University of Insubria

Department of Biotechnology and Life Sciences

PhD course in

Analysis, Management and Protection of the Biodiversity Resources

(XXVII cycle)

Heterologous expression, characterisation and applications of metagenome- and genome-sourced enzymes

Coordinator: Prof. Flavia Marinelli

Tutor: Prof. Flavia Marinelli

Dissertation of

Francesca Berini

2011-2014

INDEX

LITERATURE REVIEW

1. Metagenomics: a key to the unexplored biodiversity

"The abundance and diversity of life on Earth has come not from fossil horses or club mosses but from the flourishing of the oldest, most omnipresent life forms, the bacteria. For all intents and purposes the bacteria invented everything of importance: growth, metabolism and reproduction, swimming and chemical sensitivities, oxygen respiration and desiccation-resistant propagules. Some perfected predatory behavior and the kill. They are masters of efficiency and recycling of waste. They cover the mountaintops, the prairie, and the plains with their offspring. They swim with no thought of sleep. They fashion fuels like methane and ethanol from far less energetic forms of carbon such as CO_2. The prodigious bacteria have created sexual communication and gender, genetic recombination, and consortial living. Some thrive exposed to ferocious winds and blinding sunlight on open cliffs, others burrow into hard limestone rock and photosynthesise right through their chalky covers. As metal workers, bacteria have no peers: some precipitate gold and others mine iron; some manufacture metallic sheen of manganese and others work copper or etch glass."

Lynn Margulis

From: *Microbial diversity. Form and Function in Prokaryotes*. Edited by Oladele Ogunseitan, Blackwell Science Ltd. 2005.

1.1 The microbial biodiversity

The microbial world is immense, in both the physical and scientific senses [1]. Microorganisms include archea, bacteria, protozoans, and certain algae and fungi. The total number of microbial cells in the Earth's biosphere has been estimated to be 4-6 x 10^{30} [1-3]. Prokaryotic microorganisms – archea and bacteria – not only represent the largest proportion of individual organisms, comprising 10^6 to 10^8 separate genospecies, *i.e.* distinct taxonomic groups [2, 3], but they also make up most of the Earth's biomass [1]. Being the

oldest form of life (they live on this planet since more than three billions of years), prokaryotes have evolved and accumulated remarkable physiological and functional heterogeneity, thereby constituting the world's major reserve of genetic diversity. This unparalleled biodiversity arises primarily because of the wide variety of microbial ecosystems which are inhabited by bacteria and archea. Soil is traditionally considered the habitat that harbours the largest microbial diversity, a "hidden treasure" which could be a great source of natural products for several agricultural and biotechnological applications, with 4×10^7 and 2×10^9 prokaryotic cells contained in one gram of forest and cultivated soils, respectively [4, 5]. Based on the reassociation kinetics of DNA isolated from various soil samples, the number of distinct prokaryotic genomes has been estimated to range from 2000 to 18000 genomes per gram of soil. This number might be even an underestimate, since genomes representing rare and unrecovered species might have been excluded from these analyses [4]. Also marine environment contains an enormous pool of as yet largely underexploited microbial biodiversity: bacteria can achieve densities of up to 10^6 per millilitre of seawater [6]. Microorganisms are present not only in nutrient-rich environments as soils, lakes, oceans or inside other organisms, but also in the less-hospital habitats on Earth, such as hot springs, nearly saturated salt brines, acid mine waters at pHs near zero, ocean tranches with depth of up to 11000 m and pressures exceeding 100 MPa, deep-sea hydrothermal vents with temperatures as high as approximately 400 °C, as well as deep in Antarctic ice and kilometres below the Earth's surface [6].

Human life and activities depend on microorganisms, as they play fundamental roles in biogeochemical cycles for converting the key elements of life – carbon, nitrogen, oxygen and sulphur – into forms accessible to all other living things. Even more interestingly, the majority of the photosynthetic capacity of the planet does not depend on plants but on microbes. Microbial communities are closely associated with plants and animals making necessary nutrients, metals and vitamins available for the hosts. For humans, the billions of gut microbes assist us to digest food, break down toxins, and fight off pathogens. Humanity depends on microbes not only for nutritional and health reasons but also for cleaning up pollutants in the environment, such as oil and chemical spills, activities usually carried out not by individual microbes but by complex microbial communities [7]. Moreover, human civilisation has greatly improved by the development of numerous technologies that have their source in microbes, and microbial derived enzymes find applications in all major

industrial sectors. For instance, bacteria are used to synthesise a vast array of antibiotics, antitumor agents and immunosuppressants for clinical use, to produce biofuels, to enhance and protect agricultural crops directly or by production of biopesticides and antiparasite agents, and they are even used as markers for detection of diseases. Microbial enzymes find application in detergent production as well as for food and feed processing; but also pulp, paper, textile and leather sectors are fast growing market for the microbial derived compounds [7].

1.1.1 The Great Plate Count Anomaly

The microbial world is enormous also from the perspective of the current limitations of human knowledge. In 1898 a microbiologist, Heinrich Winterberg, for the first time described the discrepancy in the number of microorganisms between culturable bacteria on nutrient media and the total bacteria in nature counted by microscopy. Since then, microbial unculturability, the so-called "great plate count anomaly", has long been recognised in microbiology [8]. Unculturable microorganisms can be found in nearly every group within the bacteria and archea, and it has been estimated that only 0.1 to 1% of the prokaryotes are culturable by traditional cultivation and isolation methods and therefore accessible for biotechnology or basic research [4, 5, 8, 9]. There are several reasons for microbial unculturability under laboratory conditions: for instance, extremely high substrate concentrations, the lack of specific nutrients required for growth or in general the inability to recreate *in vitro* the complex ecological niche in which the microorganism lives. Some authors suggested that "unculturable" bacteria should be more specifically called "not-yet-culturable", as the ongoing development of techniques for isolation and culturing could in the future allow the successful cultivation of recalcitrant microbes, which may simply be in a physiological state that eludes our actual ability to culture them [5, 10]. Improved culture conditions include the use of nutrient-limited media as well as long incubation periods compatible with slow-growing bacteria. These culture conditions can be combined with high-throughput technologies such as microchip-based culturing or single-cell encapsulation and fluorescence-activated cell sorting (FACS); this method physically separates microorganisms in agar microdroplets, though maintaining molecular exchanges between the cells and their

environment, and facilitates the detection of previously uncultured microorganisms. Nevertheless, even when applying these methods, the ratio of uncultivated to cultivated bacteria remains high [5, 11].

To overcome the difficulties and limitations associated with cultivation approaches, several culture-independent methods have been developed, including phospholipid fatty acid analysis (PLFA) and numerous DNA- and RNA-based molecular approaches [5]. PCR methods based on the analysis of the "molecular clock" 16S rRNA gene (18S rRNA for eukaryotes), for example, provide extensive and valuable information about the taxa and species present in an environment. However, these data usually do not cover the complexity of prokaryotic diversity and provide only little if any information about the functional role of the different microbes within the community and the genetic information they contain. The 16S gene itself, in fact, accounts for approximately 0.05% of the average prokaryotic genome and cannot be used to get information about the physiology of the microorganism from which it was obtained [9, 12].

Therefore, alternative technologies, the so-called Meta-omics (metagenomics and the more recent metatranscriptomics, metaproteomics and metabolomics), have lately gained more and more success. These innovative methods utilise genomic, proteomic, metabolomic and trascriptomic toolsets to transcend cultivation limitation by studying the collective material of organisms from environmental samples. Hence, meta-omic technologies can enable the identification of novel natural products, new enzymatic activities and metabolic pathways as well as a better understanding of relationships between microorganisms [13].

1.2 Metagenomics

The term "Metagenomics" was first coined by Handelsman and co-workers in 1998 and could be defined as the analysis of the genetic complement of an entire habitat by direct extraction and subsequent cloning of DNA from an assemblage of microorganisms [14].

In the next subsection, the passages required for the construction and screening of metagenomic libraries are described along with their applications and future developments.

1.2.1 Metagenomic library construction

Metagenomic libraries have already been constructed from a broad range of environments to access the genetic potential of the microbial communities. These studies have included temperate soils, sediments, freshwaters, marine environments and the gut of animals and humans. Recently, also extreme environments such as the Arctic, glacial ice and soils, acidic and hypersaline environments, as well as solfataric hot springs and hyperthermal ponds have been addressed by metagenomics-based studies [2, 8].

Sample collection and DNA extraction

The first step in the construction of a metagenomic library is the extraction of the environmental DNA (eDNA) from the selected sample (Figure 1). Three major problems need to be taken into consideration: the DNA should be extracted from as broad a range of microorganisms as possible to be representative of the original microbial population; DNA shearing has to be avoided during the extraction procedure since high molecular weight DNA is required for suitable community analysis; thirdly, the DNA must be free from contaminating substances which interfere with downstream processing [15]. In other terms, the metagenomic DNA needs to be of sufficient quality with regard to purity, integrity, representativeness and length of the fragments, in order to be suitable for cloning into a relevant vector [8, 10]. For example, when the metagenomic library is constructed from a water environment, up to more than 1000 litres should be collected to obtain a number of cells sufficient to generate enough nucleic acids for downstream applications. On the contrary, the major problem encountered with soil metagenomic libraries is the co-extraction of humic and fulvic acids during DNA isolation, which must be removed before DNA could be further processed [11, 15].

Irrespective of the nature of the sample, DNA extraction can be achieved by two general strategies. The first one, which is the most commonly used and the fastest, is direct DNA extraction and consists of cell lysis directly within the sample matrix, followed by separation of the DNA from the matrix and cell debris. With the second strategy, indirect DNA extraction, cells are first removed from the matrix and subsequently lysed. Methods for cell lysis include chemical or enzymatic lysis and mechanical disruption (thermal shock, bead-mill

homogenisation, bead-beating, microwave heating, ultrasonication) [4, 16, 17]. Normally, 10-100-fold more DNA is obtained using direct lysis and extraction methods, in spite of a lower purity. Indirect extraction can be helpful when eukarya need to be excluded (bacterial and archeal cells could be separated from eukarya by using a density gradient) or when high DNA fragments are demanded [4, 11, 15]. It should be kept in mind that, since each DNA extraction method exhibits its own specific biases, a different picture of the microbial community can be obtained according to the used DNA extraction protocol [11].

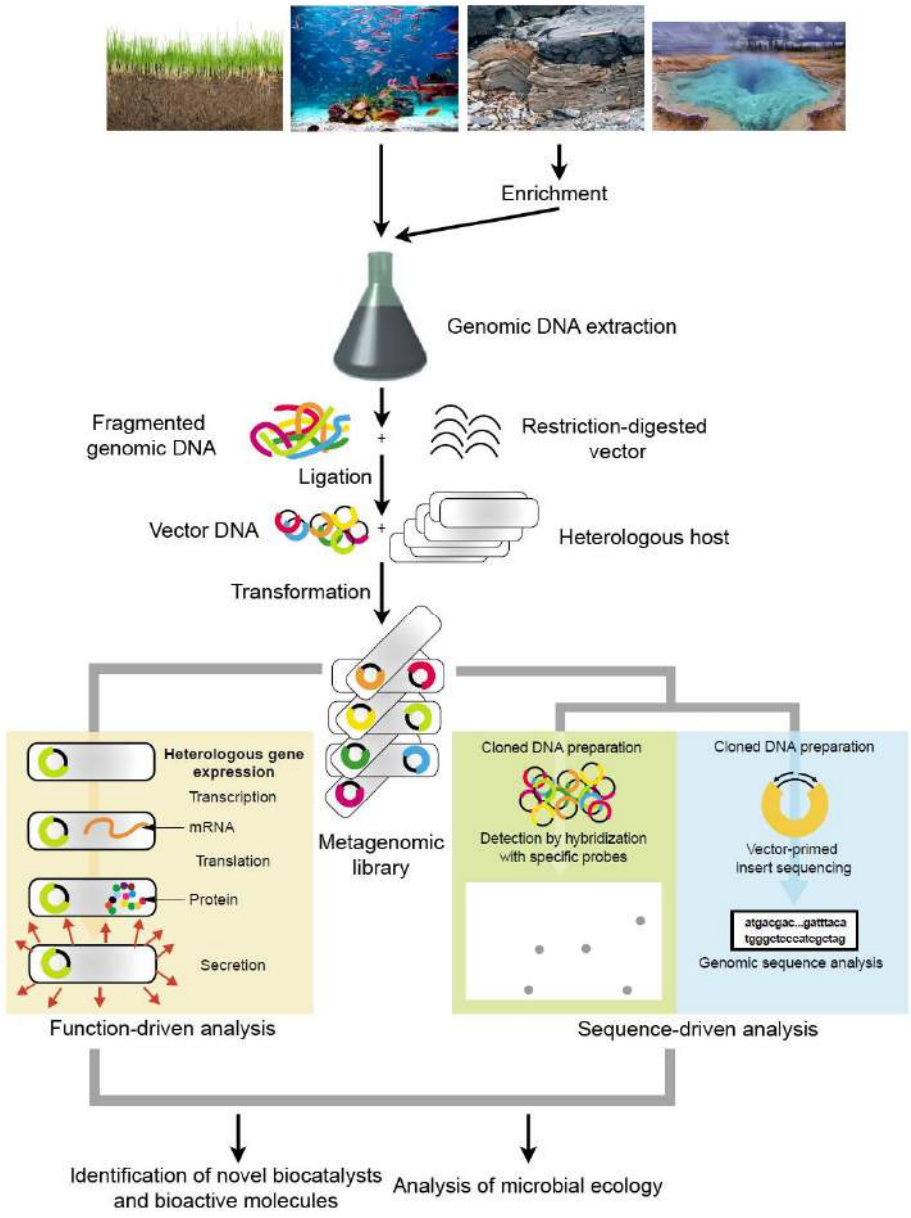


Figure 1. Construction and analysis of metagenomics libraries (modified from Schloss *et al*., 2003 [18]).

Enrichment strategies

A way to improve the quality of the environmental DNA and to enhance the screening hit rate, is the ecological enhancement, also called habitat biasing or targeted metagenomics. The microbial community is manipulated generally prior to the extraction of the metagenomic DNA in order to increase *in situ* the prevalence of target functions [3, 4, 8, 18]. Several methods have been developed for this purpose; the mostly used are the following:

- Microbial communities could be exposed to physical-, chemical- or nutritional-pressure to select microorganisms with the desired phenotypes. For example, using DNA isolated from enriched cultures grown on cellulose or chitin as their major carbon source increases from three- to four-fold the isolation of respectively cellulases and chitinases, compared with the isolates from libraries made directly from total eDNA [18].
- The prokaryotic community members can be separated by size-selective filtration from eukaryotic cells to produce relatively pure microbial DNA. Similarly, this approach is useful to separate multicellular consortia from individual cells [17, 18].
- Extracted DNA can be subjected to ultracentrifugation to enrich for high G+C content DNA, thus increasing the representation of certain genomes of particular interest (e.g. Actinomycetes and Acidobacteria) in the library [17, 18].
- Two methods for the selection of metabolically active organisms are 5-bromo-3-deoxyuridine (BrdU) labelling and stable-isotope probing (SIP), respectively based on providing BrdU- and ^{13}C-, ^{15}N-, ^{18}O- or ^{2}H-labelled substrates to bacteria. Only metabolically active bacteria will incorporate the labelled nucleotides into their DNA, which could be then isolated by immunocapture or density gradient centrifugation, thus excluding nucleic acid material from inactive or dead members of the microbial community [10, 17-19].

Library construction

After DNA extraction, the subsequent step for metagenomic library construction is DNA fragmentation by restriction enzyme digestion or mechanical shearing, followed by cloning into an appropriate host-vector system (Figure 1). Libraries can be classified into two groups with respect to average insert size: small-insert libraries in plasmid vectors (less than 15 kb) and large-insert libraries in cosmids (15-40 kb), fosmids (25-45 kb) or bacterial artificial chromosomes (BAC, more than 100 kb). The choice of the vector system depends on the

quality of the isolated DNA, the desired average insert size of the library, the required vector copy number, the host and the screening strategy that will be used. Small-insert libraries are useful for the isolation of single genes or small operons; large-insert libraries are more appropriate to recover complex pathways encoded by large gene clusters or large DNA fragments for the characterisation of genomes of uncultured soil microorganisms [4]. It has been estimated, for example, that more than 10^7 plasmid clones (5 kb inserts) or 10^6 BAC clones (100 kb inserts) are required to represent the genomes of all the different prokaryotic species present in one gram of soil. If the goal is to achieve substantial representation of the genomes of rare members of the soil community, the library should contain 10000 Gb of DNA (10^{11} BAC clones) [4].

In most metagenomic studies performed thus far, *Escherichia coli* has been used as the cloning host. Because of its status as the most well-known model host, in fact, there is ample knowledge about different useful gene expression strategies and an extended genetic toolkit is available for this microorganism. However, significant differences in the levels of expression can occur depending on the taxonomic groups present within the metagenomic DNA sample, a problem that needs to be considered especially when a functional screening (see "Library screening" below) is planned. It has been estimated that only about 30-40% of bacterial genes could be efficiently expressed in *E. coli*, a value dropping to 7% for high G+C DNA, indicating that *E. coli* is at best a suboptimal host for the heterologous expression of genes from many non-enteric bacteria. This might be due to a plethora of factors, such as codon usage differences, improper promoter recognition, lack of proper initiation factors, different preference for start codons, ribosomal entry, improper protein folding, absence of essential co-factors, accelerated enzymatic breakdown of the gene product, inclusion bodies formation, toxicity of the gene product, lack of essential post-translational processing and/or transport functions, or the inability of the host to secrete the gene expression product [8, 10, 20]. One way to enhance the possibility of a successful expression of genes may be to engineer *E. coli* expression machinery on the basis of the expected prevalence of genes from source hosts; for example, altering the host's transcription and translation systems for increasing the recognition of foreign ribosome binding sites (RBS) predicted to be prevalent in the metagenome, or co-expressing the proteins with a chaperone to promote protein folding [8]. An alternative is the development of other prokaryotic non-*E. coli* hosts. These include soil bacteria belonging to the genera *Agrobacterium, Caulobacter, Ralstonia,*

Sphingomonas, Burkholderia, Bacillus, Acidobacterium and *Verrucomicrobium.* Other alternative hosts, such as *Rhizobium, Pseudomonas* spp. and mutants of *Lysobacter enzymogenes* and *Pseudomonas fluorescens* have also been used [6, 8, 10]. Even few archeal genera (*Methanococcus, Pyrococcus, Sulfolobus* and *Thermococcus*) have been employed as alternative expression systems [2]. Factors to be considered in the choice of a suitable expression host include simplicity in handling, favourable growth, availability of genetic tools, and appropriate cellular machinery for protein and metabolite production and activities. Among the different proposed hosts, *Streptomyces* spp. appear to be one of the most promising for metagenome libraries construction, also because of the well-developed methods of genetic transfer from *E. coli* to *Streptomyces*. *S. lividans* proved to be a particularly useful host for functional screening of soil metagenomic libraries for novel polyketide synthase genes, as well as for the detection of a range of other novel metabolites [20].

In addition to single-host expression, also multi-host expression strategies could be exploited. This is due to the idea that a substantial part of the transformed genes cannot be successfully expressed in a single organism and that the use of multiple hosts either sequentially or in parallel offers great advantages. The employment of multiple hosts diversifies the available expression machinery and helps to overcome the effect of gene product toxicity and enzymatic breakdown, thus enhancing the chance of identifying bioactive molecules by matching eDNA-derived clusters with native host biochemistries [8]. To express genes from metagenomes in multiple hosts, shuttle vectors with broad-host-range are of use. A nice example of a metagenomic study in which broad-host-range vectors were employed was provided by Craig *et al.*, 2010: metagenomic libraries derived from soil were constructed in a IncP1-α broad-host-range cosmid vector using six selected proteobacterial host strains, *i.e. Agrobacterium tumefaciens, Burkholderia graminis, Caulobacter vibrioides, E. coli, Pseudomonas putida* and *Ralstonia metallidurans*. Remarkably, a high diversity of expression profiles between the different hosts was found, with little overlap [21].

Library screening

Screens of metagenomic libraries could be based either on metabolic activity (function-based or activity-based approach) or on nucleotide sequence (sequence-based or molecular

approach). The first one depends on the successful expression of target gene(s) in the metagenomic host; instead, molecular screening is based on the detection via hybridisation or PCR approaches of conserved DNA regions [10, 22].

Three major function-driven approaches have been used to recover novel biomolecules: (i) phenotypical detection, (ii) heterologous complementation, and (iii) induced gene expression [2, 8]. Phenotypical detection is the most commonly employed approach and, as the name itself suggests, is based on the identification of specific phenotypic traits associated with the activity of interest. These include particular colony pigmentation and morphology, degradation- and inhibition-halo formation, or the reaction of an added substance (like a chromophore or a chemical dye) with the expressed gene product. Heterologous complementation relies on selection of clones that have acquired capability to grow under selective conditions, as in the presence of antibiotics. Substrate-induced gene expression screening (SIGEX) was first introduced in 2005 by Uchiyama and co-workers [23] and is based on the use of an operon-trap GFP (green fluorescent protein) expression vector, where the metagenomic DNA is cloned upstream of the *gfp* gene. Positive clones, co-expressing the GFP upon substrate-induced expression, could be isolated by fluorescence-assisted cell sorting (FACS). Similarly, the metabolite-regulated expression system (METREX) aims at detecting biologically active small molecules by an intracellular biosensor system composed of the *gfp* gene under the control of a quorum-sensing promoter. When a threshold concentration of the signal molecule encoded by the metagenomic DNA fragment is exceeded, GFP is produced and positive clones are identified by fluorescence microscopy. With PIGEX (production-induced gene expression), enzymatic activities are similarly detected by the expression of *gfp*, which in this case is triggered by product formation. As sequence information is not required, functional-based methods are the only strategy with the potential to identify new classes of genes encoding either known or new functions. Another advantage is the possibility to recover full-length genes and therefore functional gene products. However, the major drawback of this approach is its dependence on the eDNA expression in a foreign host, which, as stated previously, is not always feasible. Many genes are low, if not entirely expressed or the translated gene products are inactive in the host strain, thus limiting the detection frequencies of such approach. This limitation is particularly relevant when the expression of an entire biosynthetic gene cluster, which requires the coordinated production of multiple proteins, is needed [22]. Nevertheless, function-based

approaches have uncovered a wide range of biocatalysts, like genes coding for degradative enzymes, as well as antibiotic resistance mechanisms and even new antibiotics [4].

On the other hand, sequence-based approaches have the advantage of being expression-independent. However, being based on the identification of conserved nucleotide sequences, with this type of screening it is possible to identify only members of already known gene families. Additionally, the detection of sequences of interest does not guarantee a functional and efficient expression of the target gene. Molecular screening is generally achieved by PCR techniques or by hybridisation studies with primers and probes specific for conserved regions of the genes being targeted, which for enzymes are usually the catalytic domains [6]. This approach has been used to identify phylogenetic anchors such as 16S rRNA genes and genes encoding enzymes with highly conserved domains, like polyketide synthases, glucuronic acid reductase and nitrile hydratases [4].

Another approach for the analysis and profiling of metagenomic libraries is the use of high-throughput microarrays, which represent an effective method for rapidly screen large numbers of clones [4, 9, 19]. However, microarray technology shows a 100 to 10000-fold lower sensitivity than PCR for gene detection and thus this difference might prevent the analysis of sequences from low-abundance microorganisms [4, 5]. Moreover, the recent development of sequencing technologies, like the next-generation sequencing (NGS), made it possible to sequence entire metagenomic libraries with relatively-low costs. The thereby collected sequence information can be used to recover full-length genes and to gain comprehensive insights into the flanking regions of the target genes. One important consequence of the large-scale metagenome sequencing projects is the development of new open infrastructures for data storage and analysis, as CAMERA (community cyber infrastructure for advanced microbial ecology research & analysis), MG-RAST (metagenomic rapid annotation using subsystem technology) and IMG/M (integrated microbial genomes) [2].

1.2.2 Applications of metagenomics

From its outset, metagenome-based approaches have led to the accumulation of an extraordinary number of DNA sequences. This genomic information can be exploited for

biotechnological and pharmaceutical applications and to increase our knowledge of microbial ecology and physiology.

Using the metagenome sequences to fully understand how complex microbial communities function and how microbes interact within these niches represents a major challenge for microbiologists today. In the last decade, metagenomic approaches have provided comprehensive data on microbial diversity and population dynamics in a large variety of ecosystems such as soil, global ocean and extreme environments. More recently, these technologies have been applied to medical and forensic investigations, for example for the identification of new viruses from human samples and for the reconstruction of the human gut and skin microbiome, as well as for the analysis of extinct species and the study of microbial communities in food and animal feed [2, 9, 17].

Metagenomics is currently thought to be one of the most likely technologies to provide new biotechnological products and processes. Since their introduction, in fact, metagenomic approaches have led to the discovery and characterisation of a significant number of novel genes encoding for biocatalysts or molecules with high potential for use in pharmaceutical products or production processes. It is conceivable that metagenomics together with protein engineering and *in vitro* evolution technologies might be used to find suitable natural enzymes that can serve as a backbone to produce ideal biocatalysts, *i.e.* improved tailored enzymes that optimally fit specific process requirements [24]. Till now, research in metagenome-derived DNA libraries has been focused on several classes of enzymes (Table 1), such as lipolytic enzymes (lipases and esterases), oxidoreductases and dehydrogenases, polysaccharide degrading/modifying enzymes (amylases, cellulases, chitinases, xylanases, agarases) as well as proteases, nitrilases, amidases, DNA polymerases and enzymes involved in vitamin biosynthesis, such as biotin and vitamin C [9, 15].

Target	Origin	Metagenomic library type	Number of screened clones	Screening approach
Esterase/lipase	Soil	Plasmid	286000	Function-based
Lipase	Sediment	Fosmid	7000	Function-based
Esterase	Sediment	Plasmid	60000	Function-based
Alcohol oxidoreductase	Soil/enrichment	Plasmid	400000	Function-based
Amidase	Soil/enrichment	Plasmid	193000	Function-based
Amylase	Soil	BAC	3648	Function-based
Biotin production	Soil/enrichment	Cosmid	50000	Function-based
Protease	Soil	Plasmid	100000	Function-based
Cellulase	Sediment/enrichment	λ phage	310000	Function-based
Cellulase	Aquatic community	Cosmid	3744	Function-based
Chitinase	Seawater	λ phage	825000	Function-based
Dehydratase	Soil-sediment/enrichment	Plasmid	560000	Function-based
Alkaline hydroxylanase	Ocean	Cosmid		Heterologous complementation
Serine protease inhibitor	Seawater	Plasmid	50000	Sequence-based

Table 1. Examples of industrially relevant enzymes and biocatalysts from metagenomic libraries [24].

