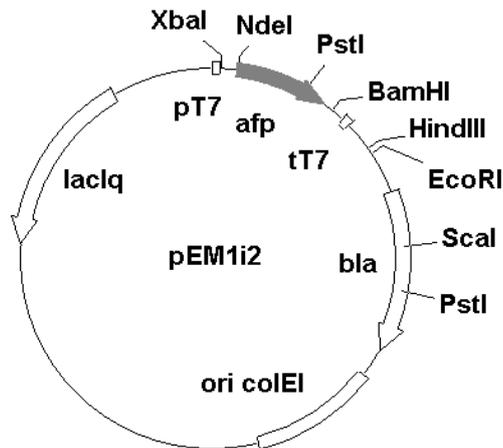
- Trivedi, O. A., P. Arora, V. Sridharan, R. Tickoo, D. Mohanty and R. S. Gokhale (2004). "Enzymatic activation and transfer of fatty acids as acyl-adenylates in mycobacteria." Nature **428**: 441-445.
- Truitt, C. L. and P. W. Paré (2004). "In situ translocation of volicitin by beet armyworm larvae to maize and systemic immobility of the herbivore elicitor *in planta*." Planta **218**: 999-1007.
- Truitt, C. L. H-X, Wei and P. W. Paré (2004). "A PlasmaMembrane Protein from Zeamays Binds with the Herbivore Elicitor Volicitin." The Plant Cell **16**: 523-532.
- Tsuge, K., T. Ano, M. Hirai, Y. Nakamura and M. Shoda (1999). "The genes *degQ*, *ppS*, and *lpA-8* (*sfp*) are responsible for conversion of *Bacillus subtilis* 168 to plipastatin production." Antimicrob. Agents Chemother. **43**(9): 2183-2192.
- Turlings, T. C. J., J. H. Tumlinson and W. J. Lewis (1990). "Exploitation of herbivore-induced plant odors by host-seeking parasitic wasps." Science **250**: 1251-1253.
- Turner, S., E. Reid, H. Smith and J. Cole (2003). "A novel cytochrome c peroxidase from *Neisseria gonorrhoeae*: A lipoprotein from gram negative bacterium." Biochem. J. **373**: 865-873.
- Vazeux, G., X. Iturrioz, P. Corvol and C. Llorens-Cortes (1997). "A tyrosine residue essential for catalytic activity in aminopeptidase A." Biochem. J. **327**: 883-889.
- Veerapandian, B., J. B. Cooper, A. Sali, T. L. Blundell, R. L. Rosati, B. W. Dominy, D. B. Damon and D. J. Hoover (1992). "Direct observation by x-ray analysis of the tetrahedral "intermediate" of aspartic proteinases." Protein Sci. **1**: 322-328.
- Verma, S. C., A. Singh, S. P. Chowdhury and A. K. Tripathi (2004). "Endophytic lolonization ability of two deep-water rice endophytes, *Pantoea* sp. and *Ochrobactrum* sp. using green fluorescent protein reporter." Biotechnol. Lett. **26**: 425-429.
- Wang, L., J. McVeyã and L. C. Vining (2001). "Cloning and functional analysis of a phosphopantetheinyl transferase superfamily gene associated with jadamycin biosynthesis in *Streptomyces venezuelae* ISP5230." Microbiology **147**: 1535-1545.
- Weissenmayer, B., J.-L. Gao, I. M. Lopez-Lara and O. Geiger (2002). "Identification of a gene required for the biosynthesis of ornithine-derived lipids." Mol. Microbiol. **45**(3): 721-733.
- Welsh, J., K. Chada, S. S. Dalal, R. Cheng, D. Ralph and M. McClelland (1992). "Arbitrarily primed PCR fingerprinting of RNA." Nucl. Acids Res. **20**: 4965-4970.
- Wenzel, M., I. Schonig, M. Berchtold, P. Kampfer and H. König (2002). "Aerobic and facultatively anaerobic cellulolytic bacteria from the gut of the termite *Zootermopsis angusticollis*." J. Appl. Microbiol. **92**(1): 32-40.
- Wilkinson, C. J., Z. A. Hughes-Thomas, C. J. Martin, I. Bohm, T. Mironenko, M. Deacon, M. Wheatcroft, G. Wirtz, J. Staunton and P. F. Leadlay (2002). "Increasing the efficiency of heterologous promoters in actinomycetes." J. Mol. Microbiol. Biotechnol. **4**(4): 417-426.
- Wilkinson, S. G. (1972). "Composition and structure of the ornithine-containing lipid from *Pseudomonas rubescens*." Biochem. Biophys. Acta **270**: 1-17.
- Wittgren, B., J. B. Borgstroem, L. Piculell and K.-G. Wahlund (1998). "Conformational change and aggregation of K-carrageenan studied by flow field-flow fractionation and mutiangle light scattering." Biopolymers **45**: 85-96.
- Wolf, S. G., D. Frenkiel, T. Arad, S. E. Finkel, R. Kolter and A. Minsky (1999). "DNA protection by stress-induced biocrystallization." Nature **400**: 83-85.