In addition to novel enzymes, also the isolation of genes encoding novel therapeutic molecules and pharmacologically active secondary metabolites is a valuable area of research. In this context, the genes of interest are often type I and type II polyketide synthases (β-ketoacyl synthetases). These are key genes involved in the synthesis of polyketide antibiotics and can be used as sequence tag tools for the identification of large biosynthetic gene clusters [9, 12, 15]. A range of already-known and novel antibiotics have been detected in metagenomic libraries, for instance indirubin, used in the treatment of leukemia, turbomycin, palmitoylputrescine and *N*-acyltyrosine, but also glycopeptide antibiotics, lantibiotics and terragines [10, 15, 18, 22]. Examples of microbial bioactivities identified from soil-derived metagenomic libraries are summarised in Table 2. Furthermore, with metagenomic approaches lots of information has been gained about the diversity of natural antibiotic resistance mechanisms [4].

Origin	Bioactivity
Forest soil	Indirubin and indigo production
Soil	Deoxyviolacein and violacein production
Soil	Long-chain fatty acid enol ester production
Desert sand soil	Indolotryptoline antiproliferative agents, dihydroxyindolocarbazole anticancer/antibiotics
Arable field soil	Fatty dienic alcohol isomers production
Desert sand soil	Fluostatins production
Agricultural soil	Turbomycin A and B production
Desert sand soil	Erdacin production
Rice paddy soil	Coproporphyrin III production
Garden soil	Indigoidine production

Table 2. Examples of microbial bioactivities identified from bioprospecting metagenomics from various soils [12].

1.2.3 Future perspective of metagenomics

The rapid advancement and increasing affordability of next generation sequencing technologies have given birth to the so-called "shotgun metagenomics", *i.e.* the direct sequencing of isolated eDNA, bypassing the laborious steps of library construction and screening. Through the development of a number of sequencing platforms, like Roche 454, Illumina/Solexa, PacBio and Ion Torrent, shotgun metagenomics has become the method of choice for varied applications, being used for instance in the human microbiome project or for the study of the complex microbial communities associated with coral reefs. The popularity of this approach is expected to rise significantly as advancing sequencing technology and better bioinformatics analysis pipelines will improve effectiveness and throughput [13].

Another derivative of the traditional metagenomic approaches is the single cell genomics, which is designed to assess the genomes of individual microbial cells isolated from environmental samples. This technology is dependent on multiple displacement amplification (MDA), which can produce micrograms of DNA necessary for sequencing applications from the few femtograms present in an individual cell. Single cell genomics enables to directly link genes and biosynthetic clusters to any taxonomic information uncovered from the genome, potentially leading to the selection of new and targeted hosts for their own heterologous expression; similarly, the identification of metabolic genes could

shed light on the conditions or substrates needed to successfully culture the target unculturable microorganism in the laboratory [13, 25].

An exciting extension of metagenomics is the high-throughput analysis of the mobilome or mobile metagenome, *i.e.* the set of genes present on mobile genetic elements, such as plasmids, transposons, insertion sequences and integrons. Analysis of mobile metagenomics is expected to provide valuable knowledge of horizontal gene transfer events and their implication on microbial ecology and evolution [16].

Besides these metagenomics-derived applications, other Meta-omics technologies are rapidly advancing. Functionality of soil microorganisms and their catalytic potential can be characterised by the analysis of collective proteins (metaproteomics) directly isolated from the environment. Metatranscriptomics has become a useful tool to assess the actual metabolic activity of a microbial community, by differentiating between expressed and non-expressed genes [2]. Finally, metabolomics offers powerful toolsets that enable the identification and characterisation of metabolites of biological significance [13]. The combination of the different Meta-omics technologies offers significant promise to advance the measurement and prediction of the *in situ* microbial responses, activities, and productivity of microbial consortia. In addition, analyses of the thereby-generated comprehensive data sets have an unprecedented potential to shed light on ecosystem functions and evolutionary processes, and increase the possibility to discover new molecules and proteins with the desired characteristics.

1.3 The MetaExplore project

The implementation of sustainable technologies and processes has been one of the thematic priorities of the European Union's Seventh Framework Programme for Research (FP7). One of the selected European projects included in the FP7 was the 5-year MetaExplore project, run from May 2009 to May 2014 (http://www.rug.nl/research/metaexplore/). MetaExplore involved eighteen research groups of eleven different European countries, with the aim to develop and apply advanced molecular tools for the cloning and sequencing of the metagenomes of microbial communities of selected soil and aquatic habitats, followed by

educated activity- and sequence-based screenings for target enzymatic activities, their analysis and engineering. European academia as well as the biotechnology industry involved in food and waste management, in fact, is continuously in search of novel enzymes able to degrade recalcitrant natural polymers such as chitins and lignins, and/or polluting man-made compounds such as halogenated aliphatic and aromatic compounds. A first class of desired activities for the MetaExplore project consisted of new chitinases and chitin deacetylases, key enzymes in the food industry. For instance, the downstream processing of exoskeletal waste (the carapace) in the industrialisation of foodstuff can be greatly helped by such novel enzymes with high specific activities. Another class concerned the activities related to various enzymatic reactions that lead to the decomposition of lignin, a polydisperse phenolic macromolecule and one of the major compounds in lignocellulosic biomass. Degradation of lignin proceeds via oxidation reactions catalysed by ligninases (laccases, peroxidases and H_2O_2 producing enzymes), which are primarily produced by white-rot fungi, but also by filamentous bacteria like *Streptomyces* spp. Potential applications of ligninolytic enzymes are related to the removal of lignin in pulp processing or to improve total hydrolysis of lignocellulosic biomass for energy and chemical production. The third enzymatic class of interest was the one of dehalogenases and halogenases (haloalkane and haloalcohol dehalogenases, ammonia lyases and epoxide hydrolases), which can find application for environmental sanitation and for production of chiral pharmaceuticals and fine chemicals.

In other terms, the final goal of this European project was the unlocking and mining of the genetic potential of microbial environmental metagenomes, with a particular attention also to the mobilome. To achieve this ultimate goal, the members of the MetaExplore consortium focused on the development of modern technologies related to every step of the metagenomic analysis, from the identification of efficient functional and sequence-based screening methodologies, to the development of alternative host/vector systems, protein engineering and bioinformatics tools.

In this context, the Laboratory of Microbial Biotechnology of University of Insubria collaborated at the MetaExplore project in the validation of functional screening methods for the detection of chitinases, chitin modifying enzymes and ligninases. Another task, object of this dissertation, regarded the expression (in different heterologous systems) of metagenome-sourced selected enzymes and their functional/biochemical/structural characterisation, followed by their scale-up production at industrial level, in collaboration

with the University of Insubria's subcontractor Actygea (http://www.actygea.com/). This dissertation mostly focusses on the class of enzymes able to degrade chitin.

2. Chitin and chitin degrading enzymes

2.1 Chitin, chitosan and chitooligosaccharides

Chitin, after cellulose, is the second most abundant natural biopolymer on Earth, widely distributed both in aquatic and terrestrial ecosystems. The annual chitin production in aquatic habitats has been estimated to range from 2.8 x 10^7 Mg/yr for freshwaters to 1.3 x 10^9 Mg/yr for seas and oceans. Although no reliable estimates exist for annual chitin production in terrestrial ecosystems, this contribution seems negligible in comparison with the marine one [26].

Chitin was discovered in 1811 in mushrooms by H. Braconnot, who called it "Fungine", even if a first description of a "material particularly resistant to usual chemicals" by the English scientist A. Hachett dated back to 1799. In 1823, A. Odier found the same material in insects and plants and named it "chitine", from the greek word χιτών meaning tunic/envelope [27].

The most important derivative of chitin is chitosan, obtained by partial deacetylation of chitin under alkaline conditions or by enzymatic hydrolysis in the presence of chitin deacetylases (Figure 2). Chitosans with degrees of polymerisation <20 and an average molecular weight less than 3900 kDa are called chitosan oligomers, chitooligomers or chitooligosaccharides (COS). COS are generated by depolymerisation of chitin or chitosan using acid hydrolysis, hydrolysis by physical methods or enzymatic degradation with chitosanases. Chitin and its derivatives have become of great interest as new functional biomaterials of high potential in various fields, from industry to medicine. It has been estimated that by 2015 the market for chitin and COS can reach up to 63 billion US$, while that for chitosan up to 21 billion US$ [28, 29].

2.1.1 Chemical structure

Chitin is an unbranched polysaccharide composed of β(1→4)-linked *N*-acetyl-2-amino-2-deoxy-β-D-glucopyranose (*N*-acetylglucosamine, GlcNAc) residues, while chitosan, the

principal derivative of chitin, is a heteropolymer of two repeating units, *i.e. N*-acetyl-2-amino-2-deoxy-β-D-glucopyranose (GlcNAc) and 2-amino-2-deoxy-β-D-glucopyranose (glucosamine) (Figure 2). Chitin and chitosan are effectively the same macromolecular entity, varying only in the fraction of acetylated repeating units [28].

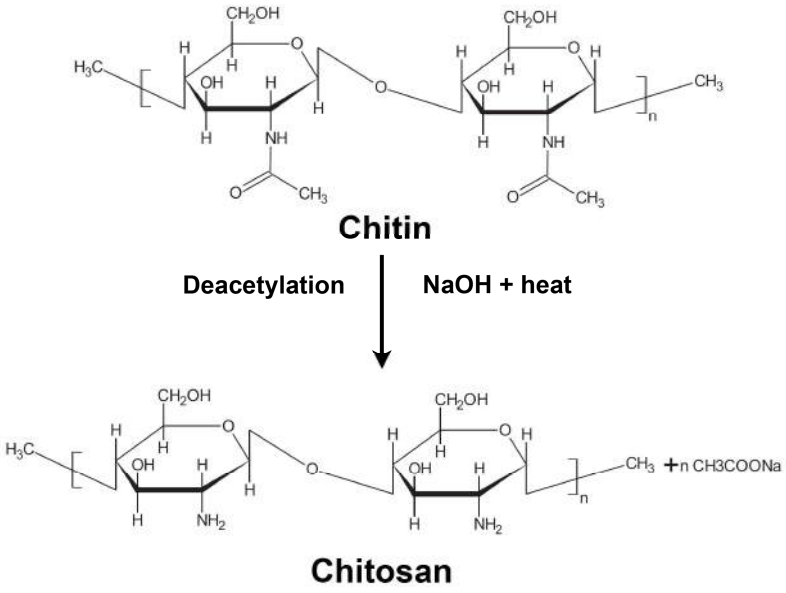


Figure 2. Deacetylation of chitin to chitosan (modified from Shukla *et al*., 2013 [28]).

Chitin producer organisms synthesise chitin according to a common pathway that ends with the polymerisation of GlcNAc from the activated precursor UPD-GlcNAc. The synthetic pathway includes the action of chitin synthases that accept the substrate UPD-GlcNAc and feed nascent chitin into the extracellular matrix. The chitin molecules, after their synthesis, arrange in a highly ordered biopolymer, showing rigid crystalline structure through inter- and intra-molecular hydrogen bonding, which varies considerably according to the origin of the polymers. In particular, three different allomorphic forms of chitin (and chitosan) are present in nature, differing in terms of polymer chain structure and crystallinity: α-, β- and γ-chitin. The α-form is widely distributed among crustaceans and insects, and is characterised by a two-chain anti-parallel structure, that gives rise to strong hydrogen bonding thus making it more stable. The β-form, quite rare and mainly obtained from molluscs such as squids and some diatoms, is characterised by a loose-packing parallel chain fashion with weak intermolecular interactions and higher solubility and swelling than α-form. The γ-chitin

is characterised by a mixture of anti-parallel and parallel chains and can be found in the cocoons of insects. Conversion from the β-form to the α-form is possible, but not the reverse; γ-chitin can be converted to α-chitin by treatment with lithium thiocyanate [27, 30].

2.1.2 Distribution in nature

As previously stated, chitin is widely distributed in nature, particularly as a structural polysaccharide in fungal cell walls, the exoskeleton of arthropods and the outer shell of crustaceans, but it is found also in certain algae, molluscs, yeasts and nematodes. In arthropods (insects, crustaceans, arachnids, and myriapods), chitin is found not only in the exoskeleton, but also in the tendons and in the linings of their respiratory, excretory and digestive systems, where it is usually present in complexes with other polysaccharides and proteins. Moreover, it is part of the reflective material (iridophores) both in the epidermis and the eyes of arthropods and cephalopods (phylum *Mollusca*). In filamentous fungi and basidiomycetes, chitin comprises 16 to 44% of the dry weight of the organism; while 20 to 58% of the total weight of shellfish, such as shrimps, crabs and krill is composed by this polysaccharide. In yeast, the amount of chitin in the cell wall is much lower, but bud scars have been shown to be largely composed of chitin [31]. Also few vertebrates contain chitin: for example, the epidermal cuticle of the fish *Paralipophrys trigloides* was found to be chitinous [27].

2.2 Chitin degrading enzymes

2.2.1 Occurrence of chitinases

Chitinolytic enzymes are ubiquitous proteins widely distributed among all kingdoms of life, being produced by bacteria, fungi, insects, viruses, plants and animals for different purposes including nutrition, morphogenesis, pathogenesis, parasitism and defence. Many of these organisms possess several genes that encode chitinolytic enzymes, usually acting in synergetic or successive manner to degrade chitin. The occurrence of multiple chitinolytic

enzymes is thought to reflect the flexibility of the organisms to deal with variability in chitin structures and different degrees of acetylation. For example, most filamentous fungi have 10 to 20 different chitinolytic genes, while in mycoparasitic species the number of such genes may reach 30 or even more [31-33].

Possession of chitinases is taxonomically widespread among bacteria, with chitinolytic representatives being found, among others, in the phyla *Actinobacteria, Firmicutes, Acidobacteria* and *Proteobacteria*. In these microorganisms, chitinases play roles not only in the decomposition of natural chitin resources, but also in antagonistic interaction with fungi and in parasitism [26]. Similarly, chitinolytic enzymes in fungi are thought to have autolytic, nutritional and morphogenetic roles, while in viruses are involved in pathogenesis [31, 33]. Further details about the chitinolytic systems of model bacteria and fungi can be found in subsection 2.2.4. Insect and crustacean chitinases are important during various stages of metamorphosis, especially in chitin turnover during molting. Their expression is tightly controlled by hormones, since premature exposure can lead to growth inhibition and mortality. Chitinases are also produced by the venom and salivary glands of some insect species, probably to facilitate the degradation of the host cuticle and therefore the ingression or egression of parasitoid larvae through such barrier [31, 32]. In plants, chitinolytic enzymes show antifungal, antibacterial, insecticidal, nematicidal and antiviral effects, and participate in the systemic acquired resistance against a broad range of pathogens. Additionally, plant chitinases serve other physiological functions that may not be directly related to their hydrolytic activity: for example, some enzymes show ice structuring activity and provide cold or freeze tolerance for the organism, others may counteract oxidative stress and act as storage proteins through their ability to bind metals [32]. Finally, mammalians chitinases are supposed to be involved in defence mechanisms against chitin-containing pathogens, chitin digestion (especially in insectivorous mammals), tissue remodelling, fertilisation and innate immunity [32].

2.2.2 Classification of chitinolytic enzymes

Chitinases have wide ranges of molecular weights (20-115 kDa), optimal temperatures (18-90 °C), optimal pH (2.0-10.5) and pI values (3.5-8.8) (http://www.brenda-enzymes.org/).

According to the Nomenclature Committee of International Union of Biochemistry and Molecular Biology (IUBMB), chitinolytic enzymes can be classified into two major categories, based on the catalysed reaction [31, 34] (see also Figure 3):

- Endochitinases (EC 3.2.1.14; (1→4)-2-acetamido-2-deoxy-β-D-glucan glycanohydrolases) cleave chitin randomly at internal sites, generating soluble and low molecular mass multimers of GlcNAc (such as chitodextrins, chitotetraose, chitotriose and diacetylchitobiose).
- Exochitinases (EC 3.2.1.52; β-*N*-acetyl-D-hexosaminide *N*-acetyl-hexosamino-hydrolases) can be further divided into two subcategories: chitobiosidases (EC 3.2.1.29), which catalyse the progressive release of diacetylchitobiose starting at the non-reducing end of chitin microfibril; and β-(1,4) *N*-acetyl glucosaminidases (EC 3.2.1.30), which cleave the oligomeric products of endochitinases and chitobiosidases, generating monomers of GlcNAc.

An alternative pathway for chitin degradation involves the deacetylation of chitin to chitosan by chitin deacetylases (EC 3.5.1.14; chitin amidohydrolases), which is finally converted to glucosamine residues by the action of chitosanases (EC 3.2.1.132; chitosan *N*-acetylglucosaminohydrolases).

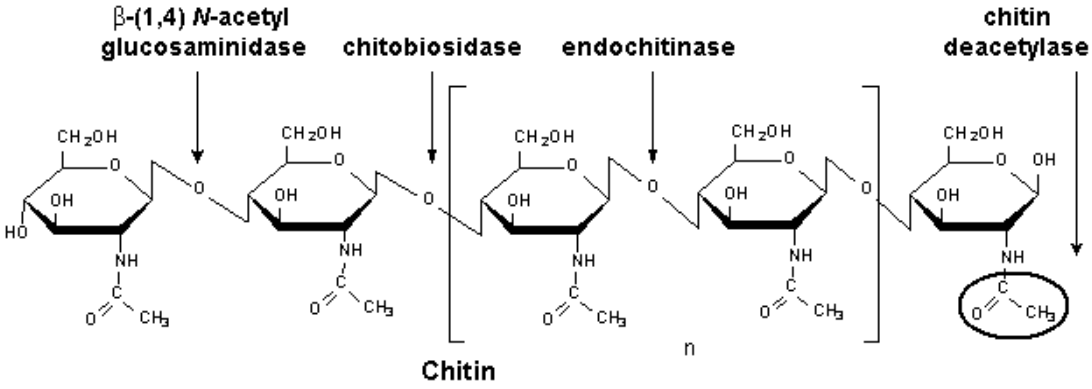


Figure 3. The enzymes involved in chitin degradation and modification, according to the IUBMB nomenclature.

Another classification of chitinolytic enzymes, firstly proposed by Henrissat in 1991 [35] and based on amino acid sequence similarity, groups chitinases into families 18, 19 and 20 of glycosyl hydrolases. Family GH18 contains endochitinases and chitobiosidases from bacteria, fungi, viruses, animals, insects and some plants, while family GH19 consists of chitinolytic

enzymes from plants and few bacterial groups (some actinomycetes, green non-sulphur and purple bacteria). The chitinases of the two families do not share amino acid sequence similarity, have completely different 3-D structures and molecular mechanisms (see subsection 2.2.3 for further details), and are therefore likely to have evolved from different ancestors. Finally, family GH20 includes β-(1,4) *N*-acetyl hexosaminidases and β-(1,4) *N*-acetyl glucosaminidases from streptomycetes and humans [31, 36].

Bacterial chitinases, mainly occurring in the GH18 family, can be further divided into three major subcategories, named A (the most abundant in the environment), B and C, according to the amino acid sequences of their catalytic domains, the modular structure and enzymatic activities (endo vs. exo) [31-33]. Plant chitinases are classified into seven different classes (Class I to VII) based on several parameters like N-terminal sequences, subcellular localisation of the enzyme, isoelectric point, signal peptide and inducers. Fungal chitinases, all belonging to family GH18, can be further divided into three subgroups on the basis of sequence comparison, namely A, B and C (not to be confused with bacterial subfamilies) [32]. All known insect chitinases belong to the family GH18 and can be assigned to eight distinct groups denoted by Roman numerals I-VIII [32].

2.2.3 Structure, catalytic mechanisms and inhibitors

Chitinolytic enzymes belonging to the same family show conserved features. For example, sequence alignments of family GH18 catalytic domains revealed the presence of two highly conserved regions, SxGG and DxxDxDxE, corresponding to the substrate-binding site and the catalytic domain, respectively. Similarly, family GH19 members are all characterised by the highly conserved motifs [FHY]GRG[AP]xQ[IL][ST][FHYW][HN][FY]NY and $L(x)_9LV(x)_{12}W[FY]W$, forming a substrate binding region [36].

Another common feature to most chitinases is their multi-domain structure. Next to the catalytic domain, in fact, other auxiliary regions, arranged in different order, could be found [26, 34]:

- N-terminal signal peptide, which mediates the enzyme secretion and is cleaved off by signal peptidases after transportation across the membrane;

- one or more carbohydrate-binding modules (CBMs), such as chitin binding domains (CBDs) and chitin insertion domains (CIDs), which increase the affinity toward chitin and, presumably, facilitate enzyme movements along the chitin chain during processive action and stimulate decrystallisation of the substrate;
- fibronectin type III (FnIII) domain, involved in binding to insoluble substrates like colloidal chitin and in the exo-hydrolytic mechanism;
- serine/threonine-rich linkers, usually post-translationally glycosylated with sugar chains and probably necessary for the secretion and maintenance of protein stability.

Crystallographic analyses of plant, bacterial and fungal chitinases in the past years have thrown light on the three dimensional structure of these hydrolytic enzymes. In particular, all GH18 and GH20 proteins have an eight-stranded beta/alpha-barrel ($(\alpha/\beta)_8$ or TIM barrel) fold [33, 34]. On the other hand, members of family GH19 have a bilobal structure with a high α-helical content [36].

Like other glycosyl hydrolases, chitinolytic enzymes generally catalyse the depolymerisation of their substrate through one of two pathways known as single- or double-displacement mechanisms (Figure 4). In both pathways, two distinct catalytic groups are involved. One of these is a carboxyl group that acts as a proton donor and is usually provided by a conserved glutamate residue at the active site of the enzyme, although in some cases an aspartate residue may fulfil this role. The second catalytic group may act either as a base (in the single-displacement mechanism) or as a nucleophile (in the double-displacement mechanism). Since single-displacement mechanism, typical of family GH19 chitinases, results in the inversion of the anomeric configuration of the hydrolysed GlcNAc residue, it is also known as the inverting mechanism. On the other hand, in the double-displacement mechanism of GH18 chitinases, also referred to as the retaining mechanism, the anomeric configuration is retained [31-33].

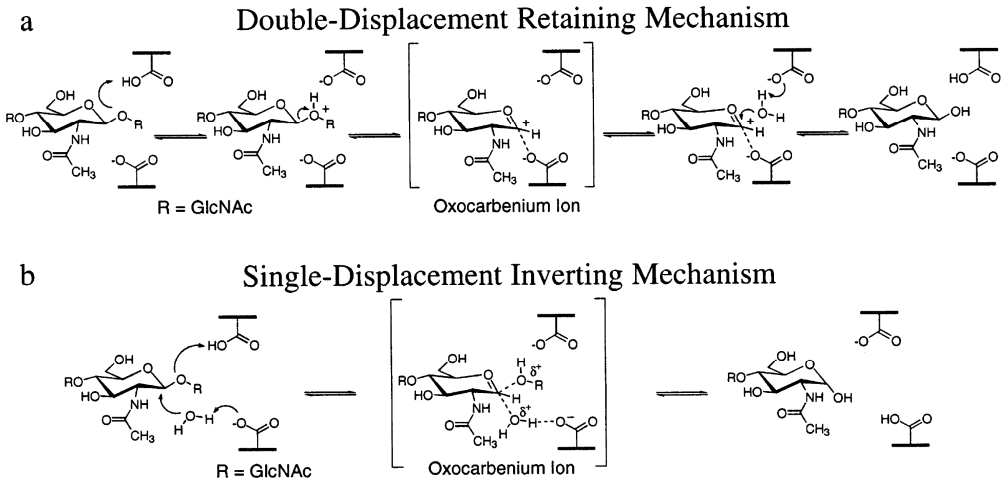


Figure 4. a) The double-displacement hydrolysis mechanism proposed for family GH18 chitinases. Protonation of a GlcNAc residue in a boat conformation leads to an oxazoline intermediate, which may be hydrolysed to form a product with retention of the anomeric configuration. **b)** The single-displacement hydrolysis mechanism proposed for family GH19 chitinases. Two acidic residues are required in the active site, and the hydrolysis product shows inversion of the anomeric configuration [31].

In addition to general enzyme inhibitors, such as organic compounds and oxidising/reducing agents, a number of reports are available on natural chitinase inhibitors. Allosamidin, for instance, is a pseudotrisaccharide antibiotic produced by *Streptomyces* spp. with a structure similar to GlcNAc, which specifically inhibits of GH18 chitinases from insects, yeast, fungi, and human serum. Psammaplin A, a brominated tyrosine-derived compound, was found to be a noncompetitive inhibitor of GH18 chitinases, binding near the active site. Other inhibitors include the cyclic pentapeptides argifin and argadin, isolated from *Gliocladium* sp. and *Clonostachys* sp. respectively, CI-4 from *Pseudomonas* sp. and C2-caffein, a low molecular weight compound composed of two linked caffeine moieties [31, 33].

2.2.4 The microbial chitinolytic system

Most of the natural degradation of chitin has microbial origin, and bacteria and fungi have developed systems for the depolymerisation, transport and metabolism of chitin and chitooligosaccharides [37, 38]. Owing to the structural complexity of the substrate, the

complete enzymatic hydrolysis of chitin is performed by a complex chitinolytic system, whose action is known to be synergistic and consecutive. The occurrence of multiple genes in a single organism may be the result of gene duplication or acquisition of genes from other organisms via lateral gene transfer [38]. Chitin turnover is highly regulated and the involved hydrolytic enzymes are usually induced by the products of chitin hydrolysis (in particular by GlcNAc) or by soluble chitin oligomers $(GlcNAc)_{2-6}$. If other more readily available growth substrates such as glucose are present, chitinolytic gene expression is usually repressed [31, 38].

In general, microbial chitin degradation starts with the secretion of chitin depolymerases (EC 3.2.1.14) that release GlcNAc, chitobiose and COS from the polymer. Then, in Gram-negative bacteria like *Serratia marcescens* and *Vibrio* spp., two of the most intensively studied chinolytic microorganisms, these compounds enter into the periplasm, where chitodextrinases (EC 3.2.1.14) and *N*-acetylglucosaminidases (EC 3.2.1.52) produce GlcNAc and, to a lesser extent, chitobiose. Finally, when transported into the cytoplasm, GlcNAc and chitobiose are metabolised or modified for cell wall biogenesis. PTS (phosphoenolpyruvate : glucose phosphotransferase system) transporters are thought to be responsible for the main GlcNAc uptake, but also other GlcNAc transporters as well as transporters with a broader substrate range (including sugar monomers like glucose, glucosamine, fructose and mannose) have been described [38].

In particular, when chitin is the main carbon source, *S. marcescens* produces four family GH18 chitinases (called ChiA, ChiB, ChiC1 and ChiC2, the last one derived from a post-translational modification of ChiC1), all of which are released into the surrounding medium. It has been suggested that ChiA and ChiB are exochitinases acting processively from the opposite sides of the chitin chain, while ChiC1 and ChiC2 are endochitinases. Their combination determines synergistic effects on chitin degradation [38, 39].