References

- Wouters, M. A. and A. Husain (2001). "Changes in zinc ligation promote remodeling of the active site in the zinc hydrolase superfamily." J. Mol. Biol. **314**: 1191-1207.
- Xia, J., Y. Xia and I. A. Nnanna (1995). "Structure-function relationship of acyl amino acid surfactants: Surface activity and antimicrobial properties." J. Agric. Food. Chem. **43**: 867-871.
- Xue, Q., G. Ashley, C. R. Hutchinson and D. V. Santi (1999). "A multiplasmid approach to preparing large libraries of ployketides." Proc. Natl. Acad. Sci. USA **96**: 11740-11745.
- Yagi, H., G. Corzo and T. Nakahara (1997). "N-acyl amino acid biosynthesis in marine bacterium, *Deleya marina*." Biochim. Biophys. Acta **1336**: 28-32.
- Yang, X., N. E. L. Brun, A. J. Thomson, G. R. Moore and N. D. Chasteen (2000). "The iron oxidation hand hydrolysis chemistry of *Escherichia coli* bacterioferritin." Biochemistry **39**: 4915-4923.
- Zhao, G., P. Ceci, A. Ilari, L. Giangiacomo, T. M. Laue, E. Chiancone and N. D. Chasteen (2002). "Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells." J. Biol. Chem. **277**: 27689-27696.
- Zimmer, C. (2001). "*Wolbachia*: A tale of sex and survival." Science **292**: 1093-1095.
- Zurek, L., C. Schal and D. W. Watson (2000). "Diversity and contribution of the intestinal bacterial community to the development of *Musca domestica* (diptera: Muscidae) larvae." J. Med. Entomol. **37**: 924-928.

Appendix I:

Maps of the plasmids constructed in this research



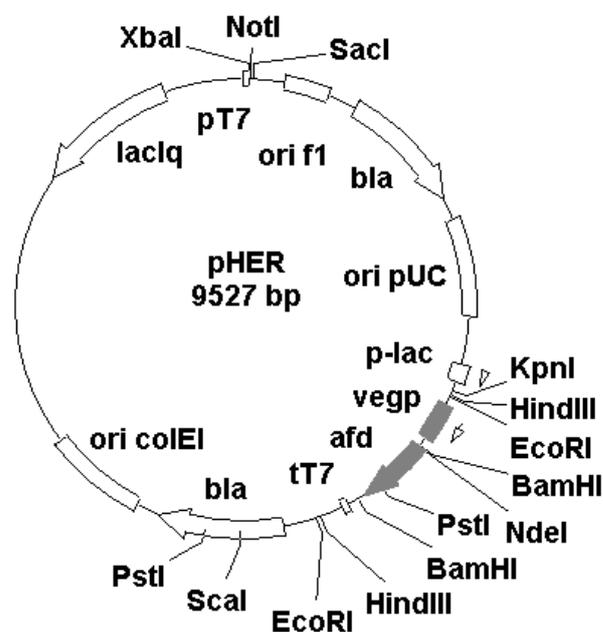
Feature	Nucleotide position
T7 promoter (pT7)	1–43
<i>afp</i> ORF	135–621
T7 terminator (tT7)	748–794
ampicillin resistance (<i>bla</i>)	1206–2063
pBR322 origin of replication (<i>ori colEI</i>)	2214–2881
<i>lacIq</i> repressor (<i>lacIq</i>)	5717–4761