In *Vibrio furnisii* and *Vibrio cholerae*, the typical degraders of chitin in marine environments, dozens of enzymes are likely to be involved in chitin degradation. These include several extracellular chitinases, at least two specific hydrolases in the periplasmic space producing GlcNAc and chitobiose, and six cytoplasmic enzymes that convert these sugars to fructose-6P, NH_3 and acetate. Additionally, a chemotaxis system and a nutrient sensor to detect extracellular chitin, a specific chitoporin in the outer membrane and three transport complexes in the inner membrane have been identified [40]. Expression of the chitinolytic

genes is regulated by the two-component signalling system ChiS/ChiR. When in the environment chitin oligosaccharides deriving from chitin partial hydrolysis are present at low levels (*minus* phenotype), the binding protein CBP (chitin oligosaccharide binding protein) binds to the periplasmic domain of the sensor kinase ChiS, locking it into an inactive conformation. On the contrary, in the *plus* phenotype, COS enter the periplasmic space, where they bind to CBP and dissociate it from ChiS: the active conformation of the kinase activates a cascade of signals involving ChiR, which resulted in chitinolytic gene expression [40].

Among Gram-positive bacteria, the best studied chitinolytic system is that of streptomycetes. Also these microorganisms, well-known decomposers of chitin in soil, possess diverse and multiple chitinase genes with different characteristics. In particular, the analysis of *S. coelicolor* A3(2) genome revealed the presence of at least thirteen different genes coding for chitinases, not organised in operons but randomly distributed across the genome: eleven members of GH18 family and two of GH19 [41]. Chitinase production in these filamentous microorganisms is induced by chitin and repressed in the presence of readily utilisable carbon sources such as glucose [41]. Although several aspects of the regulation of chitinase production by streptomycetes are still unclear, what is known is that this regulation occurs at level of transcription, and a pair of 12 bp conserved direct repeat sequences has been found to play a key role. These conserved sequences, present in the promoter regions of various type of genes and called "dre" (DasR responsive elements), are recognised by a GntR-like transcriptional regulator, DasR [41, 42]. DasR is a pleiotropic multifunctional regulator, part of the nutrient-sensing system of actinomycetes. It is also involved in morphological development, which acts as both an activator (for chitinase genes) and a repressor (for genes of the sugar phosphotransferase – PTS – and ATP-binding cassette – ABC – transport systems and for the *actII-4* and *redZ* genes, respectively involved in actinorhodin and prodigiosin antibiotics production in *S. coelicolor*). Besides DasR, other molecular regulators have been identified, including the two-component regulator ChiS/ChiR of *chiC* in *S. coelicolor* A3(2), the Cpb1 DNA-binding protein and Reg1 both in *S. lividans* [41].

In fungi, chitinase gene expression has been reported to be controlled by a repressor/inducer system, in which chitin or other products of degradation (such as GlcNAc

or glucosamine) act as inducers whereas glucose or easily metabolised carbon sources act as repressors [34]. The most studied mycoparasitic fungus is *Trichoderma harzianum*, whose chitinolytic system is composed of seven chitinases, including four endochitinases, one chitobiosidase and two β-(1,4) *N*-acetyl-glucosaminidases [34, 43]. Multiple chitinolytic enzymes have been identified also in *Trichoderma viride* culture broth, including three major extracellular enzymes of ~ 30 kDa, ~ 45 kDa and ~ 65 kDa, in addition to minor ones [44, 45].

2.2.5 Biotechnological applications of chitin, chitin derivatives and chitinases

Chitin is a white, hard, inelastic, nitrogenous polysaccharide, highly hydrophobic and insoluble in water and most organic solvents owing to its intermolecular hydrogen bonds. It is soluble in strong acids such as dichloroacetic and trichloroacetic acids and in highly polar fluorinated solvents like hexafluoroisopropyl alcohol, hexafluoroacetone and chloroalcohols. The deacetylated chitosan is soluble in aqueous solutions of weak acids such as acetic, nitric, hydrochloric and phosphoric acid, but insoluble in water, organic solvents and aqueous bases. Its properties vary depending on the degree of acetylation and molecular weight [30]. Chitin and chitosan are biocompatible, non-toxic, non-allergenic, biodegradable, biorenewable and bioabsorbable, with antibacterial and wound-healing abilities, as well as mechanical strength and low immunogenicity [30]. Thanks to these characteristics and despite the limited solubility, a very broad range of applications in different fields have been reported, such as food technology, material science, microbiology, agriculture, wastewater treatment, drug delivery systems, tissue engineering and bio-nanotechnology (see below for further details) [27, 28, 30]. Also COS find wide-range applications, because of their high solubility in water and low viscosity. Additionally, it is possible to produce several COS-derivatives (amino-derived COS, carboxylated COS, gallyc COS and sulphated COS) with specific and targeted properties [29].

Biomedicine. Chitin and chitosan are effective agents for haemostasis maintenance through aggregating platelets, and therefore could be employed for facilitating wound healing. Similarly, chitosan membranes have been designed to protect wounded and burn skin by preventing bacterial invasion and halting the evaporation of the skin's water. Chitin and its

derivatives have been used also as scaffolds for bone, skin and other natural tissue regeneration, as well as diet additives to battle obesity and hypercholesterolemia, being able to bind lipid micelles and therefore reducing the amount of absorbed cholesterol. Chitosan and chitin derivatives such as *N*-succinyl-chitosan, carboxymethyl chitin and chitosan hydrogel, have been used as protein and drug carrier and as safe DNA carriers for gene therapy.

Cosmetics. Chitosan and chitin are employed as additives in cosmetic products for hair, skin and oral care, as natural substitutes of hydrocolloids, emulsifiers, moisturisers or fungicide agents.

Paper industry. Biodegradable chitin and chitosan strengthen recycled paper and increase the environmental friendliness of packaging; moreover, the paper produced with chitosan has a smoother surface and is more resistant to moisture.

Textile industry. Derivatives of chitin are used to impart antistatic and soil repellent characteristics to the textiles. Chitin can be employed also in printing and finishing preparations, while chitosan is able to remove dyes from dye processing effluents.

Food processing. Chemical food preservatives can be replaced with chitin-based ones, which have the advantage of being safer and both antibacterial and fungicidal. The deacifying ability of chitin is utilised in coffee industry and to clarify beverages such as wine, beer and fruit juices, while microcrystalline chitin is employed as emulsifying and gelling agent for stabilizing foods.

Agriculture. COS have been shown to play an important role in defence mechanisms of plants against microbial invasion. For example, chitin fragments can desensitise the perception system of tomato and rice, leading to an improvement of the defence mechanisms in plant cells. Additionally, chitin treated seeds were found to have growth accelerating and enhancing effects.

Bio-nanotechnology. Chitosan has gained growing interest in the field of nanomaterials, because of its biocompatibility, high permeability, cost-effectiveness, non-toxic property and excellent film-forming ability.

Bioremediation. Several studies have reported the ability of chitin and chitosan to remove heavy metals, such as copper, iron, lead, silver, titanium, but also uranium, plutonium, mercury, arsenic and tungsten, from the environment. They also have been successfully

tested for the adsorption of organic pollutants and petroleum products and as flocculating agents, thus finding application in the wastewater treatment.

Energy production. An emerging application of chitin is its role as potential source of alternative energy. For example, the digestion of chitin by bacteria proved to be effective in producing electrons that act as horsepower in a microbial fuel cell-containing robot. Additionally, chitin has been utilised by *Clostridium paraputrificum* to produce hydrogen gas. The advantage of using chitin as energy resource is that most chitin sources are waste and non-food materials, such as shrimp shells.

Today, several companies are producing chitin and chitosan at a commercial scale, mainly starting from the waste residues of seafood exoskeletons, such as shrimps, crabs, oysters and squids. The process of chitin isolation from the shells consists of three steps: demineralisation, deproteinisation and bleaching. For each passage, different protocols can be applied, according to the chitin source or its further application. Demineralisation is usually achieved in 1 to 3 hours treatment with acids at room temperature, while deproteinisation is performed using aqueous sodium or potassium hydroxide solutions. Finally, the removal of pigment residues from chitin can be achieved by extraction at room temperature with acetone, chloroform, ethyl acetate, ethanol and ether mixture. Decolourisation is usually carried out through a bleaching treatment with hydrogen peroxide, potassium permanganate or sodium hypochlorite. Chitosan production at industrial level is performed through a 10-hour long process, at high temperatures (from 80 to 140 °C) and using concentrated (30 to 60% (w/v)) sodium or potassium hydroxide solution. Similarly, COS are usually obtained by treating chitin in cold 70% (w/v) sulphuric acid solution [46].

It is therefore essential to develop more sustainable and environment-friendly processes for the extraction and derivatisation of chitin, to be used as valid and efficient alternatives to the traditional chemical-based ones. An answer to this demand are chitinolytic enzymes.

Indeed, chitinases have many industrial and agricultural applications [31, 32]. Chitinolytic enzymes can be employed for the production of chitin derivatives (COS, glucosamines and GlcNAc) with immense pharmaceutical and nutritional potential, as described above. Other applications include the production of single cell proteins, fungal protoplasts,

neoglycoproteins useful for structure-function studies, and synthetic polysaccharides employed for the development of prophylactic and therapeutic agents, as well as the direct application of chitinases in medicine for the treatment of systemic fungal infections.

Moreover, since pathogenic fungi and insect pests contain chitin in their protective covers, chitinases can be employed as alternative biopesticides, alone or as supplements to the commonly used fungicides and insecticides. For example, it has been demonstrated that both first (I) and fourth (IV) instar larvae of the yellow fever and dengue vector *Aedes aegypti* can be killed within 48 h with the help of a crude preparation from the saprophytic fungus *Myrothecium verrucaria*, containing insect cuticle degrading enzymes including chitinases [47]. Similarly, two chitinolytic enzymes from *Streptomyces albidoflavus* proved to have an inhibitory effect both on the insect coffee berry borer *Hypothenemus hampei* and on the rust fungus *Hemileia vastatrix,* two of the most limiting coffee pests worldwide [48]. Also the constitutive overexpression of antifungal chitinases in agriculturally important plants represents a promising strategy for conferring them genetic resistance against phytopathogen fungi. Chitinases from the mycoparasitic fungus *Trichoderma* spp. have been overexpressed in several agriculturally important plants, e.g. lemon, cotton, apple and carrot [33].

2.3 Traditional and innovative methods for studying and producing microbial chitinolytic enzymes

The first chitinase was identified by Bernard in 1911, who found a thermosensitive and diffusible antifungal factor from orchid pulp; next, in 1929, a similar enzyme from snails was reported by Karrer and Hoffman [49]. Since then, several works have been published concerning the identification and characterisation of chitinase-producing organisms and chitinolytic enzymes. The interest in this class of enzymes and the demand for biocatalysers with new and desirable properties have kept growing as additional uses of chitinases and chitin derivatives become apparent. Microorganisms, the primary degraders of chitin in the environment, are a rich source of valuable chitin-modifying enzymes [37]. In the following

subsections, a brief review of traditional and innovative (*i.e.* metagenomic) approaches for the identification of new microbial chitinases is provided.

2.3.1 Detection and production of chitinases in culturable microorganisms

The most simple and traditionally employed method for the detection of chitin-degrading microorganisms is their screening using solid media supplemented with colloidal or swollen chitin. These assays are based on the fact that in many chitinolytic systems chitin is hydrolysed by secreted chitin depolymerases. Since these enzymes are able to diffuse through agar, the hydrolysis of polymeric chitin incorporated into the medium can be visualised as halos or zones of clearing around colonies. Though these assays have a limited sensitivity, they represent a simple and inexpensive method to identify chitinolytic microorganisms [37]. Examples can be found in [50], [51], [52] and [53], just to cite a few.

Alternative complex substrates that may be included in agar media to facilitate the visualisation of chitinolytic activities are ethylene glycol chitin (EGC) or Chin Azure, prepared by covalently linking a soluble dye to colloidal chitin. Similarly, a variety of synthetic chitin analogues can be used to screen for the production of chitin-degrading enzymes, including the fluorigenic 4-methylumbelliferyl (4-MU) analogues, 4-methylumbelliferyl β-D-*N,N'*-diacetylchitobioside (4-MU-$(GlcNAc)_2$) and 4-methylumbelliferyl β-D-*N,N',N''*-triacetylchitotrioside (4-MU-$(GlcNAc)_3$), prepared by linking a 4-MU moiety to the reducing end of chitobiose and chitotriose, respectively; or the paranitrophenol (PNP) chitin analogues 4-nitrophenyl *N*-acetyl-β-D-glucosaminide, 4-nitrophenyl *N,N'*-diacetyl-β-D-chitobioside and 4-nitrophenyl *N,N',N''*-triacetyl-β-D-chitotrioside, whose liberated PNP moiety can be detected spectrophotometrically at 410 nm [37].

Once identified a suitable chitin-degrading microorganism, two different approaches are feasible. The first one is chitinase production and purification from the chitin-degrading microorganism itself. Reports are available about microbial chitinase production in liquid batch, continuous and fed-batch fermentation, as well as solid-state fermentation and biphasic cell systems [31]. Optimised culture conditions for chitinase expression are not universal and vary depending on the species of the microorganism. Chitin supplementation to the culture broth usually exerts an inducing effect on enzyme production, but also other

parameters such as carbon and nitrogen sources, pH, incubation temperature and aeration rate proved to affect chitinase production. For example, the addition of amino acids (tryptophan, tyrosine, glutamine and arginine, at 0.1 mM concentration) stimulated chitinase production from *Bacillus* sp. BG-11, whereas the supplementation of wheat bean in combination with chitinous substrates doubled the chitinase yield obtained from *Enterobacter* sp. NRG4 [54]. When the enzyme is secreted into the culture broth, its recovery usually starts with a fractional ammonium sulphate precipitation step to concentrate the proteins, followed by one or more steps of purification with ion-exchange or gel filtration chromatography (examples can be found in [51], [55], [56] and [57]). An alternative is purification by affinity chromatography using chitin-containing columns.

When the sequence of the chitinase coding gene is available, the second approach, *i.e.* heterologous cloning and expression, is a valuable option, as described in details in subsection 2.3.3.

2.3.2 Metagenomic approaches for new chitinase gene identification

Metagenomics is a promising, even if still underexploited, tool for the identification of novel chitinolytic enzymes otherwise encrypted in natural microbial communities. Naturally-occurring suppressive soils and chitin-amended soils and sediments are thought to be particularly valuable resources for the construction of metagenomic libraries: in these environmental samples the fitness of chitin degrading bacteria and hence the proportion of genes related to chitin degradation in the metagenomic DNA are expected to be increased [58, 59]. Screening of the so-constructed metagenomic libraries can be achieved with (i) functional approaches, based on the degradation of the same polymeric chitins and synthetic chitin analogues described in the previous subsections; and/or with (ii) sequence-based screening strategies, exerted by PCR with degenerate primers designed on the basis of the conserved regions of the catalytic domains of family 18 and 19 glycoside hydrolases (reported in 2.2.3). In the latter method, once identified a positive clone, the reconstruction of the complete gene sequence can be achieved by primer walking techniques.

To date, only a few studies have employed metagenomics to identify novel putative chitinase sequences and even less works had led to the isolation and proper characterisation

of biologically active chitinolytic enzymes (a review of the published papers can be found in the conclusive chapter of this dissertation). Besides some technical challenges in the metagenomic library construction and screening, one of the mail limiting step is the identification of a suitable expression platform for the successful expression of the target enzyme, as described in the following subsection.

2.3.3 Chitinase heterologous expression

When the aim of a study is a full biochemical, structural and functional characterisation of an enzyme or, more in general, when it is necessary to produce high amount of a protein for industrial processes or for the development of commercial goods, heterologous cloning and expression in model hosts are the best pathway to follow. The same applies to chitinases, either if produced by cultivable strains or sourced by a metagenomic approach.

There are many microbial hosts available for the production of recombinant enzymes, but the preferred choice is usually the Gram-negative bacterium *Escherichia coli*. Several bacterial chitinolytic enzymes have been cloned, expressed and purified in this heterologous host, including chitinases from *Alteromonas* sp. [60], *Streptomyces* sp. [61], *Vibrio* sp. [62], *Bacillus* sp. [63, 64] and *Enterobacter* sp. [65], as well as from *Serratia* sp. [66], *Chitiniphilus shinanonensis* [67], *Rhodothermus marinus* [68], *Citrobacter freundii* [69] and *Chromobacterium violaceum* [70]. Also some fungal chitinolytic enzymes, for instance from *Trichoderma* spp. [71, 72], *Beauveria bassiana* [73] and *Paecilomyces thermophila* [74], and more recently some archeal chitinases [75, 76] have been produced in *E. coli*. The advantages of this expression platform are several and well-known, including the unparalleled fast growth kinetics (in glucose-salts media the doubling time is about 20 minutes) and the possibility to easily reach high cell densities in inexpensive fermentation media [77]. Moreover, there are many molecular tools and protocols at hand for the manipulation of this microorganism and for the high-level production of heterologous proteins, such as a vast catalogue of expression plasmids and a great number of engineered strains. Nowadays, the most common expression plasmids are the result of multiple combinations of replicons, promoters, selection markers, multiple cloning sites, and fusion protein/fusion protein removal strategies. Commercially available plasmids include, among

the others, the most commonly used p*T7*-based pET (Novagen) and pUC (Thermo Scientific) series, the pACYC (Addgene) and the p*araBAD*-based pBAD (Invitrogen) series employed for dual expression of recombinant proteins using two separate plasmids, and the more recent pET-DUET vectors (Novagen) where two different genes can be cloned in the same plasmid [77]. Dozens of *E. coli* strains are then available as hosts, all of them with advantages and disadvantages, starting from the most commonly employed BL21 and BL21(DE3) strains, to the AD494 and Origami™ (Novagen) strains developed to enhance disulphide bond formation in the cytoplasm, or the C41(DE3) and C43(DE3) strains specific for membrane and toxic protein expression [77, 78]. Moreover, *E. coli* often has the advantage of lacking endogenous proteins functionally or chemically similar to the ones heterologously expressed, or able to interact with them. This is particularly true for chitinases, absent in *E. coli* genome: the lack of an endogenous counterpart, facilitates the detection and the purification of the herein produced recombinant chitinases.

However, there are also some limitations to using *E. coli* as an expression host, some of which have been already mentioned in subsection 1.2.1. These include first of all the inability to perform certain post-translational modifications (such as glycosylation and acetylation) and defects in protein maturation and disulphide bond formation. Codon bias is another factor which affects protein expression levels and arises when the occurrence of synonymous codons in the foreign coding DNA is significantly different from that of *E. coli*, resulting in translation rate reduction, amino acid misincorporation and/or truncation of the polypeptide. Strategies for solving codon usage bias can be the codon optimisation of the heterologous sequence or increasing the availability of underrepresented tRNAs by choosing *E. coli* strains carrying plasmids with extra copies of rare tRNAs, such as the BL21(DE3)CodonPlus strain (Stratagen) with extra copies of the tRNAs for AGG/AGA (Arg), AUA (Ile) and CUA (Leu), or the Rosetta(DE3) strain (Novagen) supplying all the above-mentioned codons plus CCC (Pro) and GGA (Gly) [77]. Furthermore, the maintenance of a episomial plasmid often induces a stress response, especially when the recombinant protein is highly expressed. Some proteins directly influence host cellular metabolism through their enzymatic properties, but in general the expression of a recombinant protein induces a so-called “metabolic burden”, *i.e.* the amount of resources (raw material and energy) that are withdrawn from the host metabolism for maintenance and expression of the heterolgous DNA. This usually results in an impaired growth rate and a lowered biomass production, as

well as an enhanced *in vivo* proteolysis of the target protein. Another major drawback of recombinant protein overexpression in *E. coli* is their accumulation into inclusion bodies (IBs). IBs are electron-dense protein granules rod- or sphere-like in shape, with diameters ranging from 0.2 to 1.2 µm and observable by optical microscopy. IBs are usually homogenous, being composed for up to 80-95% of the recombinant protein (mostly in inactive form, partially folded or unstructured), with only little contaminating host proteins, membrane phospholipids, ribosomial components, DNA or RNA [77, 79, 80]. The major disadvantage of IBs formation is the need to develop a protein refolding protocol, which can be tedious and not always effective in yielding native folded and active proteins [80]. The common protocol to obtain protein solubilisation from IBs is based on their treatment with strong denaturing agents (urea, guanidium hydrochloride, guanidium thiocianate), as applied for example in [71], [72] and [73] for the solubilisation of the chitinolytic enzymes Chit33 and Chit42, Ech30 and Bbchit1, respectively. As an alternative, it is possible trying to prevent IBs formation, (i) by reducing the culture growth and protein production rate, and/or (ii) favouring the folding process by co-expressing chaperones, and/or (iii) fusing the heterologous proteins with tags enhancing their solubility [77]. Another possibility is targeting the heterologous protein to the periplasmic space or to the growth medium, thus reducing protein concentration into the cytoplasm and, therefore, the protein-protein interaction leading to IBs formation. Periplasmic recombinant expression is favourable also due to the lower proteolytic activity and the oxidising environment that facilitates the formation of disulphide bonds [79]. Methods to achieve this goal include (i) the selection and modification of the signal peptide, (ii) co-expression of helper proteins to assist translocation trough the cytosolic membrane and folding, (iii) improvement of periplasmic release and (iv) protection of target protein from degradation and contamination [81]. Boer *et al*., 2007 [71] reported about the expression of two *Trichoderma harzianum* chitinases, Chit33 and Chit42, in the periplasm of *E. coli*, by replacing the native signal peptide sequence with the filamentous phage fd pIII protein signal sequence. In [70] an example of chitinase secretion by *E. coli* in the culture broth directed by its native signal peptide is reported.

The limits of heterologous protein production in *E. coli* have stimulated the development of alternative prokaryotic and eukaryotic expression systems, each with advantages and disadvantages. A chitinase gene from *Strenotrophomonas maltophila* was expressed and purified in the Gram-negative bacterium *Burkholderia cepacia* [82], whereas four different

chitinases from *Aeromonas hydrophila, Pseudomonas maltophila* [83] and *Bacillus licheniformis* [84] were expressed in *Bacillus thuringiensis*. Also the methylotrophic yeast *Pichia pastoris* has been successfully employed for the production of two thermostable chitinases, TaCHIT1 and CtCHIT1, from *Thermoascus auranticus* and *Chaetomium thermophilum* [85], as well as for the expression and purification of Bbchit1 from *Beauveria bassiana* [73].

Streptomyces spp. are other promising cell factories for the expression of chitinases, as demonstrated in [86] and [87]. These Gram-positive, high GC-content microorganisms have the main advantage to possess a developed secretory system: secreted proteins are usually natively folded, can be produced at comparable or even higher levels than intracellular ones, and can be more easily purified, reducing the risks of contamination by host proteins, nucleic acids and endotoxins [88]. Their main limitation for the heterologous expression of chitinolytic enzymes is the presence of endogenous chitinases (see subsection 2.2.4), which can interfere with the detection of the recombinant enzyme. For the total or partial repression of this complex endogenous system, two different approaches can be followed: the exploitation of carbon catabolite repression effect exerted by glucose or other monosaccharides [89], or the knock-out of one or more regulators [41].

3. References

1. Bunge J, Willis A, Walsh F: **Estimating the number of species in microbial diversity studies**. *Annual Review of Statistics and Its Applications* 2014, **1**:427-445.
2. Simon C, Daniel R: **Metagenomic analyses: past and future trends**. *Appl Environ Microbiol* 2011, **77**(4):1153-1161.
3. Cowan D, Meyer Q, Stafford W, Muyanga S, Cameron R, Wittwer P: **Metagenomic gene discovery: past, present and future**. *Trends Biotechnol* 2005, **23**(6):321-329.
4. Daniel R: **The metagenomics of soil**. *Nat Rev Microbiol* 2005, **3**(6):470-478.
5. Mocali S, Benedetti A: **Exploring research frontiers in microbiology: the challenge of metagenomics in soil microbiology**. *Res Microbiol* 2010, **161**(6):497-505.
6. Kennedy J, O'Leary ND, Kiran GS, Morrissey JP, O'Gara F, Selvin J, Dobson AD: **Functional metagenomic strategies for the discovery of novel enzymes and biosurfactants with biotechnological applications from marine ecosystems**. *J Appl Microbiol* 2011, **111**(4):787-799.
7. Nikolaki S, Tsiamis G: **Microbial diversity in the era of omic technologies**. *Biomed Res Int* 2013, **2013**:958719.
8. Ekkers DM, Cretoiu MS, Kielak AM, Elsas JD: **The great screen anomaly--a new frontier in product discovery through functional metagenomics**. *Appl Microbiol Biotechnol* 2012, **93**(3):1005-1020.
9. Streit WR, Schmitz RA: **Metagenomics--the key to the uncultured microbes**. *Curr Opin Microbiol* 2004, **7**(5):492-498.
10. van Elsas JD, Costa R, Jansson J, Sjöling S, Bailey M, Nalin R, Vogel TM, van Overbeek L: **The metagenomics of disease-suppressive soils - experiences from the METACONTROL project**. *Trends Biotechnol* 2008, **26**(11):591-601.
11. Lombard N, Prestat E, van Elsas JD, Simonet P: **Soil-specific limitations for access and analysis of soil microbial communities by metagenomics**. *FEMS Microbiol Ecol* 2011, **78**(1):31-49.
12. Steele HL, Streit WR: **Metagenomics: advances in ecology and biotechnology**. *FEMS Microbiol Lett* 2005, **247**(2):105-111.
13. Schofield MM, Sherman DH: **Meta-omic characterization of prokaryotic gene clusters for natural product biosynthesis**. *Curr Opin Biotechnol* 2013, **24**(6):1151-1158.
14. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM: **Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products**. *Chem Biol* 1998, **5**(10):R245-249.
15. Schmeisser C, Steele H, Streit WR: **Metagenomics, biotechnology with non-culturable microbes**. *Appl Microbiol Biotechnol* 2007, **75**(5):955-962.
16. Rajendhran J, Gunasekaran P: **Strategies for accessing soil metagenome for desired applications**. *Biotechnol Adv* 2008, **26**(6):576-590.
17. Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC *et al*: **Comparative metagenomics of microbial communities**. *Science* 2005, **308**(5721):554-557.
18. Schloss PD, Handelsman J: **Biotechnological prospects from metagenomics**. *Curr Opin Biotechnol* 2003, **14**(3):303-310.
19. Uhlik O, Leewis MC, Strejcek M, Musilova L, Mackova M, Leigh MB, Macek T: **Stable isotope probing in the metagenomics era: a bridge towards improved bioremediation**. *Biotechnol Adv* 2013, **31**(2):154-165.
20. McMahon MD, Guan C, Handelsman J, Thomas MG: **Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression**. *Appl Environ Microbiol* 2012, **78**(10):3622-3629.

21. Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF: **Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria**. *Appl Environ Microbiol* 2010, **76**(5):1633-1641.
22. Banik JJ, Brady SF: **Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules**. *Curr Opin Microbiol* 2010, **13**(5):603-609.
23. Uchiyama T, Abe T, Ikemura T, Watanabe K: **Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes**. *Nat Biotechnol* 2005, **23**(1):88-93.
24. Lorenz P, Eck J: **Metagenomics and industrial applications**. *Nat Rev Microbiol* 2005, **3**(6):510-516.
25. Lasken RS: **Genomic sequencing of uncultured microorganisms from single cells**. *Nat Rev Microbiol* 2012, **10**(9):631-640.
26. Bai Y, Eijsink VG, Kielak AM, van Veen JA, de Boer W: **Genomic comparison of chitinolytic enzyme systems from terrestrial and aquatic bacteria**. *Environ Microbiol* 2014.
27. Khoushab F, Yamabhai M: **Chitin research revisited**. *Mar Drugs* 2010, **8**(7):1988-2012.
28. Shukla SK, Mishra AK, Arotiba OA, Mamba BB: **Chitosan-based nanomaterials: a state-of-the-art review**. *Int J Biol Macromol* 2013, **59**:46-58.
29. Lodhi G, Kim YS, Hwang JW, Kim SK, Jeon YJ, Je JY, Ahn CB, Moon SH, Jeon BT, Park PJ: **Chitooligosaccharide and its derivatives: preparation and biological applications**. *Biomed Res Int* 2014, **2014**:654913.
30. Wan AC, Tai BC: **CHITIN--a promising biomaterial for tissue engineering and stem cell technologies**. *Biotechnol Adv* 2013, **31**(8):1776-1785.
31. Dahiya N, Tewari R, Hoondal GS: **Biotechnological aspects of chitinolytic enzymes: a review**. *Appl Microbiol Biotechnol* 2006, **71**(6):773-782.
32. Adrangi S, Faramarzi MA: **From bacteria to human: a journey into the world of chitinases**. *Biotechnol Adv* 2013, **31**(8):1786-1795.
33. Hartl L, Zach S, Seidl-Seiboth V: **Fungal chitinases: diversity, mechanistic properties and biotechnological potential**. *Appl Microbiol Biotechnol* 2012, **93**(2):533-543.
34. Duo-Chuan L: **Review of fungal chitinases**. *Mycopathologia* 2006, **161**(6):345-360.
35. Henrissat B: **A classification of glycosyl hydrolases based on amino acid sequence similarities**. *Biochem J* 1991, **280 (Pt 2)**:309-316.
36. Udaya Prakash NA, Jayanthi M, Sabarinathan R, Kangueane P, Mathew L, Sekar K: **Evolution, homology conservation, and identification of unique sequence signatures in GH19 family chitinases**. *J Mol Evol* 2010, **70**(5):466-478.
37. Howard MB, Ekborg NA, Weiner RM, Hutcheson SW: **Detection and characterization of chitinases and other chitin-modifying enzymes**. *J Ind Microbiol Biotechnol* 2003, **30**(11):627-635.
38. Beier S, Bertilsson S: **Bacterial chitin degradation-mechanisms and ecophysiological strategies**. *Front Microbiol* 2013, **4**:149.
39. Vaaje-Kolstad G, Horn SJ, Sørlie M, Eijsink VG: **The chitinolytic machinery of *Serratia marcescens*--a model system for enzymatic degradation of recalcitrant polysaccharides**. *FEBS J* 2013, **280**(13):3028-3049.
40. Li X, Roseman S: **The chitinolytic cascade in *Vibrios* is regulated by chitin oligosaccharides and a two-component chitin catabolic sensor/kinase**. *Proc Natl Acad Sci U S A* 2004, **101**(2):627-631.
41. Nazari B, Saito A, Kobayashi M, Miyashita K, Wang Y, Fujii T: **High expression levels of chitinase genes in *Streptomyces coelicolor* A3(2) grown in soil**. *FEMS Microbiol Ecol* 2011, **77**(3):623-635.

42. Rigali S, Titgemeyer F, Barends S, Mulder S, Thomae AW, Hopwood DA, van Wezel GP: **Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces***. *EMBO Rep* 2008, **9**(7):670-675.
43. Haran S, Schickler H, Oppenheim A, Chet I: **New components of the chitinolytic system of *Trichoderma harzianum***. *Mycological Research* 1995, **99**:441-446.
44. Omumasaba CA, Yoshida N, Ogawa K: **Purification and characterization of a chitinase from *Trichoderma viride***. *J Gen Appl Microbiol* 2001, **47**(2):53-61.
45. Giridhar D, Ravi Sankar N, Kumar K, Kartheek D, Rajanikanth P, Nagalakshmi Devamma M: **Purification, Characterization and antifungal activity of chitinase from *Trichoderma viride* N9**. *Journal of Cell and Tissue Research* 2012, **12**:3187-3192.
46. Kaur S, Dhillon GS: **The versatile biopolymer chitosan: potential sources, evaluation of extraction methods and applications**. *Crit Rev Microbiol* 2014, **40**(2):155-175.
47. Mendonsa E, Vartak P, Rao J, Deshpande M: **An enzyme from *Myrothecium verrucaria* that degrades insect cuticles for biocontrol of *Aedes aegypti* mosquito**. *Biotechnology Letters* 1996, **18**:373-376.
48. Martínez CP, Echeverri C, Florez JC, Gaitan AL, Góngora CE: **In vitro production of two chitinolytic proteins with an inhibiting effect on the insect coffee berry borer, *Hypothenemus hampei* (Ferrari) (Coleoptera: Curculionidae) and the fungus *Hemileia vastatrix* the most limiting pests of coffee crops**. *AMB Express* 2012, **2**(1):22.
49. Felse PA, Panda T: **Regulation and cloning of microbial chitinase genes**. *Appl Microbiol Biotechnol* 1999, **51**(2):141-151.
50. Jankiewicz U, Brzezinska MS, Saks E: **Identification and characterization of a chitinase of *Stenotrophomonas maltophilia*, a bacterium that is antagonistic towards fungal phytopathogens**. *J Biosci Bioeng* 2012, **113**(1):30-35.
51. Waghmare SR, Ghosh JS: **Chitobiose production by using a novel thermostable chitinase from *Bacillus licheniformis* strain JS isolated from a mushroom bed**. *Carbohydr Res* 2010, **345**(18):2630-2635.
52. Park H, Kim D, Kim I, Lee C, Kim I, Kim J, Kim S, Lee H, Yim J: **Characteristics of cold-adaptive endochitinase from Antarctic bacterium *Sanguibacter antarcticus* KOPRI 21702**. *Enzyme and Microbial Technology* 2009, **45**(5):391-396.
53. Liu CL, Shen CR, Hsu FF, Chen JK, Wu PT, Guo SH, Lee WC, Yu FW, Mackey ZB, Turk J *et al*: **Isolation and identification of two novel SDS-resistant secreted chitinases from *Aeromonas schubertii***. *Biotechnol Prog* 2009, **25**(1):124-131.
54. Dahiya N, Tewari R, Tiwari RP, Hoondal GS: **Chitinase production in solid-state fermentation by *Enterobacter* sp. NRG4 using statistical experimental design**. *Curr Microbiol* 2005, **51**(4):222-228.
55. Li DC, Zhang SH, Liu KQ, Lu J: **Purification and partial characterization of a chitinase from the mycoparasitic fungus *Trichothecium roseum***. *J Gen Appl Microbiol* 2004, **50**(1):35-39.
56. Barghini P, Moscatelli D, Garzillo AM, Crognale S, Fenice M: **High production of cold-tolerant chitinases on shrimp wastes in bench-top bioreactor by the Antarctic fungus *Lecanicillium muscarium* CCFEE 5003: bioprocess optimization and characterization of two main enzymes**. *Enzyme Microb Technol* 2013, **53**(5):331-338.
57. Singh AK, Chhatpar HS: **Purification and characterization of chitinase from *Paenibacillus* sp. D1**. *Appl Biochem Biotechnol* 2011, **164**(1):77-88.
58. Kielak AM, Cretoiu MS, Semenov AV, Sørensen SJ, van Elsas JD: **Bacterial chitinolytic communities respond to chitin and pH alteration in soil**. *Appl Environ Microbiol* 2013, **79**(1):263-272.
59. Cretoiu MS, Korthals GW, Visser JH, van Elsas JD: **Chitin amendment increases soil suppressiveness toward plant pathogens and modulates the actinobacterial and oxalobacteraceal communities in an experimental agricultural field**. *Appl Environ Microbiol* 2013, **79**(17):5291-5301.

60. Orikoshi H, Nakayama S, Miyamoto K, Hanato C, Yasuda M, Inamori Y, Tsujibo H: **Roles of four chitinases (chia, chib, chic, and chid) in the chitin degradation system of marine bacterium *Alteromonas* sp. strain O-7**. *Appl Environ Microbiol* 2005, **71**(4):1811-1815.
61. Yamashita Y, Okazaki K: **Purification and antifungal activity of recombinant chitinase from *Escherichia coli* carrying the family 19 chitinase gene of *Streptomyces* sp. J-13-3**. *Biosci Biotechnol Biochem* 2004, **68**(10):2193-2196.
62. Itoi S, Kanomata Y, Koyama Y, Kadokura K, Uchida S, Nishio T, Oku T, Sugita H: **Identification of a novel endochitinase from a marine bacterium *Vibrio proteolyticus* strain No. 442**. *Biochim Biophys Acta* 2007, **1774**(9):1099-1107.
63. Lee YS, Park IH, Yoo JS, Chung SY, Lee YC, Cho YS, Ahn SC, Kim CM, Choi YL: **Cloning, purification, and characterization of chitinase from *Bacillus* sp. DAU101**. *Bioresour Technol* 2007, **98**(14):2734-2741.
64. Huang CJ, Chen CY: **High-level expression and characterization of two chitinases, ChiCH and ChiCW, of *Bacillus cereus* 28-9 in *Escherichia coli***. *Biochem Biophys Res Commun* 2005, **327**(1):8-17.
65. Chernin LS, De la Fuente L, Sobolev V, Haran S, Vorgias CE, Oppenheim AB, Chet I: **Molecular cloning, structural analysis, and expression in *Escherichia coli* of a chitinase gene from *Enterobacter agglomerans***. *Appl Environ Microbiol* 1997, **63**(3):834-839.
66. Synstad B, Vaaje-Kolstad G, Cederkvist FH, Saua SF, Horn SJ, Eijsink VG, Sørlie M: **Expression and characterization of endochitinase C from *Serratia marcescens* BJL200 and its purification by a one-step general chitinase purification method**. *Biosci Biotechnol Biochem* 2008, **72**(3):715-723.
67. Huang L, Garbulewska E, Sato K, Kato Y, Nogawa M, Taguchi G, Shimosaka M: **Isolation of genes coding for chitin-degrading enzymes in the novel chitinolytic bacterium, *Chitiniphilus shinanonensis*, and characterization of a gene coding for a family 19 chitinase**. *J Biosci Bioeng* 2012, **113**(3):293-299.
68. Hobel CF, Hreggvidsson GO, Marteinsson VT, Bahrani-Mougeot F, Einarsson JM, Kristjánsson JK: **Cloning, expression, and characterization of a highly thermostable family 18 chitinase from *Rhodothermus marinus***. *Extremophiles* 2005, **9**(1):53-64.
69. Meruvu H, Donthireddy SR: **Purification and characterization of an antifungal chitinase from *Citrobacter freundii* str. nov. haritD11**. *Appl Biochem Biotechnol* 2014, **172**(1):196-205.
70. Lobo MD, Silva FD, Landim PG, da Cruz PR, de Brito TL, de Medeiros SC, Oliveira JT, Vasconcelos IM, Pereira HD, Grangeiro TB: **Expression and efficient secretion of a functional chitinase from *Chromobacterium violaceum* in *Escherichia coli***. *BMC Biotechnol* 2013, **13**:46.
71. Boer H, Simolin H, Cottaz S, Söderlund H, Koivula A: **Heterologous expression and site-directed mutagenesis studies of two *Trichoderma harzianum* chitinases, Chit33 and Chit42, in *Escherichia coli***. *Protein Expr Purif* 2007, **51**(2):216-226.
72. Hoell IA, Klemsdal SS, Vaaje-Kolstad G, Horn SJ, Eijsink VG: **Overexpression and characterization of a novel chitinase from *Trichoderma atroviride* strain P1**. *Biochim Biophys Acta* 2005, **1748**(2):180-190.
73. Fan Y, Zhang Y, Yang X, Pei X, Guo S, Pei Y: **Expression of a *Beauveria bassiana* chitinase (Bbchit1) in *Escherichia coli* and *Pichia pastoris***. *Protein Expr Purif* 2007, **56**(1):93-99.
74. Kopparapu NK, Zhou P, Zhang S, Yan Q, Liu Z, Jiang Z: **Purification and characterization of a novel chitinase gene from *Paecilomyces thermophila* expressed in *Escherichia coli***. *Carbohydr Res* 2012, **347**(1):155-160.
75. Staufenberger T, Imhoff JF, Labes A: **First crenarchaeal chitinase found in *Sulfolobus tokodaii***. *Microbiol Res* 2012, **167**(5):262-269.
76. García-Fraga B, da Silva AF, López-Seijas J, Sieiro C: **Functional expression and characterization of a chitinase from the marine archaeon *Halobacterium salinarum* CECT 395 in *Escherichia coli***. *Appl Microbiol Biotechnol* 2014, **98**(5):2133-2143.

77. Rosano GL, Ceccarelli EA: **Recombinant protein expression in *Escherichia coli*: advances and challenges**. *Front Microbiol* 2014, **5**:172.
78. Gopal GJ, Kumar A: **Strategies for the production of recombinant protein in *Escherichia coli***. *Protein J* 2013, **32**(6):419-425.
79. Berlec A, Strukelj B: **Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells**. *J Ind Microbiol Biotechnol* 2013, **40**(3-4):257-274.
80. Ramón A, Señorale-Pose M, Marín M: **Inclusion bodies: not that bad...**. *Front Microbiol* 2014, **5**:56.
81. Yoon SH, Kim SK, Kim JF: **Secretory production of recombinant proteins in *Escherichia coli***. *Recent Pat Biotechnol* 2010, **4**(1):23-29.
82. Kobayashi DY, Reedy RM, Bick J, Oudemans PV: **Characterization of a chitinase gene from *Stenotrophomonas maltophilia* strain 34S1 and its involvement in biological control**. *Appl Environ Microbiol* 2002, **68**(3):1047-1054.
83. Wiwat C, Lertcanawanichakul M, Siwayapram P, Pantuwatana S, Bhumiratana A: **Expression of chitinase-encoding genes from *Aeromonas hydrophila* and *Pseudomonas maltophilia* in *Bacillus thuringiensis* subsp. *israelensis***. *Gene* 1996, **179**(1):119-126.
84. Tantimavanich S, Pantuwatana S, Bhumiratana A, Panbangred W: **Cloning of a chitinase gene into *Bacillus thuringiensis* subsp. *aizawai* for enhanced insecticidal activity**. *J Gen Appl Microbiol* 1997, **43**(6):341-347.
85. Li AN, Yu K, Liu HQ, Zhang J, Li H, Li DC: **Two novel thermostable chitinase genes from thermophilic fungi: cloning, expression and characterization**. *Bioresour Technol* 2010, **101**(14):5546-5551.
86. Miyashita K, Fujii T, Sawada Y: **Molecular cloning and characterization of chitinase genes from *Streptomyces lividans* 66**. *Microbiology* 1991, **137**(9):2065-2072.
87. Tsujibo H, Okamoto T, Hatano N, Miyamoto K, Watanabe T, Mitsutomi M, Inamori Y: **Family 19 chitinases from *Streptomyces thermoviolaceus* OPC-520: molecular cloning and characterization**. *Biosci Biotechnol Biochem* 2000, **64**(11):2445-2453.
88. Anné J, Maldonado B, Van Impe J, Van Mellaert L, Bernaerts K: **Recombinant protein production and streptomycetes**. *J Biotechnol* 2012, **158**(4):159-167.
89. Miyashita K, Fujii T, Saito A: **Induction and repression of a *Streptomyces lividans* chitinase gene promoter in response to various carbon sources**. *Biosci Biotechnol Biochem* 2000, **64**(1):39-43.

LAYOUT OF MY PHD THESIS

My PhD has been developed in the frame of the FP7 European project MetaExplore. This project, run from May 2009 to May 2014, involved eighteen research groups with the aim to develop metagenomic tools and techniques for the identification of novel enzymes involved in the biodegradation of recalcitrant natural molecules such as chitin and lignin, and xenobiotics anthropogenic compounds as halogenated aliphatic and aromatic molecules. European academia as well as biotechnology industries is, in fact, in continuous search of novel biocatalysts with innovative features and enhanced activities, to be employed in a vast range of industrial and environmental processes. One class of desired activities for the MetaExplore project consisted of novel chitinolytic enzymes. Chitinases, chitosanases and chitin deacetylases have many potential biotechnological and environmental applications, being for instance key enzymes in the food industry, where they can greatly improve the downstream processing of exoskeletal wastes. They can also be used as alternative and eco-friendly biocontrol agents, as well as for medical applications. Also chitin derivatives (chitosan and chitooligosaccharides) have a multiplicity of applications, which make chitinases even more attractive at industrial level.

In this context, the Laboratory of Microbial Biotechnology of University of Insubria collaborated at the MetaExplore project in the isolation, heterologous expression and characterisation (in conventional and alternative hosts) of metagenome-sourced enzymes, followed by their scale-up production.

The following chapters report Material and Methods, Results and Discussion related to the experimental work performed for the MetaExplore project during the three years of my PhD course. For clarity, the work has been divided into three sections, for a total of five different papers written along my fellowship.

The first section, entitled "**Heterologous expression and characterisation of metagenome-sourced chitinases**", contains two papers concerning the heterologous production in the conventional host *Escherichia coli* and the biochemical, functional and structural characterisation of two chitinases, identified in the frame of the MetaExplore project in two different soil metagenomic libraries.

In particular, the first paper **High-level production and characterisation of a metagenome-sourced chitobiosidase** by Francesca Berini, Ilaria Presti, Fabrizio Beltrametti, Loredano Pollegioni, Sara Sjöling, Flavia Marinelli, presents the research work performed in collaboration with the School of Life Sciences of Södertörn University (Huddinge, Sweden) and with the company Actygea (Gerenzano, Varese, Italy), on the heterologous production of Chi18H8. This enzyme represents, to our knowledge, the first metagenome-derived chitinase to be expressed and fully characterised. Chi18H8 was identified during a previous collaboration with the Swedish group of Prof. Sara Sjöling in a metagenomic library originated from a suppressive soil for clubroot disease of cabbage. The mechanism behind soil suppressiveness is usually related to the increased number and/or activity of chitinolytic microorganisms, able to hydrolyse the chitinous hyphae of pathogenic fungi; hence, naturally-occurring suppressive soils are considered a promising source of novel chitinases. The initial expression of this chitobiosidase in *E. coli* in transcriptional fusion with the gluthathione-S-transferase tag, allowed the purification of few micrograms of protein, sufficient only for an incomplete characterisation of the enzyme. However, the antifungal properties against common plant phytopathogens revealed by these preliminary studies, suggesting a possible application of Chi18H8 as biocontrol agent, encouraged us to continue working on this enzyme. Therefore, during the first part of my PhD, I cloned *chi18H8* gene in another expression system, in *E. coli* in the pET24b(+) plasmid, thus obtaining a protein in transcriptional fusion with an histidine examer tag. In this recombinant system, Chi18H8 was mainly accumulated into inclusion bodies in a prevalent inactive form. The development of a suitable and relatively easy solubilisation protocol, combined with the chitinase production scale-up both in the 3 L bench bioreactors of University of Insubria, and in the Actygea's 30 L industrial bioreactor, allowed the recovery of high amounts of recombinant chitinase (more than 7 mg of protein per g of cells), sufficient for its complete biochemical and functional characterisation and also for future applications. The characterisation studies herein

performed highlighted that Chi18H8 is a metallo-chitobiosidase, highly stable in sub-acidic conditions and with a good solvent-tolerance, which, together with its antifungal behaviour, make this enzyme an interesting candidate for future industrial or agricultural applications.

In the second chapter, **Genetic screening of a metagenomic library derived from chitin-amended agricultural soil produces a novel salt-tolerant chitinase** by Mariana Silvia Cretoiu, Francesca Berini, Anna Maria Kielak, Flavia Marinelli, Jan Dirk van Elsas, the work done in collaboration with the group of Prof. van Elsas of the University of Groningen (Groningen, The Netherlands) is presented. A metagenomic library was constructed from a chitin-amended disease-suppressive soil, and screened for genes encoding novel chitin-active enzymes. Hence, in this paper, the combined effect of natural suppressiveness and ecological enhancement (substrate enrichment) was evaluated with the aim to increase the efficiency of mining for desired chitinolytic enzymes with improved features. Among the five putative bacterial chitinase clones identified, chitinase 53D1 gene was selected for successive analysis. During my PhD, I managed to clone *53D1* in three different *E. coli* strains, either under the control of the inducible promoter of the commercial pET24b(+) and pCOLDI vectors, or under the control of its endogenous native promoter. The purified protein was characterised in terms of stability and activity, revealing that 53D1 is a metallo-chitobiosidase, highly active also on complex substrates, in a wide range of pHs and temperatures. Remarkably, the enzyme proved to be halotolerant, a quite uncommon behaviour for a chitinolytic enzyme. These key properties of 53D1 make it an interesting candidate for the treatment of seafood wastes such as shrimp carapace.

As described in the literature review, one of the major bottlenecks for the exploitation of metagenome-source biocatalysts and, generally speaking, for the high-level production and complete characterisation of candidate proteins, is their over-expression in microbial hosts. Protein production in the Gram-negative bacterium *E. coli*, the most used platform for recombinant protein expression, is often hampered by codon usage differences, cytotoxicity, inclusion bodies formation or inability to secrete the translated proteins. For this reason, a second goal of the MetaExplore project aimed at developing alternative cloning hosts, with different codon usages and higher protein secretion capacity than *E. coli*. The second section

of the present dissertation, entitled "***Streptomyces* spp. as alternative heterologous expression systems**", describes the employment of these Gram-positive filamentous bacteria for the successful expression of two proteins. Advantages of streptomycetes over *E. coli* include the innate secretion capacity, which reduces the risk of local accumulation of the over-expressed proteins and simplifies the purification procedures, the limited restriction-modification system and the low endogenous protease activity, as well as the presence of natural mechanisms of genetic exchange.

The third chapter includes the paper ***Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance** by Elisa Binda, Giorgia Letizia Marcone, Francesca Berini, Loredano Pollegioni and Flavia Marinelli, already published on BMC Biotechnology. $VanY_n$ is a protein involved in the mechanism of self-resistance in the uncommon actinomycete *Nonomuraea* sp. ATCC 39727, the natural producer of the glycopeptide antibiotic A40926, precursor of the second-generation dalbavancin. After being produced, purified and characterised in *E. coli*, in this work *$vanY_n$* gene has been cloned and expressed in three microorganisms taxonomically closely related to *Nonomuraea*, the streptomycetes *S. coelicolor* A3(2) *ΔvanRS*, *S. venezuelae* ATCC 10595 and *S. lividans* TK24, with a histidine examer tag added at the C- or N-terminus of the protein. The highest yield of protein expression and purification was achieved in *S. venezuelae*, from which it was possible to recover up to 12 mg of high-pure protein per litre of culture, a yield three fold higher than in *E. coli*. Also specific productivity was much higher than the one achieved in the Gram-negative host (1 mg of $VanY_n$ per g of cells in *S. venezuelae*, 0.13 mg/g cells in *E. coli*), thus confirming that streptomycetes are preferable hosts for $VanY_n$ heterologous production. Additionally, the work done confirmed the role of this D,D-peptidase/D,D-carboxypeptidase in reprogramming actinomycetes cell wall biosynthesis, thus conferring them an increased level of glycopeptide resistance.

In the fourth chapter, ***Streptomyces lividans* as host for the heterologous expression of a metagenome-sourced chitobiosidase** by Francesca Berini, Ilaria Presti, Giorgia Letizia Marcone, Loredano Pollegioni, Flavia Marinelli, this streptomycete is presented as alternative candidate to *E. coli* for the expression of Chi18H8. Since streptomycetes possess a complex endogenous chitinolytic system, which could interfere with heterologous

chitinase production and detection, before proceeding with *chi18H8* cloning, it was necessary to develop a method for the repression of these endogenous activities. Two different approaches were evaluated, the first one based on the metabolic repression by glucose added to the culture media, the second one involving the knock-out of the regulatory gene *dasR*. Once identified the best repression conditions, the *chi18H8* gene was cloned in a multicopy plasmid under the control of a heterologous constitutive promoter not-repressed by glucose. Enzymatic activity and zymogram analysis confirmed that the recombinant protein was secreted in the extracellular broth, which significantly simplified its purification. However, the purification yield achieved (16.9 μg of protein per g of cells), even if comparable to the one obtained for Chi18H8 purification from the first *E. coli* expression system (21 μg/g cells), was significantly lower than the one achieved with enzyme solubilisation from the inclusion bodies, described in the first chapter of the dissertation. Nevertheless, even if in this case *S. lividans* TK24 cannot be considered competitive with *E. coli* for Chi18H8 high-level production for future applications, the protein secretion into culture medium, the easy purification procedure, as well as the possibility to significantly repress the endogenous chitinolytic system by simply adding glucose to the culture, make this microorganism an interesting and valuable candidate for the expression of other metagenome-sourced chitinases, worthy of further exploration.

In the last section, called "**Chitinolytic enzymes as biocontrol agents**", the possibility to employ these hydrolytic enzymes as alternatives to traditional chemical-based pesticides is evaluated.

The paper **Effect of *Trichoderma viride* chitinases on the peritrophic matrix of the silkworm, *Bombyx mori***, by Francesca Berini, Silvia Caccia, Morena Casartelli, Terenzio Congiu, Eleonora Franzetti, Flavia Marinelli, Gianluca Tettamanti (chapter five of this thesis) deals with the study of the effect exerted by fungal chitinolytic enzymes on the peritrophic membrane of Lepidoptera. This acellular sheath that lines the midgut epithelium of most insects consists of a network of chitin fibrils associated with different proteins and glycoproteins and plays fundamental roles in insect digestion. Its chitin network is considered a possible target for the development of innovative strategies for integrated pest

management but, even if the employment of chitinases of viral, bacterial and plant origin have been already exploited for this purpose, fungal chitinases represent a still underexplored and promising resource. For this reason, during my PhD I worked on the biochemical characterisation of a commercial mixture of chitinolytic enzymes purified from the fungus *Trichoderma viride*, whose substrate specificity, pH and temperature optimum, as well as long-term stability were evaluated as premise for their subsequent application. *In vitro* effects of the fungal chitinase on the structure and permeability of the peritrophic membrane were then analysed and the results demonstrate the efficacy of these fungal chitinolytic enzymes as possible sustainable and environment-friendly alternatives to traditional chemical pesticides.