The *afp* ORF and primers

CCCATCCATATGACCGACACCAACATCACCACCCCCGCCCTCACCGCCGACCCCGAGGTCGCCGCCG
 CTGCCGCGCAGTTCCTCACTCCGGTCGTGCACAAGATGCAGGCGCTCGTCGTCAACGGCAAGCAGGC
 GCACTGGAACGTCCGCGGCTCGAACTTCATCGCGATCCACGAGCTGCTCGACTCGGTTCGTCCGCTCAC
 GCCCAGGACTACGCCGACACCGCCGCGGAGCGCATCGTCCCTGGGCCCTGCCGATCGACTCCCGCG
 TCTCGACCATGGCCGAGAAGACCAGCACCGCCGTCCCCGCCGGCTTCGCGCAGTGGCAGGACGAGAT
 CAAGGCCATCGTCTCCGACATCGACGCAGCTCTCGTCGACCTGCAGGCCGCGATCGACGGTCTCGAC
 GAGGTGACCTGACGAGCCAGGATGTCGCCATCGAGATCAAGCGCGGCGTCGACAAGGACCGCTGGT
 TCCTTCTCGCGCACCTCGCCGAG**TAATACCCGGCGCCGACGGATCCTGACT**

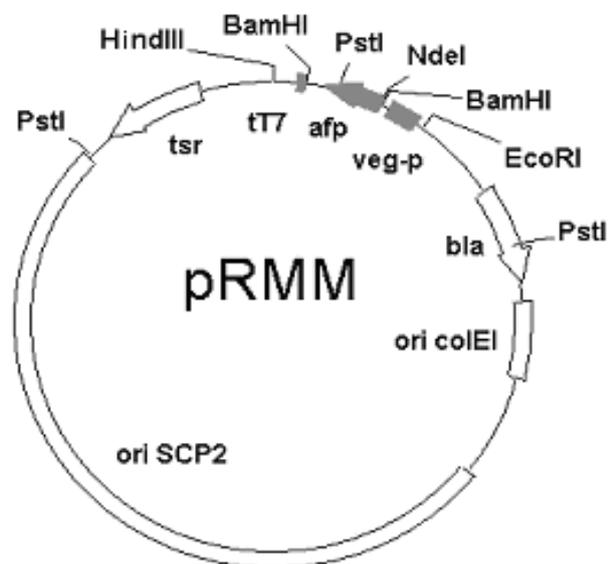
[The upper and lower primers used for introducing restriction sites (underlined) is in bold]

pEM1i2 is the plasmid used in *E. coli* expression. The *afp* gene has been inserted into plasmid **pET11a** (Stratagene)



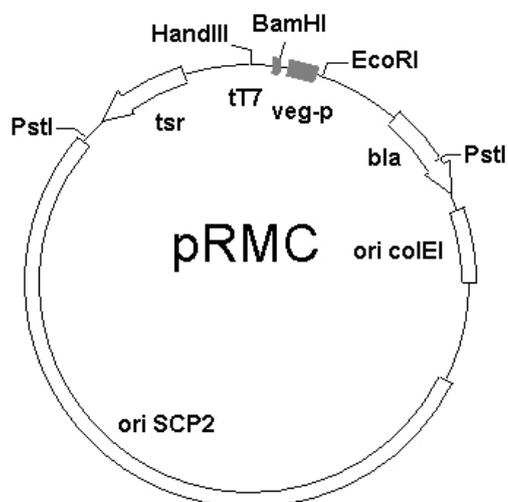
Feature	Nucleotide position
T7 promoter (pT7)	1–43
f1 origin (ori f1)	284–590
ampicillin resistance ORF 1 (bla)	789–1646
pUC origin (ori pUC)	1797–2464
<i>lac</i> promoter (p-lac)	2795–2916
<i>veg</i> Promoter (vegp)	3054–3334
<i>afd</i> ORF	3413–3899
T7 terminator (tT7)	4026–4072
ampicillin resistance ORF 2 (bla)	4484–5341
pBR322 origin of replication (ori colEI)	5492–6159
<i>lacIq</i> repressor (lacIq)	8995–8039

pHER is the intermediate plasmid produced during construction of pRMM and pRMC for *M. arborescens* DSM20754 expression. This is a fusion plasmid of **pEM1i2** and **pHN15** (Werner BioAgents, Jena, Germany)



Feature	Nucleotide position
T7 terminator (tT7)	257–211
<i>afp</i> ORF	915–430
<i>veg</i> promoter (veg-p)	1250–970
ampicillin resistance ORF (<i>bla</i>)	2057–2914
pBR332 origin (ori coEI)	3065–3732
SCP2* origin (ori SCP2*)	4632–11232
Thiostrepton resistance ORF (<i>tsr</i>)	12332–11482

pRMM is the plasmid used in *M. arborescens* DSM20754 expression. This plasmid was constructed by introducing a sliced expression cassette from **pHER** into the vector of **pSEK4** (Floss company), a derivative of **pRM5** (McDaniel *et al.* 1993).



Feature	Nucleotide position
T7 terminator (tT7)	257–211
veg promoter (veg-p)	347–627
ampicillin resistance ORF (bla)	1405–2262
pBR332 origin (ori coEI)	2413–3080
SCP2* origin (ori SCP2*)	3980–10580
Thiostrepton resistance ORF (tsr)	11680–10830

pRMC is the control plasmid in which the *afp* expression cassette has been removed from the *Bam*HI sites. All other features are identical with **pRMM**.

(These maps were drawn with freesoftware, plasmid processor developed by Department of Biotechnology, University of Kuopio, Finland)

Appendix II

Resistance of *M. arborescens* strains against selected antibiotics

Table I. The resistance profile of *M. arborescens* DSM20754

Antibiotic	Solvent	Concentration	Resistance
Apramycin	water	50µg/ml	+
Streptomycin	water	50µg/ml	-
Geneticin G418	water	40µg/ml	--
Thiostrepton	ethanol	50µg/ml	---
Gentamicin	water	40µg/ml	--
Kanamycin	water	50µg/ml	--
Chloramphenicol	ethanol	50µg/ml	---
Carbenicillin	water	100µg/ml	---
ClonNAT(HKI)	water	40µg/ml	-

“+”, resistant; “-”, susceptible, grows much slower than control; “- -”, more susceptible, only tiny colonies appeared after over night growth; “- - -”, highly susceptible, no growth observed after 24-h growth.

Table II. The resistance profile of *M. arborescens* SE14

Antibiotic	Solvent	Concentration	Resistance
Apramycin	water	50µg/ml	+
Streptomycin	water	50µg/ml	---
Geneticin G418	water	40µg/ml	-
Thiostrepton	ethanol	50µg/ml	---
Gentamicin	water	40µg/ml	---
Kanamycin	water	50µg/ml	+
Chloramphenicol	ethanol	50µg/ml	---
Carbenicillin	water	100µg/ml	---
ClonNAT(HKI)	water	40µg/ml	---

“+”, resistant; “-”, susceptible, grows much slower than control; “- -”, more susceptible, only tiny colonies appeared after over night growth; “- - -”, highly susceptible, no growth observed after 24-h growth.