RESULTS

SECTION 1

“Heterologous expression and characterisation of metagenome-sourced chitinases”

Chapter 1

High-level production and characterisation of a metagenome–sourced chitobiosidase

Francesca Berini[1*], Ilaria Presti[1¥], Fabrizio Beltrametti[2], Sara Sjöling[3], Loredano Pollegioni[1], Flavia Marinelli[1]

[1]Department of Biotechnology and Life Sciences, University of Insubria Varese, Italy and "The Protein Factory Research Center", Politecnico di Milano, ICRM CNR Milano and University of Insubria Varese, Italy

[2]Actygea, Gerenzano, Varese Italy

[3]School of Life Sciences, Södertörn University, Huddinge Sweden

*Corresponding author, francesca.berini@uninsubria.it

¥ Present address: Industriale Chimica, Saronno, Varese Italy

Abstract

Metagenomic approaches provide access to environmental genetic diversity and allow the discovery of novel valuable biocatalysts for biotechnological applications. One of the major bottlenecks is the set-up of suitable expression and purification procedures, essential for the high-level production of metagenome-sourced enzymes, their complete characterisation and subsequent industrial exploitation. In this work, we report on the expression in *E. coli* of a chitinase, Chi18H8, previously identified in a suppressive-soil metagenomic library. A fast, robust and economically feasible protocol for the recombinant protein recovery from inclusion bodies was developed, followed by its production scale-up in a pilot scale industrial-bioreactor. With a prevalent activity at acid pH and mesophilic temperatures, Chi18H8 is a metallo-chitobiosidase, whose activity and structure are influenced by several metal ions and detergents. The long-term stability in acidic environments, the high solvent

tolerance, as well as the antifungal properties, make this enzyme an interesting candidate for both biotechnological and agricultural exploitations.

Keywords: chitinase, heterologous expression, inclusion bodies, biocontrol agent

Introduction

Chitin is an unbranched and insoluble biopolymer, composed of repeated units of *N*-acetylglucosamine (GlcNAc) and widely distributed in nature, for instance in the fungal cell walls and in the exoskeleton of arthropods and crustaceans, as well as in certain algae, molluscs, yeasts and nematodes [1, 2]. Most of the natural degradation of chitin has microbial origin, and bacteria and fungi have developed systems for the depolymerisation, transport and metabolism of this polysaccharide, usually including multiple chitinolytic genes acting in synergistic and consecutive manner [3, 4]. Based on the catalytic specificity, chitinolytic enzymes can be classified in: (i) endochitinases (EC 3.2.1.14), which produce multimers of GlcNAc by randomly cleaving chitin at internal sites; (ii) chitobiosidases (EC 3.2.1.29), which catalyse the release of soluble dimers starting at the non-reducing end of chitin microfibril; and (iii) β-(1,4) *N*-acetyl glucosaminidases (EC 3.2.1.30), which cleave the oligomeric products of endochitinases and chitobiosidases, thus generating monomers of GlcNAc [5]. Alternatively, chitin can be deacetylated to chitosan by the action of chitin deacetylases (EC 3.5.1.14), followed by its conversion by chitosanases (EC 3.2.1.132) to glucosamine residues. According to the classification system firstly proposed by Henrissat in 1991 [6], most of bacterial chitinases belong to family 18 of glycosyl hydrolases and can be further grouped into three subcategories from A to C, on the basis of the amino acid sequence similarity of their catalytic domains, the modular structure and enzymatic activities [2, 5].

The growing interest for chitinolytic enzymes in biotechnology is based on their wide range of applications, from medical purposes and waste recycling, to the production of single cell proteins, fungal protoplasts and biofuels. Additionally, chitinases can be employed as environment-sustainable biocontrol agents of chitin-containing plant pathogenic fungi and

insects pests, as well as for the industrial production of value-added chitin derivatives, including chitooligosaccharides, chitosan and glucosamines, with high pharmaceutical and nutritional potential [2]. Microorganisms, the primary degraders of chitin in the environment, are a particularly rich source of valuable chitin-modifying enzymes [3]. Indeed, conventional molecular and functional screening approaches have been employed for the identification of bacterial chitinase genes in different environmental samples, within both aquatic and soil habitats [7, 8]. Microbial unculturability under standard laboratory conditions, however, greatly hinders the number of microorganisms and, therefore, of microbial genes and enzymes, that may be identified by traditional approaches. Metagenomics circumvents the need of microbial cultivation and hence represents a powerful and promising tool for the identification of novel valuable biocatalysts otherwise encrypted in natural microbial communities [9]. Nevertheless, to date only a few investigations have focussed on metagenomic approaches for discovering innovative chitinase sequences [10-14] and in most of these studies the protein product was not characterised or even not expressed and purified. It is in fact commonly reported that the current persistent bottleneck occurring in the metagenomic exploitation is the heterologous expression of genes of unknown origin in the commonly used microbial hosts [9].

In a previous study we investigated the construction and functional screening of a suppressive soil metagenomic library, which led to the identification of a novel bacterial chitobiosidase, named Chi18H8, endowed with an interesting antifungal activity against several important crop pathogens [15]. Due to the extremely low production and purification yield (21 μg/g cells, from *Escherichia coli* BL21(DE3)/pGEX-6P-3::*chi18H8*), it was not possible to proceed with its complete biochemical characterisation and its evaluation as a potential antifungal biocontrol agent. In this paper, we report on the development of a process for the high-level production of Chi18H8 and its high-yield purification from *E. coli* inclusion bodies (IBs). Thanks to this work, the Chi18H8 biochemical and functional characterisation was completed.

Materials and methods

chi18H8 cDNA sub-cloning

E. coli DH5α, used for the cloning procedures, was purchased from Invitrogen-Life Technology, Carlsbad, USA. The cDNA encoding for the chitobiosidase Chi18H8 [15] was cloned into the pET24b(+) expression plasmid (kanamycin resistance; Novagen Inc., Madison, USA) using *Eco*RI and *Xho*I restriction sites. This plasmid allows a polyhistidine Tag (His_6-Tag) addition at the C-terminus of the protein. Cloning and transformation procedures were controlled by DNA sequencing (BMR Genomics, Padua, Italy). The construct was finally transformed into *E. coli* BL21 Star™(DE3) (Invitrogen-Life Technology, Carlsbad, USA). Recombinant *E. coli* strains were maintained on Luria Broth (Miller's modification) agar plate (LB: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl. For solidification, 15 g/L agar were added) supplemented with 50 µg/mL kanamycin.

Chi18H8 expression

Protein expression was carried out in the following media, supplemented with 50 µg/mL kanamycin: LB; terrific broth (TB: 12 g/L tryptone, 24 g/L yeast extract, 9.4 g/L K_2HPO_4, 2.2 g/L KH_2PO_4, 8 g/L glycerol); super broth (SB: 32 g/L tryptone, 20 g/L yeast extract, 5 g/L NaCl); autoinduction media A and B. Autoinduction medium A composition was based on [16]. The composition of autoinduction medium B was as follows: 10 g/L tryptone, 5 g/L yeast extract, 3.3 g/L $(NH_4)_2SO_4$, 6.8 g/L KH_2PO_4, 7.1 g/L Na_2HPO_4, 0.5 g/L glucose, 2 g/L α-lactose, 0.15 g/L $MgSO_4$, 2 mg/L $CaCl_2$, 2 mg/L $MnSO_4$ x H_2O, 2 mg/L $ZnSO_4$, 2 mg/L $CoCl_2$, 2 mg/L $CuCl_2$ x 2 H_2O, 2 mg/L $NiCl_2$, 2 mg/L NH_4MoO_4, 2 mg/L $FeCl_3$. Trace element ($MgSO_4$, $CaCl_2$, $MnSO_4$ x H_2O, $ZnSO_4$, $CoCl_2$, $CuCl_2$ x 2 H_2O, $NiCl_2$, NH_4MoO_4, $FeCl_3$) stock solutions were sterilised by filtration (0.2 µm) and stored at 4 °C. All reagents were from Sigma-Aldrich, St Louis, USA, unless otherwise stated.

Starter cultures were prepared from a single recombinant *E. coli* colony in 10 mL LB medium supplemented with 50 µg/mL kanamycin, grown overnight (O.N.) at 37 °C and 200 revolutions per minute (rpm). Baffled 300 mL Erlenmeyer flasks containing 50 mL of the different media were inoculated with the starter culture (initial OD_{600nm} = 0.1) and further incubated as before. For LB, TB and SB media, protein expression was induced by adding 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG) to cells at different growth phases (early- or

late-exponential growth phases), as determined by growth curve construction. After induction, the cells were cultured at various temperatures (37 °C, 25 °C or 20 °C) at 200 rpm. Cells were harvested at regular time intervals by centrifugation (1900 x *g* for 30 min at 4 °C). Supernatants (*i.e.* the cell-free fermentation broths) were treated with 10% (v/v) trichloroacetic acid. Cell pellets were instead sonicated on ice (3-5 cycles of 30 s each, with a 30-s interval, using a Branson Sonifier 250, Danbury, USA) in phosphate buffer saline (PBS) pH 7.3 (140 mM NaCl, 2.7 mM KCl, 10.1 mM Na_2HPO_4, 1.8 mM KH_2PO_4) containing 10 µg/mL deoxyribonuclease (DNase), 0.19 mg/mL phenylmethylsulfonylfluoride (PMSF) and 0.7 mg/mL pepstatin. Soluble and insoluble fractions were then separated by centrifugation at 20000 x *g* for 1 h at 4 °C. Insoluble fractions (containing membrane and IBs) were re-suspended in a volume of PBS equal to the corresponding cytoplasmic soluble fraction (2 mL/g cells) for successive analyses. Protein concentration was determined by the Biuret assay [17].

Chi18H8 solubilisation from IBs and purification

For Chi18H8 solubilisation, *E. coli* BL21 Star™(DE3)/pET24b(+)::*chi18H8* cells were grown O.N. at 37 °C and 200 rpm in 300 mL baffled Erlenmeyer flasks containing 80 mL of LB and 50 µg/mL kanamycin. Starter cultures (initial OD_{600nm} = 0.1) were inoculated in 2 L Erlenmeyer flasks with 750 mL LB medium with the selective antibiotic, incubated at 37 °C and 200 rpm. IPTG was added to a final concentration of 0.4 mM during early exponential growth phase (OD_{600nm} ~0.6). Cells were harvested after O.N. incubation at 20 °C and 200 rpm and washed with sodium chloride-tris-EDTA (STE buffer: 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 100 mM NaCl). Protein solubilisation from IBs was achieved with a protocol modified from [18]. Briefly, *E. coli* cells were re-suspended in 50 mM Tris-HCl pH 8.0, 25% (w/v) sucrose, 1 mM EDTA (5 mL/g cells) and incubated for 30 min at room temperature under vigorous shaking. After sonication on ice (6 cycles of 30 s each, with a 30-s interval), 5 mL/g cells of 0.2 M NaCl, 1% (v/v) sodium deoxycholate (DOC) and 1% (v/v) Nonidet P-40 were added. The sample was further incubated as above and centrifuged (20000 x *g* at 4 °C for 30 min). The pellet was washed with 1% (v/v) Triton X-100 and 1 mM EDTA (10 mL/g cells), followed by centrifugation at 12000 x *g* at 4 °C for 10 min; the procedure was repeated twice. The IBs were then washed with 10 mL/g cells deionised water and stored O.N. at –20 °C. Finally, the pellet was re-suspended in 10 mM lactic acid (10 mL/g cells) and incubated at 37 °C and 200

rpm for 5 h. Non-solubilised material was removed by centrifugation at 1900 x *g* at 4 °C for 5 min. The solubilised protein was then dialysed O.N. against 100 mM sodium acetate buffer pH 5.0 or 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 5.6.

Further purification of Chi18H8 was achieved by a negative purification method with hydrophobic interaction chromatography (HIC). Samples of the enzyme suspended in 100 mM HEPES pH 5.6 were loaded onto the weak anionic exchanger Diaion WA11 resin (Resindion s.r.l., Milan, Italy) used in HIC conditions. The resin has a polyacrylate matrix, tertiary amine anion exchanger, particle size of 0.4-0.6 mm and a loading capacity of 1.4 meq/mL. Pilot experiments were performed with 1 mL resin (wet volume) loaded on 3 mL columns; columns were gravity operated at room temperature. For large scale purification, a 4.5 cm diameter Amicon column was loaded with 40 mL of WA11 resin and was operated at a flow of 20 mL/min, with a BÜCHI Pump Manager C-615 (BÜCHI, Oldham, UK) coupled with: a Pump Module C-605; a Knauer variable wavelength UV detector; and a Fraction Collector BÜCHI 684. The resin was activated with methanol/water 1:1 and then equilibrated with 100 mM HEPES pH 5.6; proteins were separated with an isocratic method using 100 mM HEPES pH 5.6/ethanol 50% (v/v).

Chi18H8 concentration was estimated using the theoretical extinction coefficient at 280 nm (77015 $M^{-1}cm^{-1}$), based on the amino acid sequence of the protein.

Scale up in 3 L and 30 L bioreactors

300 mL flask cultures of recombinant *E. coli* cells, grown O.N. in LB medium supplemented with 50 µg/mL kanamycin, were used to inoculate (initial OD_{600nm} = 0.1) a 3 L P-100 Applikon glass reactor (height 25 cm, diameter 13 cm) equipped with a AD1030 biocontroller and AD1032 motor, containing 2 L LB and kanamycin (50 µg/mL). Cultivation in fermenter was carried out at 37 °C, with stirring at 500 rpm (corresponding to 1.17 m/s of tip speed) and 2 L/min aeration rate. Dissolved oxygen (measured as % of the initial pO_2 value) and pH value of the culture broths were monitored respectively using an Ingold polarographic oxygen electrode and a pH meter. Foam production was controlled by adding Antifoam SE-15 (Sigma-Aldrich, St Louis, USA) through an antifoam sensor. When the cell density reached an OD at 600 nm of 0.6, recombinant protein production was induced by adding 0.4 mM IPTG. Cultivation was then continued O.N. at 20 °C and 500 rpm.

For the fermentation at the 30 L scale, the vegetative inoculum was prepared by inoculating fresh *E. coli* BL21 Star™(DE3)/pET24b(+)::*chi18H8* slants in ten 1 L baffled flasks, each containing 300 mL LB supplemented with 50 µg/mL kanamycin and grown O.N. at 37 °C and 220 rpm. 3 L of pre-culture were used to inoculate a 30 L stirred fermenter, containing 27 L of selective LB medium. The culture was grown at 37 °C and 300 rpm, with constant 1.0 vvm aeration rate and pressure at 0.5 Bar, up to an OD_{600nm} ~0.6 and then induced with 0.4 mM IPTG. Growth was allowed for further 16 h after induction at 20 °C.

SDS-PAGE electrophoresis and zymogram analysis

Proteins samples were separated by sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS-PAGE) with 12% (w/v) polyacrylamide gels, using a Tris-glicine system and Coomassie brilliant blue R-250 staining as described in [19]. Chitobiosidase production was estimated through densitometric analysis of SDS gel bands with the software Quantity One (Bio-Rad Laboratories, Hercules, USA) and His_6-glycine oxidase (His_6-GO) from *Bacillus subtilis,* gently provided by Loredano Pollegioni, University of Insubria [20], as standard. The molecular weight markers were from GE-Healthcare Sciences, Little-Chalfont, UK.

Chitinolytic activity was detected through zymogram analysis using 10% (w/v) separating polyacrylamide gels containing 0.7 mg/mL carboxymethyl-chitin-remazol brilliant violet (CM-chitin-RBV, Loewe Biochemica, Sauerlach, Germany) [15, 21].

Chi18H8 activity measurement

Chi18H8 activity was assayed with the fluorimetric chitooligosaccharide analogues 4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide (4-MU-GlcNAc), 4-methylumbelliferyl *N,N'*-diacetyl-β-D-chitobioside (4-MU-$(GlcNAc)_2$) and 4-methylumbelliferyl *N,N',N''*-triacetyl-β-D-chitotrioside (4-MU-$(GlcNAc)_3$) as substrates as reported in [15]. One unit (U) of Chi18H8 activity was defined as the amount of enzyme required for the release of 1 µmole of 4-MU per min at 37 °C [15, 22].

Chitinolytic activity was determined also on colloidal chitin by the colorimetric method described by Anthon and Barrett [23] adapted to enzymatic hydrolysis. Colloidal chitin was prepared by the method reported in [24], starting from chitin flakes from shrimp shells (Sigma-Aldrich, St Louis, USA). 250 µL of protein sample were added to an equal volume of 10 g/L colloidal chitin, and the mixture was incubated at 37 °C for 1 h. The reaction was

quenched by boiling for 5 min and then centrifuged (20000 x *g*, 25 °C, 15 min); 200 µL of the supernatant were mixed with equal volumes of 0.5 M NaOH and MBTH reagent [23]. After 15-min incubation at 80 °C, 400 µL of a solution containing 0.5% (w/v) $FeNH_4(SO_4)_2$ x 12 H_2O, 0.5% (w/v) sulfamic acid and 0.25 M HCl were added and allowed to cool to room temperature. Subsequently, 1 mL water was added and absorbance at 620 nm (A620) determined. The released reducing sugars were estimated by comparison of A620 to a standard curve prepared with GlcNAc concentrations from 0 to 600 µM. One U of Chi18H8 activity was defined as the amount of enzyme that released 1 µmol/mL GlcNAc per mL per h at 37 °C.

Effect of pH and temperature on the enzyme activity

The optimum pH for Chi18H8 activity on the substrate 4-MU-$(GlcNAc)_2$ was determined with the fluorimetric assay above described with the following buffers (each 100 mM): glycine-HCl (pH 3.0), sodium acetate (pH 4.0 and 5.0), sodium phosphate (pH 6.0 and 7.0), Tris-HCl (pH 8.0) and sodium pyrophosphate (NaPPi, pH 9.0). The optimum temperature was calculated with the same fluorescent assay, incubating the reaction mixture at various temperatures (from 5 to 70 °C). Moreover, the influence of pH on Chi18H8 activity on CM-chitin-RBV was verified with zymogram analysis as previously illustrated, equilibrating the gel in the above listed buffers at different pHs. Long-term stability of the enzyme was tested by pre-incubating the chitobiosidase at 30 °C and at different pHs (in sodium acetate pH 5.0 and sodium phosphate pH 6.0 and 7.0, all 100 mM). The residual chitinolytic activity was fluorimetrically assayed on 4-MU-$(GlcNAc)_2$ from 0 to 144 h, incubating the reaction mixture at the same pH at which the enzyme was pre-treated.

Effect of metal ions, alcohols, detergents and other compounds on the enzyme activity

The effect of metal ions [Ca^{2+} ($CaCl_2$ x 2 H_2O), Cu^{2+} ($CuCl_2$ x 2 H_2O), Fe^{3+} ($FeCl_3$ x 6 H_2O), K^+ (KCl), Mg^{2+} ($MgCl_2$ x 6 H_2O), Mn^{2+} ($MnCl_2$ x 4 H_2O), Ni^{2+} ($NiCl_2$ x 6 H_2O), NH_4 (NH_4Cl), Zn^{2+} ($ZnCl_2$), Co^{2+} ($CoCl_2$ x 6 H_2O)], enzyme inhibitors [dithiothreitol (DTT), β-mercaptoethanol], the chelating agent EDTA, detergents [SDS, Triton X-100, Tween-20, DOC, *N*-lauroylsarcosine (NLS), Nonidet-40], organic solvents [ethanol, methanol, propanol, dimethyl sulfoxide (DMSO)], sugars [GlcNAc, chitobiose] and increasing concentrations of salt (up to 2 M NaCl) on Chi18H8 activity was investigated by adding each compound to the fluorimetric assay

mixture. Final assayed concentrations were: 20 mM for each metal ions and EDTA, 10 mM for sugars, 10% (v/v) for organic solvents, 5% (v/v) for β-mercaptoethanol and DTT, 1% (w/v) for detergents. Inhibition was calculated in percentage to the activity recorded in the absence of inhibitors.

Circular dichroism measurement

Far-UV spectra were recorded at 15 °C with a Jasco J-715 polarimeter (Jasco, Easton, USA) equipped with a temperature programmer, in the 195-250 nm wavelength range. Protein samples of 0.1 mg/mL in 10 mM HEPES pH 5.4 were employed, and the measurements corrected for buffer contribution. Secondary structure fractions were calculated from deconvolution of the CD spectra with the program k2d3 (http://k2d3.ogic.ca//index.html) [25].

Results

Recombinant protein expression

chi18H8 cDNA (G+C ratio 64.4%) was cloned into the pET24b(+) expression plasmid in *E. coli* BL21 Star™(DE3) cells, under the control of the IPTG-inducible *T7* promoter. This construct encodes for a 425-amino acid protein carrying an additional histidine examer-Tag (His_6-Tag) at the C-terminus, with a predicted molecular mass of 46.78 kDa and an isoelectric point of 7.82.

Chi18H8 expression was at first investigated in LB medium, adding 0.4 mM IPTG during early exponential growth phase (OD_{600nm} ~0.6) and harvesting the cells at regular intervals, from 0 h to O.N. incubation at 37 °C. Under these conditions, a protein band at the expected molecular mass for the recombinant chitobiosidase was detectable in the insoluble fractions collected after cell sonication, being only faintly visible in the corresponding soluble fractions (Figure 1). Consistently, the same band was not detected in the negative control (*i.e., E. coli* cells harbouring the pET24b(+) vector without an insert). As reported in Figure S1A, Appendix 1 Supplementary Materials, Chi18H8 production in the insoluble fraction slightly increased with time, reaching 2.7 mg/g cells (corresponding to 32 mg/L fermentation broth) after O.N. incubation. Indeed traces of chitobiosidase activity were detectable only into the

soluble fractions, indicating that the recombinant protein loses its biological activity when accumulated into IBs. The reduction of incubation temperature to 25 and 20 °C increased the production of the recombinant protein into the insoluble fraction (up to *ca.* 10 mg/g cells, corresponding to 55 mg/L fermentation broth after O.N. incubation) (Figure S1B&C, Appendix 1 Supplementary Materials), but yet no clear band at the expected molecular mass was detectable by SDS-PAGE into the soluble cytoplasmic fraction (data not shown). In these conditions, the sensitive fluorimetric assay revealed more chitobiosidase activity into the soluble cytoplasmic fraction than into the insoluble fraction (Figure S1B&C, Appendix 1 Supplementary Materials), indicating that most of the enzyme is packed into IBs as inactive. Even lower levels of Chi18H8 activity were observed in the other tested growth conditions, for example by inducing protein expression during late exponential growth phase (OD_{600nm} ~2.5) or replacing the LB medium with the richer TB and SB media or with the two defined autoinduction broths (data not shown).

Chi18H8 solubilisation from IBs and purification

Since when recombinant *E. coli* cells were incubated at 20-25 °C, some residual chitinolytic activity was detectable from the insoluble fraction (Figure S1B&C, Appendix 1 Supplementary Materials), we thought that it could be worthy trying Chi18H8 recovery from IBs. Thus, a screening of different protocols for protein solubilisation was applied to IBs collected by centrifugation. Chi18H8 recovery was initially attempted by using strong denaturing agents like urea, guanidium hydrochloride or guanidium thiocyanate, combined with detergents, reducing agents or chelating agents, followed by refolding *via* dialysis [26]. The yield of protein solubilisation increased with the denaturing agent concentration. However, refolding steps often resulted into protein aggregation/precipitation and finally the protein samples obtained after the dialysis were biologically inactive (data not shown). Since protein aggregation during refolding could be due to the exposure of hydrophobic amino acid stretches caused by high concentration of chaotropic agents [27, 28], milder solubilisation conditions of IBs were attempted to better preserve the existing native-like secondary structure and reduce the chance of aggregation during the solubilisation process [28, 29]. Protocols thus were based on the use of detergents [30], or alkyl alcohols [29], or extremely basic [31] or acid pH [18]. Only the last one (*i.e.*, solubilisation under extremely acid conditions using HCl or lactic acid) allowed a high-yield recovery of active Chi18H8.

Table S1 in Appendix S2 of Supplementary Materials shows that the highest solubilisation yield (>90%) was achieved with IBs re-suspension in 10 mM HCl. However, when IBs were solubilised in 10 mM lactic acid, the recovered protein had a higher specific activity, up to 40.7 U/mg. The electrophoretic analysis of the solubilised protein revealed that it was about 70% pure (Figure 2, panel A and Table 1).

To further purify the recombinant chitobiosidase, different precipitation methods and chromatographic procedures were evaluated. The addition of different salts (NaCl, KCl, $MgCl_2$, $NH_4(SO_4)_2$), different concentrations of ethanol as well as the chitinase substrate CM-chitin-RBV, led to the co-precipitation of the enzyme and contaminating proteins, showing no apparent advantage for a selective precipitation (data not shown). Purification of the chitobiosidase by ion exchange chromatography (IEC), affinity chromatography (AC) and hydrophobic interaction chromatography (HIC) was then evaluated, using resins and conditions reported in Table S2 in Appendix S2 of Supplementary Materials. In most of the conditions, the recombinant enzyme did not bound to the resins, in some cases it precipitated and, in general, it was inactive if the pH was increased above 7.0. Finally, with the weak anionic exchanger WA11 resin used in HIC conditions, most of the contaminating proteins were retained by the resin, whereas Chi18H8 was recovered in the flow through, allowing a negative purification of the enzyme. Protein recovery was estimated to be of about 70% (Table 1). The analysis of the purified protein by SDS-PAGE revealed that it was >95% pure (Figure 2, panel A). Zymogram analysis on carboxymethyl chitin confirmed the chitinolytic activity of the enzyme (Figure 2, panel B).

Chi18H8 production in 3 L bench- and 30 L industrial-bioreactors

Chi18H8 production was scaled-up in 3 L and 30 L bioreactors, where *E. coli* BL21 Star™(DE3)/pET24b(+)::*chi18H8* cells were grown in the conditions previously optimised at flask level. Figure 3 shows, as example, the time course of growth and Chi18H8 production in the 3 L bench-fermenter. Maximum protein production was 9.09 mg/g cells (corresponding to 75 mg/L) and 12.9 mg/g cells (equal to 80 mg/L) for the 3 L and 30 L bioreactors respectively, higher than the one obtained at flask level (Table 2). The subsequent IBs solubilisation with 10 mM lactic acid allowed in both cases the recovery of bioactive Chi18H8, with a yield >65% and a specific activity on 4-MU-$(GlcNAc)_2$ of 52.17 U/mg and 48.4

U/mg, respectively (Table 2). The SDS-PAGE analysis of the solubilised proteins revealed that they were about 70% pure, similarly to the results obtained in flasks (data not shown).
Also the negative method of purification by HIC with the weak anionic exchanger WA11 resin turned out to be scalable and suitable for the production of Chi18H8: 135 mL of solubilised Chi18H8 in 100 mM HEPES pH 5.6 were loaded onto a 40 mL WA11 resin and separated with an isocratic method, as described in the Material and Method section. The recovery yield was approximately 70% and the final purity rose to >95%. These results demonstrate that the protocols developed for Chi18H8 production, solubilisation from IBs and purification could be easily scaled-up.