Appendix III

Instruments

Instrument	Type	Company	Function
ÄKTA FPLC system		Amersham Biosciences	Protein purification
Edman degradation	amino acid sequencer 494A	Applied Biosystem model, USA	Sequencing protein
ESI-MS/MS sequencing	Q-TOF 1.5 hybrid mass spectrometer using "medium" nano ESI-capillaries	Micromass, UK And Protona, Denmark	Protein sequencing after trypsin digestion
FFF-MALS	18-angle DAWN-EOS and Optilab DSP interferometric refractometer	Wyatt Technology Corp.	Measurement of native protein molecular mass
Fluorescence microscope	Axionskop	Zeiss, Germany	Bacterial morphology observation
French Press		SLM Aminco	Disintegration of bacteria
Hybridisation oven		Biometra	Southern blotting and FISH
HPLC-MS	LCQ-Mass spectrometry	Finnigan	Activity assay
ICP/OES	Optima 3300 DV,	Perkin-Elmer	Iron content determination
MALDI-TOF	TofSpec 2E	Micromass, UK	Measuring the mass of lyophilized protein
Optical density measurement	SpectraMAX 250	Molecular Devices Corp.	Protein amount determination and OD ₆₀₀
Photometer		Eppendorf	Measuring the OD ₆₀₀
Semi-dry electrotransfer unit	Hoefer Semiphor TE77	Amersham Biosciences	Protein membrane transfer
Quattro-LC-MS	Hewlett-Packard HP 1100 coupled to Micromass Quattro II tandem quadrupole mass spectrometer.	Avondale, PA, USA and Waters, Micromass, Manchester, UK	Protein mass determination
Sonication		Bandelin	Disintegration of bacteria
Spectrofluorometer	V-550	Jasco	UV-absorption measurement

Appendix IV

Softwares

Software	Producer	Function
BioEdit 5.0.6.	Tom Hall Department of Microbiology, North Carolina State University, USA	Gene analysis
CLUSTAL X (1.8)	Julie Thompson Toby Gibson European Molecular Biology Laboratory Meyerhofstrasse 1 D 69117 Heidelberg Germany Des Higgins University of County Cork Cork Ireland	Multiple sequence alignment
Cn3D 4.1	The National Center for Biotechnology Information (NCBI), USA	Protein sequence alignment and structure simulation
CS ChemOffice 6.0	CambridgeSoft, USA	Calculating LogP and reaction standard free energy
DNASStar	DNASTAR. Inc	DNA analysis
Origin 5.0	Microcal oftware, Inc.	Data analysis
Plasmid processor	Department of Biotechnology, University of Kuopio, Finland	Plasmid drawing
TreeView 1.6.6	Roderic D. M. Page Division of Environmental and Evolutionary Biology Institute of Biomedical and Life Sciences University of Glasgow, Glasgow G12 8QQ, Scotland, UK	Phylogeny analysis
Xcalibur 1.1	Finnigan Corp.	Chromatograph and mass spectrometry analysis

Appendix V

Primers and probes

Primer	Sequence 5'→3'
27f	AGAGTTT <u>GATCCT</u> GGCTCAG
1492r	GGTTACCTTGTTACGACTT
Abam3	AGTCAGGATCCGTCGGCGCCGGGTATTAC
Afc1	GACATGACCGACACCAACATC
Ande1	CCCATCCATATGACCGACACCAACATC
Arc3	CCGTCGGCGCCGGGTATTAC
asmf1	ACIGA (C/T)ACIAA (C/T)AT (A/T/C)ACIAC
asmf2	ACIGCIGA (C/T)CCIGA (A/G)GT
asmr2	CGIAC (A/G)TTCCA (A/G)TG
sdu	I (A/C) (A/G/T/C)GAGG
asmr3	GTTCCAGTGCGCCTGCTTGC
asmr4	CGACCGGAGTGAGGAACTGC
asmf4	ACAGCTCGCCGATGGTCACA
asmf5	ATCACCACCCCGCCCTCAC
asmf6	ATGCAGGCGCTCGTTCGTC AAC
T7tem	GCTAGTTATTGCTCAGCGG
vegprom	GAGGTGGATGCAATGGCGAAG
asmr5	TTIAI (C/T)TC (A/G)TC (C/T)TGCCA (C/T)TG

Underlined sequences are introduced restriction sites. Degenerate nucleotides in primers are shown in brackets.

Appendix VI

Abbreviations

Afp	<u>a</u> mide <u>f</u> orming <u>p</u> rotein
ACP	<u>a</u> cy <u>c</u> arrier <u>p</u> rotein
ArCP	<u>a</u> ryl <u>c</u> arrier <u>p</u> rotein
APCI	<u>a</u> tmospheric <u>p</u> ressure <u>c</u> hemical <u>i</u> onization
BAW	<u>B</u> eet <u>A</u> rm <u>y</u> worm
BHI	<u>b</u> rain <u>h</u> eart <u>i</u> nfusion medium
BSA	<u>b</u> ovine <u>s</u> erum <u>a</u> lbumin
CTAB	cetyltrimethylammonium bromide and hexadecyltrimethylammonium bromide
DAPI	4',6-Diamidino-2-phenylindole
Dps	<u>D</u> NA-bind <u>p</u> rotein in <u>s</u> tarved cell
DTNB	5,5'- <u>d</u> i <u>t</u> hio-bis(2- <u>n</u> itro <u>b</u> enzoic acid)
ESI	<u>e</u> le <u>t</u> ro <u>s</u> pray <u>i</u> onization
FAS	<u>f</u> atty <u>a</u> cid <u>s</u> ynthase
FISH	Fluorescence <i>In Situ</i> Hybridization
FFF	<u>f</u> low <u>f</u> ield- <u>f</u> low <u>f</u> ractionation
FPLC	fast protein liquid chromatography
Flp	<u>f</u> erritin- <u>l</u> ike <u>p</u> rotein
Ftp	<u>f</u> ine <u>t</u> angled <u>p</u> ili protein
HPLC	<u>h</u> igh- <u>p</u> ressure <u>l</u> iquid <u>c</u> hromatography
ICP	<u>I</u> nductive <u>c</u> oupled <u>p</u> lasma
I.U.	<u>i</u> nternational <u>u</u> nit (enzyme)
K_{cat}	Turnover number
K_m	Michaelis- Menton constant
K_{eq}	equilibration constant
LogP	octanol/water partition coefficient
MALDI	<u>m</u> atrix- <u>a</u> ssisted <u>l</u> aser <u>d</u> esorption <u>i</u> onization
MALS	<u>m</u> ulti- <u>a</u> ngle <u>l</u> ight <u>s</u> cattering
Mr	molecular weight
MS	<u>m</u> ass <u>s</u> pectrometry