Chi18H8 characterisation

Substrate specificity

Chi18H8 substrate specificity was measured using three different-length analogues of natural chitooligosaccharides and colloidal chitin. Chi18H8 showed a prevalent chitobiosidase activity (40.7 U/mg), a weaker endochitinase activity (7.5 U/mg) and no β-*N*-acetyl-glucosaminidase activity. Moreover, the enzyme was able to hydrolyse colloidal chitin, with an estimated activity of 0.41 U/mg.

pH and temperature profile

The effect of pH on Chi18H8 was calculated using 4-MU-$(GlcNAc)_2$ as substrate. The optimum pH for the enzyme was between 5.0 and 6.0; only 15.3% of the chitinolytic activity was maintained at pH 7.0, whereas no activity was detected at pH 3.0 and 4.0 and above pH 8.0 (Figure 4, panel A). The effect of pH on Chi18H8 activity on CM-chitin-RBV was also verified by equilibrating the zymogram gel in buffers at different pHs: the enzyme was able to hydrolyse the substrate only between pH 5.0 and 7.0, being any degradation halo absent at other pHs (Figure 4 panel C).
The optimum temperature on 4-MU-$(GlcNAc)_2$ was between 35 °C and 40 °C. More than 55% and 10% of the activity was retained at 25 and 15 °C, respectively (Figure 4, panel B). The long-term stability of Chi18H8 was investigated by measuring the residual activity on the substrate 4-MU-$(GlcNAc)_2$ up to 144 h of pre-incubation of the enzyme in buffers at different pHs (5.0, 6.0 and 7.0) at the optimal temperature of 30 °C. When the recombinant chitobiosidase was pre-incubated at pH 7.0, a complete loss of activity was registered within

24 h. At pH 6.0, the activity dramatically decreased to 5% of the initial one within 144 h. Finally, at pH 5.0 the chitobiosidase confirmed to be highly stable, retaining about 70% activity after 144 h (Figure 4, panel D).

Influence of various compounds on Chi18H8 activity

The effect of several elements was tested on Chi18H8 enzyme activity (Table 3). Among the metal ions, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+}, Zn^{2+} as well as the monovalent cation NH_4^+ reduced the hydrolytic activity of Chi18H8, with the strongest inhibition being caused by Cu^{2+} and Fe^{2+}. On the contrary, Ca^{2+}, K^+ and Co^{2+} slightly increased the activity of the enzyme. An inhibitory effect was shown by the chelating agent EDTA, suggesting that Chi18H8 could be a metalloenzyme and that metal ions may be necessary for its catalytic activity. The enzyme inhibitors β-mercaptoethanol and DTT strongly reduced the chitinase activity. Among the detergents, SDS, NLS, Tween-20 and Triton X-100 showed a strong inhibitory effect, Nonidet P-40 had only a slight effect on the chitinolytic activity of Chi18H8, while DOC increased it. Chi18H8 was stable in the presence of organic solvents (10% (v/v) final concentration), being its activity only slightly reduced by them. NAG did not influence the hydrolytic activity of the enzyme, while 10 mM chitobiose slightly inhibited it. Chi18H8 activity was finally evaluated in the presence of increasing concentration of NaCl, up to 2 M: as show in Table 3, the enzyme is inhibited by high salt concentrations, retaining only 4.35% of its activity when incubated with 2 M NaCl.

Circular dichroism

Figure 5 shows the CD spectra in the far-UV region of the purified chitinase. The spectrum indicated a prevalence of β-sheets (~36.2%) and ~8.3% of α-helices (solid line). The addition of Ca^{2+} (dashed line) or Zn^{2+} (dotted line), respectively enhancer and inhibitor of Chi18H8 activity, altered the spectrum of the protein: both ions did non influence significantly the composition of β-sheets (~36.4 and ~35.9% with Ca^{2+} and Zn^{2+}, respectively), but reduced the α-helix content of the protein (~3.1 and ~2.6% with Ca^{2+} and Zn^{2+}, respectively).

Discussion

Metagenomic analysis is a powerful tool for accessing the genetic diversity encrypted in natural environments and for the discovery of novel chitinolytic enzymes useful for biotechnological applications [9, 32]. In our previous work, the screening of a fosmid metagenomic library of a suppressive soil for clubroot disease of cabbage resulted in the identification of a novel GH18 chitinase, called Chi18H8, with an interesting inhibitory activity against common phytopathogen fungi [15]. The complete characterisation of this enzyme was, however, hampered by the low purification yield: only 21 µg protein/g cells were purified from the soluble cytoplasmic fraction of the recombinant *E. coli* BL21 (DE3)/pGEX-6P-3::*chi18H8* strain, where the protein was mainly produced as inactive into the IBs. Indeed, the development of suitable expression systems for the over-expression of proteins, as well as the identification of efficient recovery and purification protocols, represent the major bottlenecks of a metagenomic approach.

In this work we investigated Chi18H8 production in *E. coli* BL21 Star^{TM}(DE3)/pET24b(+)::*chi18H8.* Also in this system, Chi18H8 was mainly accumulated into the IBs, due to the high concentration of the produced recombinant protein into the cytoplasm (in the best conditions, more than 9 mg protein/g cells) and probably also due to the presence of an hydrophobic putative signal peptide mediating protein aggregation and/or membrane sticking. Soluble Chi18H8 production in the cytoplasmic fraction was slightly increased by lowering the incubation temperature after induction, as demonstrated for other recombinant proteins [33, 34], but the yield was too low for sustaining a scalable production process. Protein production was obtained also in two auto-induction broths [16], but even in these cases more than 95% of the produced recombinant protein (up to 0.4 mg/g cells and 2.5 mg/g cells for autoinduction media A and B, respectively) was accumulated as inactive into IBs.

Protein recovery from IBs could be tedious, time-consuming and not always effective in yielding native folded and active proteins [35]. Conversely, the protocol herein developed and optimised for Chi18H8 solubilisation is quite simple, being composed of few washing steps with detergents, followed by freezing and thawing, and re-suspension in lactic acid. Even if at such acid pH (around 2.5) Chi18H8 is not active, its refolding could be achieved by simple dialysis against appropriate buffers, like pH 5.0 sodium acetate. The solubilisation

yield obtained with this method is high (>80%) and the purity of the solubilised protein superior than 70%. Interestingly, Chi18H8 could not be recovered in an active form by using strong denaturing agents like urea or guanidium hydrochloride that are currently the most commonly employed agents for IBs solubilisation. Our work confirms that mild-solubilisation protocols, like the one herein developed, can be equally if not more efficient than those based on chaotropic agents. These mild methods preserve the existing native-like secondary structure of proteins and reduce the chance of protein re-aggregation during the refolding process [28, 29].

The IBs solubilisation procedure was then coupled with a further purification step based on hydrophobic-interaction chromatography, that represents one of the most employed chromatographic methods for industrial purification of peptides and proteins [36]; it is in fact user-friendly and easy to scale-up at relatively low cost and many resins of different chemical composition are commercially available. Chi18H8 did not bind to most of the tested resins in the pH conditions required for its activity and solubility. Nevertheless, thanks to the ability of the weak anionic exchanger WA11 resin to retain most of the contaminating proteins, it was possible to recover in a single step the recombinant chitinase with a purity superior to 95%.

Enzyme assays on chito-oligosaccharide analogues using the pure Chi18H8 confirmed its prevalent chitobiosidase activity, with optimum at acid pH and mesophilic temperatures [15]. Interestingly, the recombinant Chi18H8 was active not only on these synthetic analogues, commonly used for chitinolytic enzyme detection, but also on the complex substrate colloidal chitin. Moreover, preliminary experiments highlighted a prevalent endo-type hydrolytic activity also on soluble chitosan (Kjell Morten Vårum, Norvegian University of Science and Technology, Trondheim, Norway; unpublished results).

Chi18H8 is a metallo-enzyme, whose activity is significantly reduced in the presence of the chelating agent EDTA. Additionally, Chi18H8 activity is strongly inhibited by iron and copper and, to a minor extent, by manganese, nickel and zinc. Chitinolytic enzymes are usually particularly sensitive to iron, mercury and copper, even if some inhibitory effect has been recorded also for manganese, zinc, magnesium and calcium [37, 38]. Interestingly, Chi18H8 activity is increased by calcium, as already proved for some other chitinolytic enzymes [39, 40]. Ca^{2+} ions, by binding to the enzyme, are able to altering its secondary structure, as determined by circular dichroism analysis. Also the enhancing effect exerted by cobalt and

potassium on Chi18H8 activity, even if quite uncommon, has been already described for few chitinases (for cobalt [41, 42], for potassium [43]). The high solvent tolerance of Chi18H8 suggests a possible origin of the enzyme from a solvent-tolerant microorganism and could be worth of further exploration for its application in non-aqueous solutions.

Finally, production and purification of Chi18H8 was scaled up from shaken flasks to 3 L bench- and 30 L industrial pilot scale-bioreactors, giving comparable results in terms of recombinant *E. coli* biomass, Chi18H8 production, IBs solubilisation and negative HIC purification yields. The whole process (upstream and downstream) is robust and scalable and *ca.* 37 mg of pure active protein can be obtained from 1 litre of culture at the pilot scale of 30 L fermenter, supporting the further exploitation of the metagenome-sourced Chi18H8. The antifungal behaviour, long-term stability and solvent tolerance, make this enzyme an interesting and valuable candidate for both industrial and agricultural applications, worth of future investigation.

Acknowledgments

This research activity was funded by the EU FP7/2007-2011 Metaexplore project, grant agreement No. 222625 to Flavia Marinelli and Sara Sjöling. Miur fellowship and CIB contribution to Francesca Berini are also acknowledged. The authors thank E. Binda from the University of Insubria for technical assistance.

References

1. Bai Y, Eijsink VG, Kielak AM, van Veen JA, de Boer W: **Genomic comparison of chitinolytic enzyme systems from terrestrial and aquatic bacteria**. *Environ Microbiol* 2014.
2. Adrangi S, Faramarzi MA: **From bacteria to human: a journey into the world of chitinases**. *Biotechnol Adv* 2013, **31**(8):1786-1795.
3. Howard MB, Ekborg NA, Weiner RM, Hutcheson SW: **Detection and characterization of chitinases and other chitin-modifying enzymes**. *J Ind Microbiol Biotechnol* 2003, **30**(11):627-635.
4. Beier S, Bertilsson S: **Bacterial chitin degradation-mechanisms and ecophysiological strategies**. *Front Microbiol* 2013, **4**:149.
5. Dahiya N, Tewari R, Hoondal GS: **Biotechnological aspects of chitinolytic enzymes: a review**. *Appl Microbiol Biotechnol* 2006, **71**(6):773-782.
6. Henrissat B: **A classification of glycosyl hydrolases based on amino acid sequence similarities**. *Biochem J* 1991, **280 (Pt 2)**:309-316.
7. Bhattacharya D, Nagpure A, Gupta RK: **Bacterial chitinases: properties and potential**. *Crit Rev Biotechnol* 2007, **27**(1):21-28.
8. Li H, Greene LH: **Sequence and structural analysis of the chitinase insertion domain reveals two conserved motifs involved in chitin-binding**. *PLoS One* 2010, **5**(1):e8654.
9. Schmeisser C, Steele H, Streit WR: **Metagenomics, biotechnology with non-culturable microbes**. *Appl Microbiol Biotechnol* 2007, **75**(5):955-962.
10. LeCleir GR, Buchan A, Hollibaugh JT: **Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions**. *Appl Environ Microbiol* 2004, **70**(12):6977-6983.
11. Hjort K, Bergström M, Adesina MF, Jansson JK, Smalla K, Sjöling S: **Chitinase genes revealed and compared in bacterial isolates, DNA extracts and a metagenomic library from a phytopathogen-suppressive soil**. *FEMS Microbiol Ecol* 2010, **71**(2):197-207.
12. Jacquiod S, Demanèche S, Franqueville L, Ausec L, Xu Z, Delmont TO, Dunon V, Cagnon C, Mandic-Mulec I, Vogel TM *et al*: **Characterization of new bacterial catabolic genes and mobile genetic elements by high throughput genetic screening of a soil metagenomic library**. *J Biotechnol* 2014.
13. Bhuiyan FA, Nagata S, Ohnishi K: **Novel chitinase genes from metagenomic DNA prepared from marine sediments in southwest Japan**. *Pak J Biol Sci* 2011, **14**(3):204-211.
14. Stöveken J, Singh R, Kolkenbrock S, Zakrzewski M, Wibberg D, Eikmeyer FG, Pühler A, Schlüter A, Moerschbacher BM: **Successful heterologous expression of a novel chitinase identified by sequence analyses of the metagenome from a chitin-enriched soil sample**. *J Biotechnol* 2014.
15. Hjort K, Presti I, Elväng A, Marinelli F, Sjöling S: **Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics**. *Appl Microbiol Biotechnol* 2014, **98**(6):2819-2828.
16. Li Z, Kessler W, van den Heuvel J, Rinas U: **Simple defined autoinduction medium for high-level recombinant protein production using T7-based *Escherichia coli* expression systems**. *Appl Microbiol Biotechnol* 2011, **91**(4):1203-1213.
17. Gornall AG, Bardawill CJ, David MM: **Determination of serum proteins by means of the biuret reaction**. *J Biol Chem* 1949, **177**(2):751-766.
18. Okumura S, Saitoh H, Wasano N, Katayama H, Higuchi K, Mizuki E, Inouye K: **Efficient solubilization, activation, and purification of recombinant Cry45Aa of *Bacillus thuringiensis* expressed as inclusion bodies in *Escherichia coli***. *Protein Expr Purif* 2006, **47**(1):144-151.
19. Schägger H, von Jagow G: **Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa**. *Anal Biochem* 1987, **166**(2):368-379.

20. Job V, Marcone GL, Pilone MS, Pollegioni L: **Glycine oxidase from *Bacillus subtilis*. Characterization of a new flavoprotein**. *J Biol Chem* 2002, **277**(9):6985-6993.
21. Tikhonov VE, Lopez-Llorca LV, Salinas J, Jansson HB: **Purification and characterization of chitinases from the nematophagous fungi *Verticillium chlamydosporium* and *V. suchlasporium***. *Fungal Genet Biol* 2002, **35**(1):67-78.
22. McCreath KJ, Gooday GW: **A rapid and sensitive microassay for determination of chitinolytic activity** *J Microbiol Meth* 1992, **14**:229-227.
23. Anthon GE, Barrett DM: **Determination of reducing sugars with 3-methyl-2-benzothiazolinonehydrazone**. *Anal Biochem* 2002, **305**(2):287-289.
24. Hsu SC, Lockwood JL: **Powdered chitin agar as a selective medium for enumeration of actinomycetes in water and soil**. *Appl Microbiol* 1975, **29**(3):422-426.
25. Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C: **Prediction of protein secondary structure from circular dichroism using theoretically derived spectra**. *Proteins* 2011.
26. Volontè F, Piubelli L, Pollegioni L: **Optimizing HIV-1 protease production in *Escherichia coli* as fusion protein**. *Microb Cell Fact* 2011, **10**:53.
27. Huang C, Lin H, Yang X: **Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements**. *J Ind Microbiol Biot* 2012, **39**:383-399.
28. Singh SM, Panda AK: **Solubilization and refolding of bacterial inclusion body proteins**. *J Biosci Bioeng* 2005, **99**(4):303-310.
29. Singh SM, Sharma A, Upadhyay AK, Singh A, Garg LC, Panda AK: **Solubilization of inclusion body proteins using n-propanol and its refolding into bioactive form**. *Protein Expr Purif* 2012, **81**(1):75-82.
30. Tao H, Liu W, Simmons BN, Harris HK, Cox TC, Massiah MA: **Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS**. *Biotechniques* 2010, **48**(1):61-64.
31. Patra AK, Mukhopadhyay R, Mukhija R, Krishnan A, Garg LC, Panda AK: **Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from *Escherichia coli***. *Protein Expr Purif* 2000, **18**(2):182-192.
32. Simon C, Daniel R: **Metagenomic analyses: past and future trends**. *Appl Environ Microbiol* 2011, **77**(4):1153-1161.
33. Binda E, Marcone GL, Pollegioni L, Marinelli F: **Characterization of VanYn, a novel D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727**. *FEBS J* 2012, **279**(17):3203-3213.
34. Piubelli L, Campa M, Temporini C, Binda E, Mangione F, Amicosante M, Terreni M, Marinelli F, Pollegioni L: **Optimizing *Escherichia coli* as a protein expression platform to produce *Mycobacterium tuberculosis* immunogenic proteins**. *Microb Cell Fact* 2013, **12**:115.
35. Rosano GL, Ceccarelli EA: **Recombinant protein expression in *Escherichia coli*: advances and challenges**. *Front Microbiol* 2014, **5**:172.
36. Saraswat M, Musante L, Ravidá A, Shortt B, Byrne B, Holthofer H: **Preparative purification of recombinant proteins: current status and future trends**. *Biomed Res Int* 2013, **2013**:312709.
37. Barghini P, Moscatelli D, Garzillo AM, Crognale S, Fenice M: **High production of cold-tolerant chitinases on shrimp wastes in bench-top bioreactor by the Antarctic fungus *Lecanicillium muscarium* CCFEE 5003: bioprocess optimization and characterization of two main enzymes**. *Enzyme Microb Technol* 2013, **53**(5):331-338.
38. Deane EE, Whipps JM, Lynch JM, Peberdy JF: **The purification and characterization of a *Trichoderma harzianum* exochitinase**. *Biochim Biophys Acta* 1998, **1383**(1):101-110.
39. Stressmann M, Kitao S, Griffith M, Moresoli C, Bravo LA, Marangoni AG: **Calcium interacts with antifreeze proteins and chitinase from cold-acclimated winter rye**. *Plant Physiol* 2004, **135**(1):364-376.

40. García-Fraga B, da Silva AF, López-Seijas J, Sieiro C: **Functional expression and characterization of a chitinase from the marine archaeon *Halobacterium salinarum* CECT 395 in *Escherichia coli***. *Appl Microbiol Biotechnol* 2014, **98**(5):2133-2143.
41. Waghmare SR, Ghosh JS: **Study of thermostable chitinases from Oerskovia xanthineolytica NCIM 2839**. *Appl Microbiol Biotechnol* 2010, **86**(6):1849-1856.
42. Lee YS, Park IH, Yoo JS, Chung SY, Lee YC, Cho YS, Ahn SC, Kim CM, Choi YL: **Cloning, purification, and characterization of chitinase from *Bacillus* sp. DAU101**. *Bioresour Technol* 2007, **98**(14):2734-2741.
43. Aunpad R, Rice D, Sedelnikova S, Panbangred W: **Biochemical characterisation of two forms of halo- and thermo-tolerant chitinase C of *Salinivibrio costicola* expressed in *Escherichia coli***. *Ann Microbiol* 2007, **57**(2).

Figure legends

Figure 1. SDS-PAGE analysis of *E. coli* BL21 Star™(DE3) cells carrying pET24b(+) or pET24b(+)::*chi18H8* plasmids, harvested after O.N. incubation at 37 °C in LB medium. From *E. coli* BL21 Star™(DE3)/pET24b(+): soluble (lane 1) and insoluble (lane 3) fractions; from *E. coli* BL21 Star™(DE3)/pET24b(+)::*chi18H8*: soluble (lane 2) and insoluble (lane 4) fractions. In each lane, samples corresponding to 0.5 mL of cell culture were loaded. Std: reference protein, His_6-GO from *Bacillus subtilis* (2.5 μg, 42.66 kDa).

Figure 2. SDS-PAGE (A) and zymogram (B) analysis of Chi18H8. C.E.: crude extract, raw enzyme after solubilisation from IBs with 10 mM lactic acid, loaded onto WA11 resin. 1, 2 and 3: flow-through of purification. Std 1: SDS-PAGE reference protein, His_6-GO from *Bacillus subtilis* (5 μg, 42.66 kDa). Std 2: Zymogram reference protein, chitinases from *Trichoderma viride* (Sigma-Aldrich, St Louis, USA).

Figure 3. Chi18H8 production at 3 L bench-bioreactor scale. Panel A: Time course of pH (■, solid line), pO_2 (●, solid line), temperature (□, dashed line) and wet weight (▲, dotted line). Induction with 0.4 mM IPTG was performed 2 h after the inoculum when the OD_{600nm} reached the value of ~0.6, and cells were harvested after 0, 1, 2, 4 h and O.N. from induction. Panel B: Chi18H8 production (mg protein/g cells wet weight) into insoluble fractions (grey bars) was determined by densitometric analysis of proteins separated through SDS-PAGE. Chitobiosidase activity was measured by fluorimetric assay on 4-MU-$(GlcNAc)_2$, either in the soluble cytoplasmic fractions of cells (■) or in the insoluble fractions (membrane and IBs)(●).

Figure 4. Enzymatic properties of Chi18H8. Panel A: pH profile of Chi18H8, using 4-MU-$(GlcNAc)_2$ as substrate. Buffers used (final concentration 100 mM) were glycine-HCl (pH 3.0), sodium acetate (pH 4.0 and 5.0), sodium phosphate (pH 6.0 and 7.0), TrisHCl (pH 8.0) and sodium pyrophosphate (pH 9.0). Panel B: temperature influence on chitobiosidase activity on 4-MU-$(GlcNAc)_2$. Assays were performed in 100 mM sodium acetate buffer pH 5.0. Panel C: effect of pH on Chi18H8 activity on CM-chitin-RBV. The zymogram was incubated O.N. in the following buffers (final concentration 100 mM): glycine-HCl (pH 3.0), sodium acetate (pH

4.0 and 5.0), sodium phosphate (pH 6.0 and 7.0), TrisHCl (pH 8.0) and sodium pyrophosphate (pH 9.0). Panel D**:** enzymatic stability of Chi18H8. The enzyme was pre-treated in 100 mM sodium acetate pH 5.0 (■, solid line), sodium phosphate pH 6.0 (●, dashed line) or sodium phosphate pH 7.0 (▲, dotted line) for 144 h at 30 °C. At various time intervals, samples were withdrawn and the residual chitinolytic activity was fluorimetrically measured on the substrate 4-MU-$(GlcNAc)_2$. In panel A, B and D, the activity is expressed as a percentage of the initial recorded activity and the values represent the mean of three experiments (mean ± standard error).

Figure 5. Circular dichroism (CD) far-UV spectra of Chi18H8, alone (solid line) or with the addition of 20 mM Ca^{2+} (dashed line) and Zn^{2+} (dotted line). The concentration of Chi18H8 was 0.1 mg/mL.

Figures

Figure 1

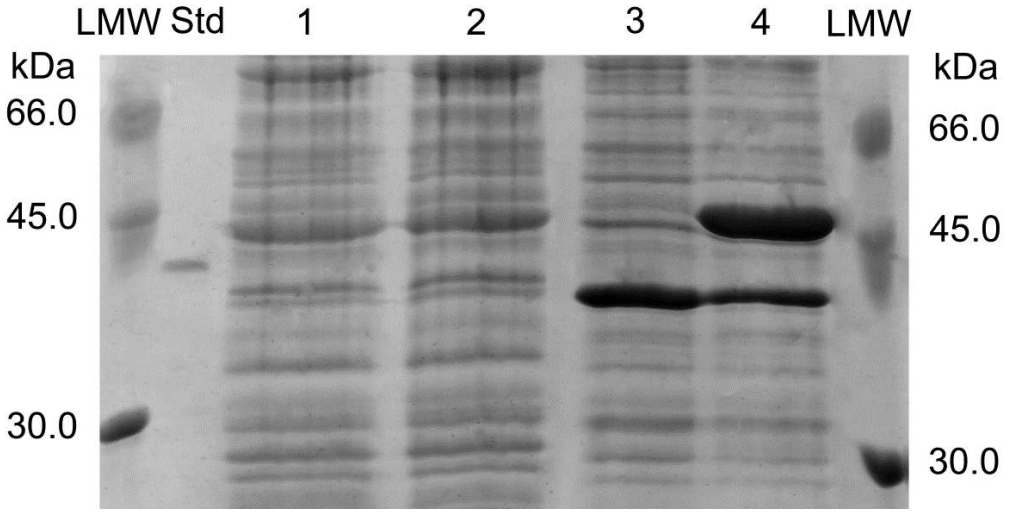


Figure 2

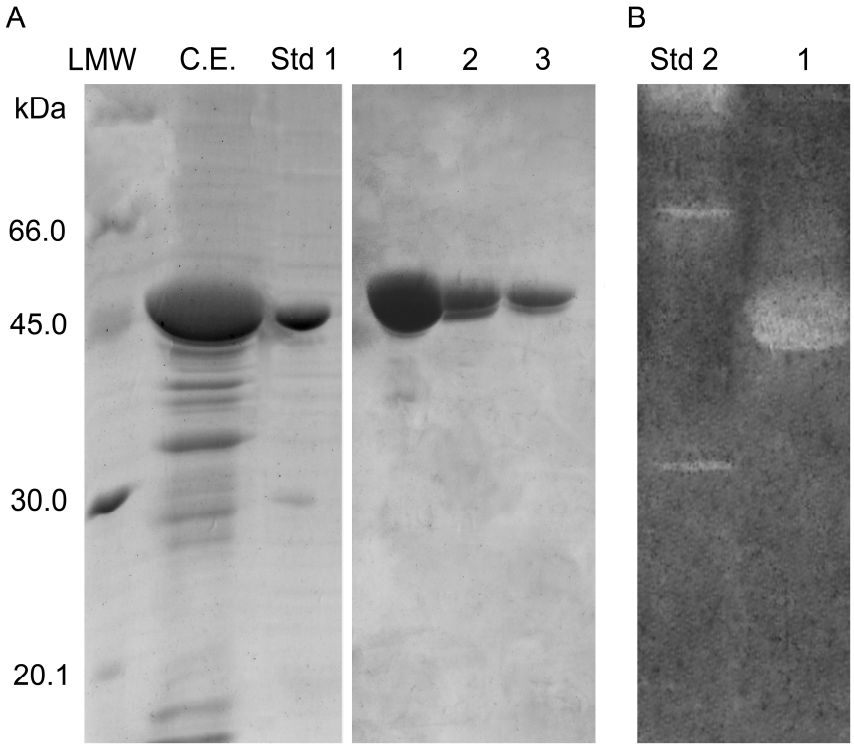


Figure 3

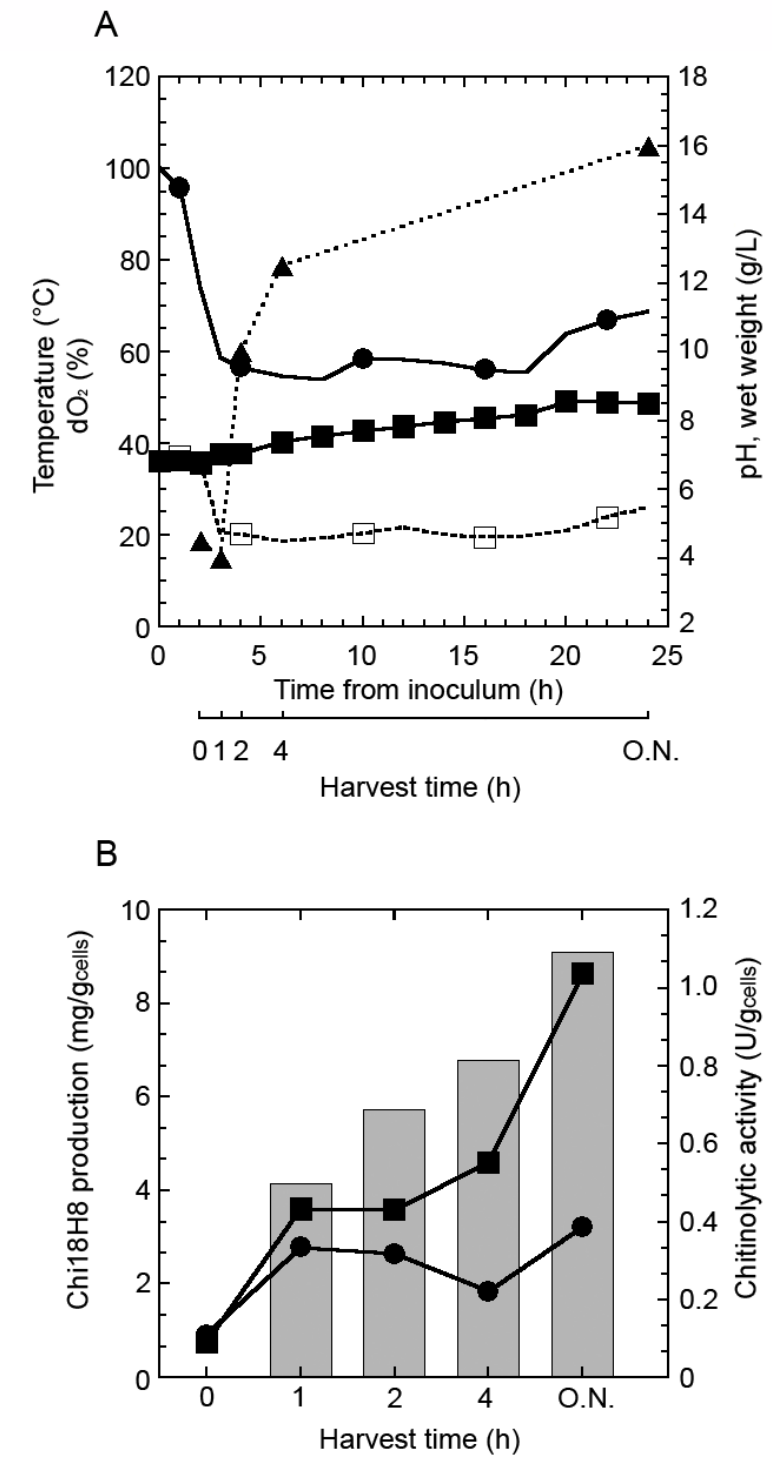

A
Temperature (°C)
dO2 (%)
pH, wet weight (g/L)
Time from inoculum (h)
0 1 2 4 O.N.
Harvest time (h)
B
Chi18H8 production (mg/gcells)
Chitinolytic activity (U/gcells)
Harvest time (h)