Appendix

NAP	<u>N</u> eutrophil- <u>a</u> ctivating <u>p</u> rotein
NRPS	<u>n</u> on- <u>r</u> ibosomal <u>p</u> eptide <u>s</u> ynthase
OD	<u>o</u> ptical <u>d</u> ensity
OES	<u>o</u> ptical <u>e</u> mission <u>s</u> pectroscopy
ORF	<u>o</u> pen <u>r</u> eading <u>f</u> rame
PAGE	<u>p</u> oly <u>a</u> crylamide <u>g</u> el <u>e</u> lectrophoresis
PAL	<u>p</u> eptidoglycan- <u>a</u> ssociated <u>l</u> ipoproteins
PBS	<u>p</u> hosphate- <u>b</u> uffered <u>s</u> aline
PE	<u>p</u> hosphatidylethanolamine
PEG	<u>p</u> oly <u>e</u> thylene <u>g</u> lycol
PFA	<u>p</u> ara <u>f</u> ormal <u>a</u> ldehyde
Pfr	Bacterial ferritin
PKS	<u>p</u> oly <u>k</u> etide <u>s</u> ynthase
PMSF	<u>p</u> henyl <u>m</u> ethyl <u>s</u> ulfonyl <u>f</u> luoride
P-pant	4'-phosphopantetheine
Quattro-MS	quadrupole-hexapole-quadrupole mass spectrometer
RACE	<u>r</u> apid <u>a</u> mpification of <u>c</u> DNA <u>e</u> nds
ROS	<u>r</u> eactive <u>o</u> xygen <u>s</u> pecies
SAP	Shrimp Alkaline Phosphatase
SD sequence	<u>S</u> hine- <u>D</u> algarno sequence
SDS	<u>s</u> odium <u>d</u> odecyl <u>s</u> ulphate
TES buffer	<i>N</i> -Tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
TOF	<u>t</u> ime- <u>o</u> f- <u>f</u> light mass spectrometry

Curriculum vitae

1. Personal Data

NAME: Liyan Ping
PLACE OF BIRTH: Zhuozhou, Hebei, China
DATE OF BIRTH: August 21, 1972
NATIONALITY: Chinese
MARITAL STATUS: Married (Mrs. Xiaoshu T. Ping)

2. Education and Research Experience:

Jan. 2002 - ongoing Max-Planck Institute for Chemical Ecology (Jena, Germany) under Prof. Dr. Wilhelm Boland, expected Ph.D. (Biology). Project entitled "The role of gut bacteria in herbivorous insect induced plant defence".

Jul. 1999 to Dec. 2001 Beijing Forestry University (Beijing, China), M.S. (Botany) under Prof. Yingbai Shen, "Tree volatile signals under biological stress", supported by National Natural Science Foundation of China.

Sep. 1998 to Apr. 1999 Chinese Academy of Agricultural Sciences, Department of Biotechnology (Beijing, China), technical assistant, worked on project "molecular marker of Lipoxgenase deletion in soybean".

Sep. 1997 to Feb. 1998 Peking University, College of life Sciences (Beijing, China), unregistered student.

Jul. 1997 to Jun. 1998 Chinese Academy of Sciences, Beijing Institute of Botany (Beijing, China), part-time research assistant, worked on 'The physiological function of Kiwi fruit pulp chloroplast' supported by National Natural Science Foundation of China.

Sep. 1994 to Jun. 1997 Shenyang Agricultural Universty (Shenyang, China).

Sep. 1989 to Jun. 1993 Hebei Forestry School (Shijiazhuang, China).

Selected Publications

- 1 Ping L., Boland W. 2004. Signals from the Underground: bacterial volatiles promote growth in Arabidopsis. *Trends in Plant Science*, 9 (6): 263-266.
- 2 Ping L., Spiteller D., Boland W. 2003. Plant defense-induced *N*-acyl glutamines from insect guts: structural diversity and microbe-assisted biosynthesis. *The BCPC International Congress-Crop Science and Technology*, 2: 1211-1216.
- 3 Ping L., Shen Y., Jin Y. 2001. Plant volatiles released in succession from artificially wounded ashleaf maple leaves. *Australian Journal of Plant Physiology*, 28 (6): 513-517.
- 4 Ping L., Shen Y., Jin Y., Hao J. 2001. Comparison of leaf volatiles induced by mechanical damage from diverse taxonomic tree species. *Acta Botanica Sinica*, 43 (3): 261-266.
- 5 Ping L., Shen Y. 2001. Plant wound-induced volatiles and their physiological functions. *Plant Physiology Communications*, 37 (2):166-172 (in Chinese).

I declare that the majority of the studies written in my dissertation has been done by myself independently, except those specified in the text.

Liyan Ping

20th Dec., 2004