Figure 4

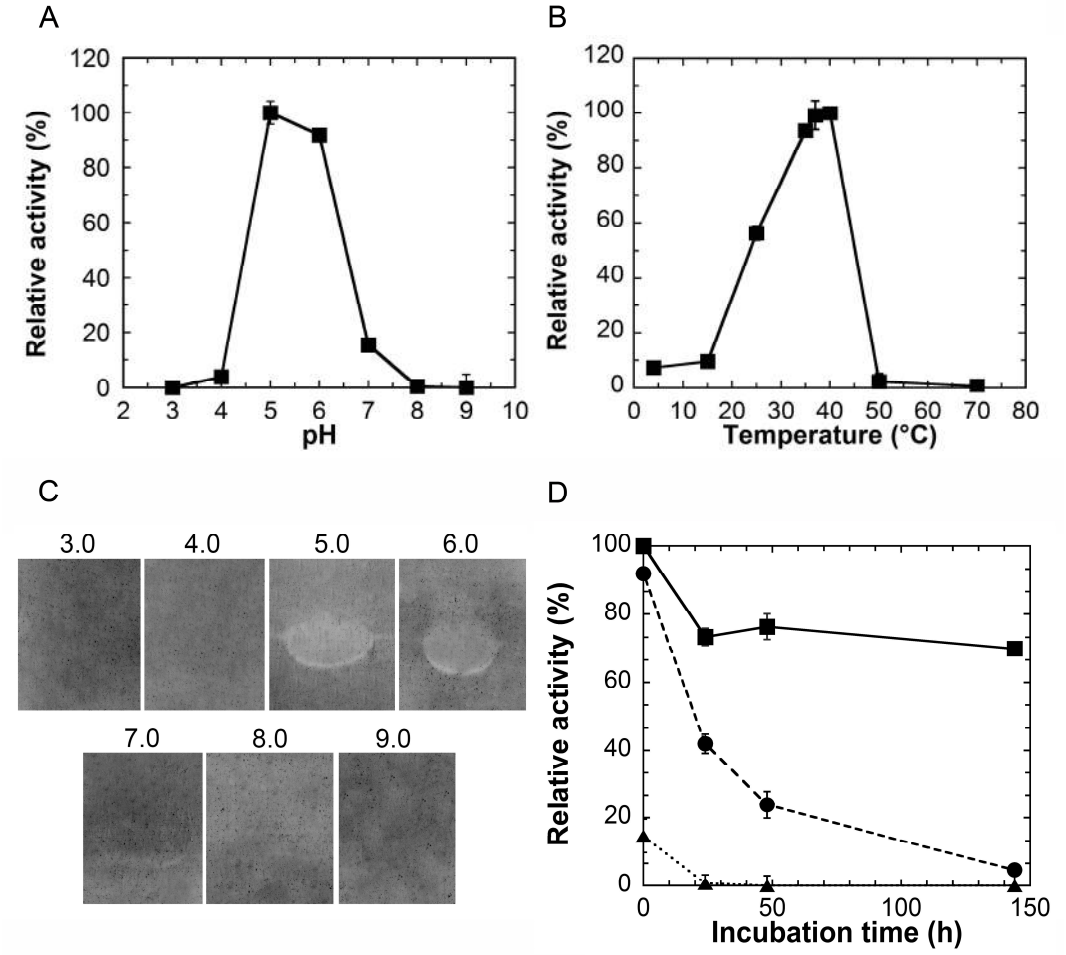
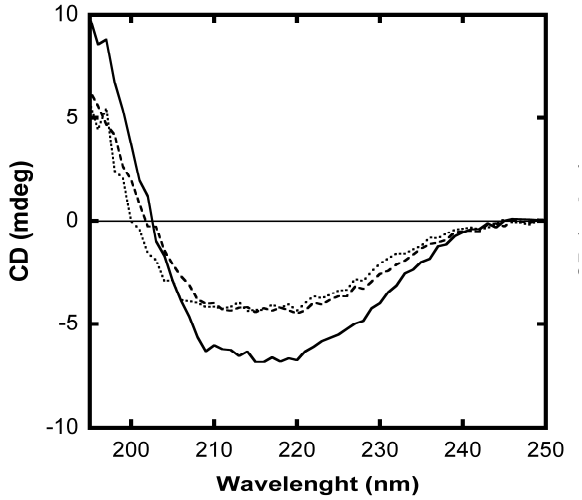


Figure 5



Tables

Table 1. Purification of Chi18H8. Enzymatic activity was measured by fluorimetric assay on 4-MU-$(GlcNAc)_2$

	Total protein concentration (mg/mL)	Volumetric activity (U/mL)	Specific activity ($U/mg_{total\ proteins}$)	Purification (fold)	Recovery (%)	Recovery (mg/g_{cells})
Inclusion bodies	6.25	0.94	0.15	1.00	100.00	9.38
Raw Chi18H8 solubilised from inclusion bodies	1.02	29.06	28.49	189.90	82.00	7.70
Purified Chi18H8	0.526	20.35	38.69	257.90	57.40	5.40

Table 2. Comparison of Chi18H8 production and solubilisation from IBs in *E. coli* BL21 Star™(DE3)/pET24b(+)::*chi18H8* cells grown in 2 L flasks, 3 L bench-bioreactor and 30 L industrial-bioreactor, induced in the early exponential growth phase (OD_{600nm} ~0.6) with 0.4 mM IPTG and harvested after O.N. incubation. Chi18H8 production (mg protein/g cells wet weight) into insoluble fractions and solubilisation yield were estimated by densitometric analysis of proteins separated through SDS-PAGE. Specific activity of the solubilised enzyme was measured by the fluorimetric assay on 4-MU-$(GlcNAc)_2$ as substrate.

	Cell weight (g_{cells}/L)	Chi18H8 production (mg/g_{cells})	Solubilisation yield (%)	Specific activity (U/mg)
Flask	4.0	9.38	> 80	40.7
3 L bench-bioreactor	5.3	9.1	> 65	52.1
30 L industrial-bioreactor	6.2	12.9	> 65	48.4

Table 3. Stability of Chi18H8 in the presence of different classes of compounds. The activity was measured on 4-MU-$(GlcNAc)_2$ as substrate, at 37 °C in 100 mM sodium acetate pH 5.0.

Compounds	**Final concentration**	**Relative activity (%)**
Control		100
Metal ions	mM	
Ca^{2+}	20	126.80 ± 0.83
Cu^{2+}	20	0.05 ± 0.00
Fe^{2+}	20	0.10 ± 0.88
K^{+}	20	124.22 ± 3.14
Mg^{2+}	20	86.43 ± 1.38
Mn^{2+}	20	49.91 ± 0.01
Ni^{2+}	20	64.32 ± 0.01
NH_4^{+}	20	68.02 ± 7.65
Zn^{2+}	20	16.59 ± 2.03
Co^{2+}	20	119.60 ± 1.48
EDTA	20	53.64 ± 0.01
Enzyme inhibitors	% (v/v)	
B-mercaptoethanol	5	1.14 ± 2.10
DTT	5	20.54 ± 6.73
Detergents	% (w/v)	
SDS	1	0.00 ± 0.00
Triton X-100	1	72.22 ± 2.47
Tween-20	1	24.02 ± 3.60
DOC	1	120.43 ± 0.01
Nonidet P-40	1	91.04 ± 1.96
NLS	1	0.00 ± 0.00
Sugars	mM	
NAG	10	104.10 ± 3.34
Chitobiose	10	81.97 ± 1.89
Organic solvents	% (v/v)	
Ethanol	10	90.91 ± 2.46
Methanol	10	88.61 ± 2.34
Isopropanol	10	69.44 ± 1.96
DMSO	10	97.73 ± 14.4
Salts	M	
NaCl	0.1	38.53 ± 1.83
NaCl	0.25	29.99 ± 8.05
NaCl	0.5	16.55 ± 0.09
NaCl	1	8.61 ± 0.37
NaCl	2	4.35 ± 2.52

Value are mean ± SD from three independent experiments

Supplementary materials

Appendix 1. Chi18H8 expression conditions

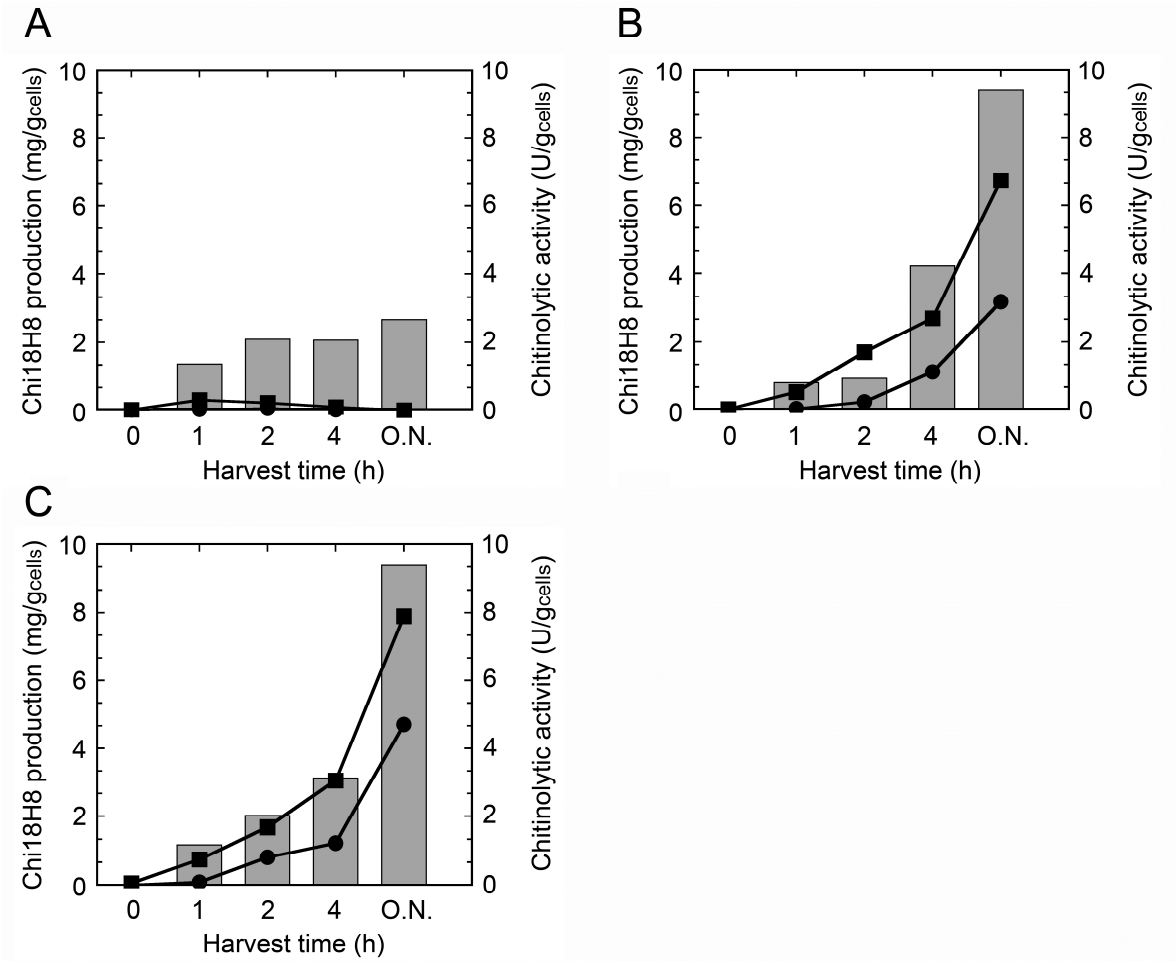


Figure S1. Chi18H8 expression levels in *E. coli* BL21 Star™(DE3)/pET24b(+)::*chi18H8* recombinant strain. The recombinant strain was incubated at 37 °C (A), 25 °C (B) and 20 °C (C) after induction with 0.4 mM IPTG and cells were harvested after 0, 1, 2, 4 hours and overnight (O.N.) from induction. Chi18H8 production (mg protein/g cells wet weight) into insoluble fractions (grey bars) was determined by densitometric analysis of proteins separated through SDS-PAGE. Chitobiosidase activity was measured by fluorimetric assay on 4-MU-$(GlcNAc)_2$, either in the soluble cytoplasmic fractions of cells (■) or in the insoluble fractions (membrane and IBs)(●).

Appendix 2. Chi18H8 solubilisation from IBs and purification

E. coli BL21 Star™(DE3)/pET24b(+)::*chi18H8* IBs were diluted in HCl or lactic acid, either 10 or 100 mM and incubated at 37 °C under vigorous shacking for several hours. The results of the experiments are summarised in Table S1. The highest solubilisation yields (>90%) were achieved when IBs were re-suspended in 10 mM HCl. However, in 10 mM lactic acid the protein showed a higher specific activity, up to 40.7 U/mg.

Table S1. Chi18H8 solubilisation from IBs in acid pH conditions. Solubilisation yield was estimated by densitometric analysis of proteins separated with acrylammide electrophoresis. The activity was measured on 4-MU-$(GlcNAc)_2$ as substrate, at 37 °C in 100 mM sodium acetate pH 5.0.

Solubilisation agent	Concentration (mM)	Solubilisation yield (%)	Specific activity (U/mg Chi18H8)
HCl	10	93	29.6
	100	20	15.91
Lactic acid	10	82	40.7
	100	60	30.71

Values are mean from at least three independent experiments

Samples of the Chi18H8 solubilised from IBs in 10 mM lactic acid and then dialysed against 100 mM HEPES pH 5.6, 10 mM calcium lactate pH 6.6, 50 mM sodium acetate pH 5.0 or 25 mM sodium acetate pH 4.0, were used in pilot experiments for the identification of a suitable purification method. Table S2 reports the different resins, the chromatographic conditions and the results of enzyme activities measured with the fluorimetric assay (Hjort *et al.*, 2014) in the flow through and in the eluate.

Only the weak cationic exchanger P11 and the hydrophobic resin HP20SS were able to retain the recombinant chitobiosidase; however, in both the cases the enzyme was impossible to be eluted in an active form. For the other resins, the chitinolytic activity was mainly found in the flow-through together with most of the contaminating proteins. Only with the weak anionic exchanger WA11 resin used in HIC conditions, most of the contaminating proteins were retained by the resin, allowing a negative purification of the chitinase from the flow through.

Resin type	Chi18H8 buffer	Activation buffer	Equilibration buffer	Chi18H8 activity in the flow through (%)	Elution buffer	Chi18H8 activity in the eluate (%)
HiTrap™ Chelating HP (GE Healthcare)	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6 / imidazole 20 mM / NaCl 500 mM	HEPES 100 mM pH 5.6 / imidazole 20 mM / NaCl 500 mM	100	HEPES 100 mM pH 5.6 / imidazole 250 mM / NaCl 500 mM	0
HiTrap™ Chelating HP (GE Healthcare)	Sodium acetate 50 mM pH 5.0	Sodium acetate 50 mM pH 5.0/ imidazole 20 mM / NaCl 100 mM	Sodium acetate 50 mM pH 5.0/ imidazole 20 mM / NaCl 100 mM	100	Sodium acetate 50 mM pH 5.0/ imidazole 500 mM / NaCl 100 mM	0
HiTrap™ Chelating HP (GE Healthcare)	Sodium acetate 50 mM pH 5.0	Sodium acetate 50 mM pH 5.0/ glycerol 5% (w/v)	Sodium acetate 50 mM pH 5.0/ glycerol 5% (w/v)	100	Sodium acetate 50 mM pH 5.0/ glycerol 5% (w/v) / imidazole 500 mM	0
HiTrap™ SP FF (GE Healthcare)	Lactic acid 10 mM	Lactic acid 50 mM	Lactic acid 50 mM / NaCl 5 mM	100	Lactic acid 50 mM / NaCl 0.5 M	0
HiTrap™ SP FF (GE Healthcare)	Acetic acid 25 mM pH 4.0	Acetic acid 25 mM pH 4.0	Acetic acid 25 mM pH 4.0	100	Acetic acid 25 mM pH 4.0 / NaCl 1 M	0
DEAE-Sephadex A-25 (Pharmacia)	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
DEAE-Sephadex A-50 (Pharmacia)	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
DEAE-Cellulose (Sigma-Aldrich)	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Diaion WA11 (Resindion)	HEPES 100 mM pH 5.6	NaOH 0.1 M	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Cellex-E (Bio-Rad)	HEPES 100 mM pH 5.6	NaOH 0.1 M	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Amberlite IRA-67 (BDH)	HEPES 100 mM pH 5.6	NaOH 0.1 M	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Amberlite IRA-900 900 (Alfa Aesar)	HEPES 100 mM pH 5.6	NaOH 0.1 M	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Amberlite IRA-904 900 (Alfa Aesar)	HEPES 100 mM pH 5.6	NaOH 0.1 M	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Dowex 1X8 (Fluka)	HEPES 100 mM pH 5.6	NaOH 0.1 M	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Chelex 100 (Bio-Rad)	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / NaCl 1 M	0
Dowex rg 50WX2 (Alfa Aesar)	Calcium lactate 10 mM pH 6.6	HCl 0.1 M	Calcium lactate 10 mM pH 6.6	87	Calcium lactate 10 mM pH 6.6 / NaCl 1 M	0
Chelex 100 (Bio-Rad)	Calcium lactate 10 mM pH 6.6	Lactic acid 100 mM	Calcium lactate 10 mM pH 6.6	85	Calcium lactate 10 mM pH 6.6 / NaCl 1 M	0

Resin type	Chi18H8 buffer	Activation buffer	Equilibration buffer	Chi18H8 activity in the flow through (%)	Elution buffer	Chi18H8 activity in the eluate (%)
CF11 (Whatman)	Calcium lactate 10 mM pH 6.6	Lactic acid 100 mM	Calcium lactate 10 mM pH 6.6	100	Calcium lactate 10 mM pH 6.6 / NaCl 1 M	0
Cellex CM (Bio-Rad)	Calcium lactate 10 mM pH 6.6	Lactic acid 100 mM	Calcium lactate 10 mM pH 6.6	70	Calcium lactate 10 mM pH 6.6 / NaCl 1 M	0
Diaion WA11 (Resindion)	Calcium lactate 10 mM pH 6.6	Lactic acid 100 mM	Calcium lactate 10 mM pH 6.6	61	Calcium lactate 10 mM pH 6.6 / NaCl 1 M	0
P11 (Whatman)	Calcium lactate 10 mM pH 6.6	Lactic acid 100 mM	Calcium lactate 10 mM pH 6.6	0	Calcium lactate 10 mM pH 6.6 / NaCl 1 M	0
Diaion WA11 (Resindion) §	HEPES 100 mM pH 5.6	Water / methanol 1:1	HEPES 100 mM pH 5.6	100	HEPES 100 mM pH 5.6 / ethanol 50 %	0
Diaion WA11 (Resindion)§	HEPES 100 mM pH 5.6	Water / methanol 1:1	HEPES 100 mM pH 5.6	100	Water / ethanol 50 %	0
HP20SS (Resindion)	HEPES 100 mM pH 5.6	Water / methanol 1:1	HEPES 100 mM pH 5.6	0	HEPES 100 mM pH 5.6 / ethanol 50 %	0
HP20SS (Resindion)	HEPES 100 mM pH 5.6	Water / methanol 1:1	HEPES 100 mM pH 5.6	0	Water / ethanol 50-80 %	0

§ The resin was used in HIC mode due to its styrenic matrix

Table S2. Affinity chromatography (AC), ion exchange chromatography (IEC) and hydrophobic interaction chromatography (HIC) pilot experiments. Enzyme activity was measured on 4-MU-$(GlcNAc)_2$ as substrate, at 37 °C in 100 mM sodium acetate pH 5.0.

Chapter 2

Genetic screening of a metagenomic library derived from chitin-amended agricultural soil produces a novel salt-tolerant chitinase

Mariana Silvia Cretoiu[1,2], Francesca Berini[3], Anna Maria Kielak[4], Flavia Marinelli[3], Jan Dirk van Elsas[1]

[1]Department of Microbial Ecology, CEES, University of Groningen, Groningen - The Netherlands

[2]Department of Marine Microbiology, Royal Netherlands Institute for Sea Research, Yerseke - The Netherlands

[3]Department of Biotechnology and Life Sciences, University of Insubria, and "The Protein Factory" Research Center, Politecnico di Milano, ICRM CNR Milano and University of Insubria, Varese - Italy

[4]Department of Microbial Ecology, Netherlands Institute of Ecology, Wageningen - The Netherlands

Corresponding author: J.D. van Elsas, j.d.van.elsas@rug.nl

Abstract

Metagenomes derived from soil constitute rich sources of genes that encode enzymes with novel characteristics. Biasing soils towards particular functions, e.g. chitin degradation, has been recommended as a strategy to improve metagenomic hit rates. Here, we report on the construction of a metagenomic library from a chitin-amended disease-suppressive agricultural soil and its screening for genes that encode novel chitin-active enzymes. The library, constructed in fosmids in an *Escherichia coli* host, comprised 145,000 25- to 40-kb insert clones, for a total of approximately 5.6 Gb of cloned soil DNA. Using genetic

screenings by repeated *chiA* gene specific PCR, we identified and characterised five putative bacterial chitinase genes. The analysis thus allowed access to the genomic context of these five genes. The five fosmids with a putative *chiA*-like gene contained 25-35 kb inserts, all producing *chiA* specific amplicons. Sequencing of the whole fosmid inserts resulted in the identification of four putative chitinase genes, next to one chitin deacetylase gene. *De novo* annotation followed by comparative genomics of the relevant fosmid regions revealed that the genetic environments of the putative *chiA* genes were all different. Remarkably, one novel *chiA*-like gene occurred in two different backgrounds, indicating genetic shufflings. Using fosmid 53D1, this gene was cloned in an expression vector in an *Escherichia coli* host, and brought to expression. On the basis of cultures of this producer organism, the protein was purified and characterised in terms of stability and activity. The 53D1 chitinase can be typified as a novel halotolerant chitobiosidase, revealing high activity as well as structural stability.

Keywords: chitinase, fosmid library, functional metagenomics, suppressive soil

Introduction

Chitin and its derivatives constitute promising natural biopolymers, which are important for application in biomedicine, agriculture and the pharmaceutical industry. Particular features of chitins are their general biodegradability as well as lack of toxicity. Enzymes active on chitin and chitin oligomers, here denoted as chitinases, are of great interest for use in large-scale modification or degradation of chitin moieties. Two main areas of application have been described, *i.e*. (1) the development of agents to antagonise chitin-containing phytopathogenic fungi or nematodes for application in agriculture, and (2) the use of chitinases as industrial biocatalysts for the production of chitin derivatives. Chitinases belong to the generic family of glycosyl hydrolases and have wide ranges of molecular weights (20-115 kDa), optimal temperatures (18-90°C), pH (2.0-10.5) and pI values (3.5-8.0) (http://www.brenda-enzymes.org/). Specifically, all chitinases fall within the glycoside hydrolase (GH) families 18 and 19. Remarkably, the family-18 and -19 enzymes have

different structures and modes of action. Related to the structure of the catalytic domain and the position of the hydrolytic site on the polysaccharide chain, the chitinases show either endo- or exo-activity (1, 2). Clearly, chitinases are prevalent in the microbiota of basically all ecosystems on Earth, and highest quantities of chitin are turned over by microorganisms (bacteria and fungi) in marine and terrestrial settings (3, 4)

The taxonomic diversity of soil microorganisms involved in the degradation of chitin has been sparsely investigated. It has been reported that diversity in the chitinolytic process was key to the functioning of the N cycle (5, 6, 7). Furthermore, chitin-degrading enzymes were shown to be involved in bacterial-fungal competition for plant root exudates in the rhizosphere (8). Moreover, in agricultural soils, the addition of chitin helps to enhance the suppressiveness against soil-borne pathogens by stimulating active chitinolytic microbial communities (9, 10, 11).

Recent developments in metagenomics-based analyses of the soil microbiota have enabled the access to novel genes and useful biomolecules (12, 13, 14, 15). Furthermore, the use of ecological enhancement (substrate enrichment) may increase the efficiency of mining for enzymes with improved features (16, 17). Hence, a metagenomics approach with biased habitats to find novel chitinases is warranted.

In the present study, we used a chitin-amended agricultural soil as a source of novel genes encoding chitin-degrading enzymes. We report the construction of a large-insert metagenomic library in fosmids in an *Escherichia coli* host. The library was subjected to PCR-based screenings for genes encoding relatives of the (exochitinase) *chiA* gene. Whole fosmid inserts were then sequenced using next-generation-sequencing (Illumina) technology. The functional metagenomics analysis of soil enriched with chitin, as applied in this study, constitutes a key approach for the discovery of novel enzymes involved in the chitin metabolic pathway.

Materials and Methods

Soil samples

Replicate soil samples were collected from the top 20 cm of an experimental agricultural field amended with chitin, located at the experimental farm “Vredepeel” in the southeast of

the Netherlands. The field has been used since 1990 by the Applied Plant Research Institute (PPO) to test and monitor diverse agricultural practices. The soil was amended with 1.8% of shrimp waste chitin (20 tons/ha, calculated over the top 20 cm). The soil was characterised as sandy, with pH 5.7±0.2 and 3.2% organic matter. Following collection, the soil samples were homogenised by sieving (2 mm pore size mesh sieve) and mixing, after which 10 g were subsampled for DNA extraction.

Soil DNA extraction

High molecular weight DNA for the construction of a metagenomic library was produced using a modification of the protocol previously described by van Elsas *et al*. (15). 10 g of soil were suspended in 10 ml extraction buffer (100 mM Tris-HCl, 100 mM ethylenediaminetetraacetic acid disodium salt (NaEDTA), 100 mM NaPO4, 1.5% NaCl, 1% CTAB, pH 8.0), vortex-mixed (5 s) and sonicated (water bath sonicator) for 15 min. After sonication, 100 µl of Proteinase K (10 mg/ml) were added, followed by 2 h incubation at 37°C with gentle shaking (200 rpm). The solution was then subjected to extraction using 1 vol. of phenol/chloroform/iso-amylalcohol (25:24:1) at 60°C for 30 min (in a water bath). The aqueous phase was obtained and DNA was precipitated with 2-propanol, followed by embedding in agarose plugs (1% low-melting-point agarose). The DNA was then subjected to pulsed-field gel electrophoresis (PFGE) on 1% agarose gel supplemented in the upper part with 2% polyvinyl pyrrolidone (PVP), in 0.5X Tris-Borate-EDTA (TBE). Conditions for electrophoresis were: 14°C (using a PFGE DRIII System; BioRad, Hercules, California, USA), gradient 6 V/cm, included angle 120°, initial switch time 0.5 s, final switch time 8.5 s, linear ramping factor 20 h. Following electrophoresis, 2-cm agarose fragments containing DNA in the size range 30-40 kb were cut out of gel, without staining or exposure to UV radiation. From the agarose cuts, DNA was recovered using β-agarase (New England Biolabs, Ipswich, UK) according to manufacturer's protocol.

Metagenomic library construction

Construction of a metagenomic library was performed using the CopyControl Fosmid Library Production Kit (Epicentre, Madison, WI, USA). The metagenomic DNA was 5'-phosphorylated, blunt-ended and subsequently ligated into the pCC1Fos vector. *Escherichia coli* EPI300-T1R (Epicentre, Madison, WI, USA) was then transformed with the ligation mix.

The transformed cells were plated on Luria-Bertani (LB) agar supplemented with 12.5 µg/ml chloramphenicol (positive selection of fosmid clones), which were incubated overnight at 37°C. Following incubation, multiple colonies were found on the LB plates, which were pooled per plate (up to 1,500 colonies per pool). The metagenomic library was then stored as pools of "transformed EPI300 cells - amplified library", with glycerol at -80°C, according to the cloning kit manufacturer recommendations.

Chitinase A gene (*chiA*) PCR-based screening strategy

Primers GA1 Fw and GA1 Rev were used for *chiA* gene based genetic screenings. These primers and the suitable PCR conditions were previously described and optimised (18). The PCR products, with predicted sizes of 450-600 bp, according to a previous study of the same chitin-amended soil (11), were detected by standard agarose gel electrophoresis. To check the identity of the respective *chiA*-like genes, the amplicons were extracted from gel using the Wizard SV Gel and PCR CleanUp System (Promega, Madison, WI, USA), after which they were directly sequenced (using the reverse primer - LGC, Berlin, Germany). As controls, DNA extracted from EPI300-T1R cells as well as pCC1Fos were also tested for the presence of *chiA* genes. In addition, the primer sequences were aligned (stand-alone BLAST v. 2.2.28+) against the genome sequence of EPI300-T1R and the pCC1 plasmid vector provided by the cloning kit manufacturer (Epicentre, Madison, WI, USA), which revealed that no amplicons of the indicated size range were predicted.

Screening for *chiA*-positive clones

The metagenomic library was screened for the presence of clones containing *chiA*-like genes using a "pool-subpool-single" PCR strategy as described in (19, 20). For this purpose, clones (pooled, subpooled or single) were cultured overnight in LB broth supplemented with 12.5 µg/ml chloramphenicol in 96-well plates. The contents of 2 plates were combined for single plasmid extractions using the QIAprep Spin Miniprep kit (Qiagen, Venlo, The Netherlands). The resulting DNA templates were then used for PCR based detections. As back-ups, the original clones were stored at 4°C. In case of positive reactions, subpools of rows of each plate were retested. Following single-plate positive results, we could thus identify positives down to the single-row or single-clone levels. At the "single-row" and "single-clone" levels, the fosmid copy numbers were induced to up to 50 copies by adding 0.4 µl of autoinduction

solution (500X) per 200 µl of LB broth (supplemented with 12.5 µg/ml chloramphenicol) according to the producer's specifications (Epicentre, Madison, WI, USA). All PCR products were checked for size and integrity by standard agarose gel electrophoresis. The presence of *chiA*-like sequences was confirmed by amplicon sequencing, as described above. Amplicon sequences were assigned to chitinase genes by TBlast-X and aligned with a suite of 22 characterised chitinases (retrieved from GenBank and CAZy), by using Clustal-W (BioLinux7; 21). Tree building was based on the Neighbour-Joining algorithm, using bootstrapping with 100 repetitions and the substitution model (MEGA 5.2).

DNA extraction from selected clones

Presumptive *chiA*-like gene containing clones were cultured in 2 ml LB broth supplemented with 12.5 µg/ml chloramphenicol, and fosmid copy numbers were raised by adding 4 µl of autoinduction solution (500X) before incubating overnight at 37°C. Following incubation, fosmid DNA was extracted using the Gene Jet Plasmid Midi Preparation Kit (ThermoScientific, St. Leon-Rot, Germany). DNA size and integrity were verified by PFGE with the electrophoresis conditions described above. DNA concentration was measured in a spectrophotometer (Nanodrop; ThermoFisher Scientific, St. Leon-Rot, Germany).

Sequencing of fosmid insert DNA

Full-length inserts (average sizes 30-40 kb) of selected fosmid clones were sequenced employing an Illumina HiSeq platform (BaseClear, Leiden, The Netherlands). Prior to sequencing, the concentration and quality of the DNA were assessed by microfluidics-based electrophoresis (Bioanalyzer, Agilent Technologies, Waldbronn, Germany). Paired-end libraries were prepared for each individual fosmid DNA using the Paired-End DNA Sample Preparation Kit and specific adaptors (Illumina, Eindhoven, The Netherlands). Sequencing was then optimised and carried out on the HiScanSQ Illumina system.

Assembly and sequence analysis

Raw data processing was supported by BaseClear (Leiden, The Netherlands). Briefly, it implied the generation of FAST-Q sequence reads, followed by quality controls and *de novo* assembly, yielding consolidated contigs. The FAST-Q sequence reads were generated using the Illumina Casava pipeline (version 1.8.2). Initial quality assessment was based on data

passing the Illumina Chastity filter. Subsequently, reads containing adapters and/or PhiX control signals were removed using an *in house* filtering protocol. The second quality assessment was based on the FAST-QC quality control tool (version 0.10.0). The quality of the sequences was also enhanced by trimming off low-quality bases using the "Trim sequences" option of the CLC Genomics Workbench (v. 5.5.1; CLC Bio, Aarhus, Denmark). The quality-filtered sequence reads were further filtered by removing all reads that could align to vector backbone sequences. All remaining sequences were then employed to generate contiguous sequences using the "*de novo assembly"* option of the CLC Genomic Workbench (v. 5.5). The contigs were considered robust if the average coverage was higher than 500x. Coverage was calculated from mapping of the reads against the contig sequences. The final contigs were then considered to be representative of the whole inserts.

***De novo* annotations of fosmids**

The final contigs, obtained after the *de novo* assemblies, were first checked for the presence of the *chiA*-like gene sequences as obtained during the screening. For each individual fosmid, the initially obtained sequence of the amplicon was aligned against the full-length contig using the BioEdit (v. 7.2.0) sequence alignment editor (22); each alignment was manually checked for errors and gaps. Following this initial analysis, open reading frames (ORFs) were assigned to the insert sequence and verified in three ways. A first assignment was obtained using GLIMMER (v.3.02) (23; http://cbcb.umd.edu/software/glimmer/) on a BioLinux v.7 platform (21). A second assignment of ORFs was based on the MetaGene software (24). The third assignment of ORFs was obtained via automatic annotation of coding sequences using the Rapid Annotation Subsystems Technology (RAST) provided by the National Microbial Pathogen Data Resource (NMPDR) (25). The predicted ORFs were then compared between the three predictive tools, after which a consensus or 'most likely' annotation was obtained. The annotation of each ORF from RAST (yielding protein-encoding genes, further referred to as coding sequences, CDSs) was manually curated and completed by a similarity search against the non-redundant protein (http://www.ncbi.nlm.nih.gov) databases using BLASTP (http://blast.ncbi.nlm.nih.gov/Blastp). BLASTP algorithm parameters (Table 1A) were optimised according to the BLAST Program Selection Guide (http://www.ncbi.nlm.nih.gov/guide/training-tutorials/ BLAST tutorials and guides/).

For each CDS, the closest protein homolog was assigned based on previously described combined criteria (26, 27), *i.e*. query coverage (%), maximum identity (%), alignment scores (maximum and total score) and e-value (Table 1B). RAST annotation also included a scan for tRNA genes and classification according to the "Cluster of Orthologous Groups" of Protein (COGs). ORFs shorter than 120 bp were discarded when the query coverage and maximum identity criteria were not in the established range. Spacers were subsequently searched against the non-redundant database (http://www.ncbi.nlm.nih.gov) using all BLAST options to ensure that no ORFs were missed. Then start and stop codons were identified for all annotated ORFs.

Prediction of putative chitinase genes

Prediction of putative chitinase/glycosyl hydrolase family 18 functions was performed through the InterProScan (EMBL) integrative tool for search of similarities within all available functional annotated protein databases (sequences of proteins, protein super-families and hidden Markov models) (28). Furthermore, the secondary and tertiary structure of protein was predicted on the I-TASSER server (29; http://zhanglab.ccmb.med.umich.edu/I-TASSER/) with default parameters.

Promoters for bacterial genes were predicted using B-Prom (SoftBerry, http://linux1.softberry.com/berry). Ribosomal binding sites (RBSs) were identified using the web version of RBS Calculator (https://salis.psu.edu/software/; 30) and manually checked according to accepted models (31, 32). The taxonomic affiliation of genes annotated as chitinases and glycosyl hydrolases family 18 was confirmed by comparison with the complete Carbohydrate Sequence Database (CAZy) available using Mothra.ornl (CAZYmes Analysis Toolkit, 33).

Phylogenetic analysis of chitinase-like genes

All predicted chitinases and glycosyl hydrolases (family 18) were aligned - using Clustal-W (BioLinux v. 7; 21) - with 65 sequences of characterised chitinases extracted from the CAZy database. Phylogenetic reconstruction was performed by Maximum Likelihood (ML) analysis using the amino acid type substitution model of Jones-Taylor-Thorton (JTT) with uniform rates, partial deletion and site coverage cut-off of 95%. The tree was bootstrapped using 100

replicates. A characterised cellulose of *Escherichia coli* P12b was employed as an outgroup sequence.

Prediction of the origin of fosmid inserts, comparative genomics and identification of potential horizontal gene transfer

The origin of the inserts was predicted based on the phylogenetic affiliation of more than 75% of identified genes. We interpreted these predictions with the cautionary note in mind that *de novo* annotation is relative at the level of bacterial genomes in soil (34) and that the size of the inserts in each fosmid represents less than 1% of an average annotated genome. Synteny of the recovered fosmid inserts with regions of existing genomes and inter-gene regions similarities were determined using the multiple genome alignment progressiveMauve software (35). Searches for G+C-rich islands were performed by CpGFinder (SoftBerry, http://linux1.softberry.com/berry). Nucleotide frequency analysis of the sequences was carried out for screening for potential horizontally-transferred regions (Scatter Plot Viewer, http://www.jcvi.org/cms/research/).

Sub-cloning of the *53D1* chitinase gene

Escherichia coli DH5α (Invitrogen-Life-Technology, Carlsbad, USA) was used for cloning of one selected *chiA*-like gene, *i.e*. *53D1*. The DNA encoding the 53D1 putative chitinase was obtained by amplifying the ORF with the fosmid DNA as template, and cloned into the expression vectors pET24b(+) (Novagen Inc., Madison, USA) carrying kanamycin resistance and adding a polyhistidine tag (His_6-Tag) at the C-terminus of the protein, and pColdI (TaKaRa Bio Inc., Otsu, Japan) carrying ampicillin resistance and adding a His_6-Tag at the N-terminus of the protein. For cloning into pColdI, the following primers were employed: 53D1_pColdI_FW (5'-AATTGAGCTCAGTCACGGTTCGGTCTCTCC-3') and 53D1_pColdI_RV (5'-CCAAAAGCTTTTACGGTCTCAGCCGGGATG-3') containing the underlined restriction sites for *Sac*I and *Hind*III, respectively. Primers 53D1_pET24b_FW (5'-ACCACATATGATGAGTCACGGTT CGGTCTCTCC-3') and 53D1_pET24b_RV (5'-AATACTCGAGCGGTCTCAGCCGGGATGAGA-3') containing the underlined restriction sites for *Nde*I and *Xho*I respectively, were used for cloning chitinase cDNA into pET24b(+). To study the expression of the chitinase gene under the control of its putative native promoter, an additional pair of primers was designed in order to amplify the 150-bp region upstream of the start codon and encompassing the

essential genetic elements for expression. The primers employed were: 53D1prom_pET24b_FW (5'-AATACATATGCGGTCGGATGACTGTGGCGCC-3') and 53D1prom_pET24b_RV (5'-AATACTCGAGCGGTCTCAGCCGGGATGAGA-3'), carrying respectively *Nde*I and *Xho*I restriction sites. The results of the cloning and transformation procedures were routinely controlled by subsequent DNA sequencing (BMR Genomics, Padua, Italy). *E. coli* BL21 Star™(DE3) (Invitrogen-Life-Technology, Carlsbad, USA) transformed with pColdI::*53D1* and pET24b(+)::*53D1*, and *E. coli* DH5α carrying pET24b(+)::*53D1prom* plasmid were maintained on LB agar supplemented with 50 μg/ml kanamycin or 100 μg/ml ampicillin.

Expression and purification of the 53D1 chitinase

Expression experiments with the three recombinant strains described above and with the respective controls (carrying the empty vectors) were conducted in LB medium and in Terrific Broth (TB, Sigma-Aldrich, St Louis, USA). For protein purification, early exponential phase cells of *E. coli* BL21 Star™ (DE3)/pET24b(+)::*53D1* in LB (OD_{600nm} ~ 0.6) were induced by 0.5mM of isopropyl β-D-thiogalactopyranoside (IPTG) and harvested after overnight incubation at 25°C and 200 rpm. After washing with sodium chloride-Tris-EDTA (STE buffer: 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 100 mM NaCl), cells were sonicated on ice for 5 cycles of 30 s each. His_6-53D1 protein was purified from the dissolved cytoplasmic fraction (after 500 mM NaCl was added) by loading onto a 5-ml Ni^{2+}-Hitrap chelating affinity column (GE Healthcare Sciences, Little Chalfont, UK), equilibrated with 50 mM Tris-HCl pH 8.0, 500 mM NaCl and 20 mM imidazole, according to the manufacturer's instructions. The protein was eluted using increasing concentrations of 50 mM Tris-HCl pH 8.0, 500 mM NaCl and 250 mM imidazole and loaded onto a size-exclusion PD10 Sephadex G25 column (GE Healthcare Sciences, Little Chalfont, UK) equilibrated with 50 mM Tris-HCl pH 8.0.

SDS-PAGE electrophoresis, Western blotting and zymogram analysis

Protein purification was followed by sodium dodecyl sulphate-polyacrylamide (SDS-PAGE) gel electrophoresis on 12% (w/v) polyacrylamide gels (36). For Western blot analysis, the protein was identified by Anti His-Tag Antibody HRP Conjugate (Novagen Inc., Madison, USA) using a chemioluminescence method (ECL Western Blotting Detection System, GE Healthcare Sciences, Little Chalfont, UK). Molecular weight markers were from GE-

Healthcare Sciences, Little Chalfont, UK. Zymogram analysis was employed to detect chitinolytic activity on polyacrylamide gels (10% w/v) containing 0.7 mg/ml carboxymethyl-chitin-remazol brilliant violet (CM-chitin-RBV) (Loewe Biochemica, Sauerlach, Germany), as described in (37).

Chitinase activity assay

Chitin-degrading activity was fluorimetrically assayed with the chitooligosaccharide analogues 4-methyl umbelliferyl *N*-acetyl-β-D-glucosaminide (4-MU-GlcNAc), 4-methylumbelliferyl *N,N'*-diacetyl-β-D-chitobioside (4-MU-$(GlcNAc)_2$) and 4-methylumbelliferyl *N,N',N''*-triacetyl-β-D-chitotrioside (4-MU-$(GlcNAc)_3$) as substrates, as reported by (37). One unit (U) of chitinase activity was defined as the amount of enzyme required for the release of 1 µmole of 4-MU per min at pH 5.0 and 37°C. Chitinolytic activity on colloidal chitin prepared from shrimp shells (Sigma-Aldrich, St Louis, USA) as described by Hsu and Lockwood (38), was determined according to the method of Anthon and Barret (39) adapted to enzymatic hydrolysis. Briefly, 250 µl of protein sample were added to an equal volume of 10 g/l colloidal chitin, and the mixture was incubated at 37°C for 1 h. The reaction was quenched by boiling for 5 min and then centrifuged (20000 x *g*, 25°C, 15 min); 200 µl of the supernatant was mixed with equal volumes of 0.5 M NaOH and of MBTH reagent. After 15-min incubation at 80°C, 400 µl of a solution containing 0.5% (w/v) $FeNH_4(SO_4)_2 \cdot 12\ H_2O$, 0.5 % (w/v) sulfamic acid and 0.25 M HCl was added and allowed to cool to room temperature. Subsequently, 1 ml H_2O was added and absorbance at 620 nm (A620) determined. The released reducing sugars were estimated by comparison of A620 to a standard curve prepared with *N*-acetyl-D-glucosamine concentrations from 0 to 600 µM. One U of chitinase activity was defined as the amount of enzyme that released 1 µmol/ml *N*-acetyl-D-glucosamine per h at 37°C.

Enzyme characterisation

The optimum pH for 53D1 chitinase activity was determined by testing activity at different pHs with the fluorescent assay on 4-MU-$(GlcNAc)_2$ and the following buffers (100 mM): glycine-HCl (pH 3.0), sodium acetate (pH 4.0 and 5.0), sodium phosphate (pH 6.0 and 7.0), Tris-HCl (pH 8.0) and sodium pyrophosphate (NaPPi, pH 9.0). The optimum temperature for 53D1 chitinase activity was determined by incubating the reaction mixture at various

temperatures (5 - 70°C). The effects of metal ions (20 mM), enzyme inhibitors (5% v/v), chelating agents (20 mM), detergents (1% w/v), organic solvents (10% v/v), sugars (10 mM *N*-acetyl-D-glucosamine and 10 mM chitobiose) and increasing concentrations of NaCl (up to 2 M) were investigated by adding each compound to the assay mixture. All reagents were purchased from Sigma-Aldrich (St Louis, USA).

Results

Construction of a metagenomic fosmid library from chitin-treated soil

Using 10 g of chitin-amended soil, we produced 2.5 µg of pure high molecular weight (HMW) soil DNA (0.25 per g soil), with an average fragment size of about 40 kb. The DNA was found to be pure enough to serve for direct cloning into the pCC fosmid system. Following the cloning and plating steps, a total of 145,000 *Escherichia coli* fosmid clones was generated. The clones were harvested into pools, each containing up to roughly 1,500 individual fosmid clones. The estimated total size of the library was 5.8 GB, which is comparable to any of the large soil metagenomic libraries that have been previously reported (14). Employing screens of subsets of clones, the percentage of insert-carrying clones was estimated to be about 100%.

Screening for *chiA*-related genes

All fosmid pools were successfully screened by *chiA* gene based PCR, in that clear negative or positive signals were obtained. The presumptive positive pools were retested and only pools that turned out to be positive again were further considered. In total, 18 of the pools yielded positive PCR results twice. The resulting 18 amplicons, all with sizes of 450-600 bp, were then cloned and subjected to sequence analysis, after which the sequences were compared to existing *chiA* sequences by BLAST-N based comparisons with database sequences. The analysis showed that 13 of the 18 sequences were very remote from the canonical *chiA* gene sequences, having <35% homology. The underlying fosmid pools were not further considered. Thus, we focused on the remaining five predicted/detectable *chiA* sequences, with homologies to any defined *chiA*-like database sequence of >35%. The five fosmid pools were subjected to several cycles of splitting up in subpools and PCR detection of the *chiA*

gene, until, finally, single fosmids that generated the same signal were obtained. The resulting *chiA*-like amplicons were then confirmed as being representative of 'presumptive' *chiA*-like genes that encode chitin-splitting enzymes (Figure 1). The (five) fosmids were subjected to further characterisation using whole fosmid sequencing and one was subjected to in-depth functional analyses, as detailed below.

De novo annotation and general characteristics of genetic fragments recovered from fosmids of the chitin-amended soil metagenomic library

None of the five fosmid clones - denoted 14A, 22G3, 28C5, 53D1 and 101F8- revealed expression of chitin-degrading activity when the fluorimetric activity assay was used directly on supernatants of grown cultures. Phylogenetic analyses of the generated *chiA* amplicon sequences showed, for all amplicons, strong support for affiliation to bacterial glycosyl hydrolases of family 18 (GH18; Figure 1). Furthermore, the sequences were sufficiently different from known chitinases to warrant further investigations. On the basis of the foregoing, the inserts of all five fosmids were subjected to full-length sequencing using Illumina paired-end technology. This yielded total sequence information amounting to 6.4 to 9.6 Mb per fosmid, at a coverage level over 100x (Table 2A). The sizes of the assembled inserts varied from 21.2 to 39.7 kb. Moreover, all inserts were confirmed to have a bacterial origin, as evidenced by the BLAST analyses. Furthermore, no tRNA or rRNA genes were identified. The G+C content of the fosmids was, on average, 58.8±6.4%. The G+C% was specific per fosmid and, for each fosmid, consistent over the full length of the insert (Table 2B). Overall, tetranucleotide counts varied from 188 (fosmid 22G3) to 355 (fosmid 101F8). Comparisons of the nucleotide frequencies across fosmid regions indicated potential regions of horizontal gene transfer. Moreover, similarities in sequences were found between fosmids 22G3, 28C5 and 53D1 (Figure 2). All three used annotation methods then showed for each of the fosmids, that ORFs for enzymes of the GH18 family, *N*-acetyl-glucosamine transport, sugar ABC transport, a molecular chaperone, one or more transcriptional regulator(s) and (overall) carbohydrate metabolism were present (Figure 3). Genes encoding putative chitinases and general carbohydrate transport/capture and metabolism proteins were singly present on fosmids 22G3, 28C5 and 53D1, twice (fosmid 14A) or up to fivefold (fosmid 101F8). Additionally, a range of other predicted gene functions were found (Supplementary Tables S1, S2, S3, S4 and S5), being a large fraction involved in housekeeping

and cell replication functions. The fraction of predicted gene products, denoted as hypothetical proteins, was relatively high, varying from 55 to 76%. The predicted proteins and their localisation on the different fosmids are presented in Table 3.

Fosmid annotation and prediction of closest homologue

Fosmid 14A

Eighteen ORFs were predicted to exist in the insert in the 22.6 kb-insert fosmid 14A (Table S1), which revealed three regions without identifiable ORFs. The overall G+C content was 52.7%. The majority of ORFs showed a positive transcription frame (Figure 3 - 14A). Gene lengths varied from extremely small, *i.e*. 188 bp (CDS8, hypothetical protein) to large, *i.e*. 2,393 bp (CDS7, closest hit β-D-galactosidase CAZy glycoside hydrolase family 2). One CDS (CDS10, 1,697 bp) was annotated as a putative gene encoding an endochitinase, with best BLAST hit (99% similarity, 99% coverage) to a recently described *Kitasatospora setae* chitinase (40). The number of genes with products potentially involved in carbohydrate metabolism and sugar ABC transporters was low (Table 3). Surprisingly, one third of the CDSs (so 6 CDSs) were assigned to proteins predicted to be involved in plasmid partitioning next to a phage-type integrase. Another 33% of the putative CDSs remained hypothetical. Half of the CDSs were affiliated to sequences from a *Burkholderia*-like source organism, at a level of similarity (protein based) between 27 (CDS6, ABC transporter) and 82% (CDS14, hypothetical protein) (Table S1).

Fosmid 22G3

Twenty putative ORFs were identified in the 21.2 kb-insert fosmid 22G3 (one gap). The overall G+C content was 58.8%. Only one ORF had a negative transcription frame (Figure 3 - 22G3; Table S2). The size of the CDSs was, on average, large, with 63% of CDSs having more than 0.5 kbp. The majority (55%) of the putative CDSs were affiliated to homologs found in an *Acidobacterium*-like organism, indicating that *Acidobacterium* was the presumed source organism. One gene of a typical chitinase was identified, which revealed 100% identity with, and coverage of, chitinase "A" of *Acidobacterium capsulatum* ATCC 51196. One CDS, for *N*-acetyl-glucosamine transport (similarity 51%, coverage 86%) was also affiliated to a homolog from *A. capsulatum* ATCC51196. One transcriptional regulator (similarity 33%, coverage 78%), which was similar to a region from *Granullicella tundricola* MP5ACTX9, was found

downstream of the putative chitinase gene. No chaperonins and sugar ABC transporter genes were found (Table 3).

Fosmid 28C5

A contiguous sequence, encompassing 35 ORFs with positive transcription frame, was assigned to the 31.9 kb-insert fosmid 28C5 (Figure 3 - 28C5; Table S3). The fosmid insert G+C content was 65.5%, namely the highest among all fosmid inserts. The sizes of the CDSs ranged from 143 (CDS9-hypothetical protein) to 2,309 bp (CDS34- transcriptional regulator). One gene (CDS24, 1,190 bp) was annotated as a gene encoding a putative chitinase, with a best BLAST hit to a gene from *Stenotrophomonas maltophilia* AU12-09 (45% similarity, 88% coverage). One *N*-acetyl-glucosamine ABC transporter with a best hit (25% similarity, 85% coverage) to a transporter from *Streptomyces bingchenggensis* BCW-1 was identified downstream of the chitinase gene. Other putative CDSs, such as transcriptional regulators and chaperonin GroEL, were affiliated at moderate similarity levels (average 38%) and high coverage (98%) to sequences from *Chloroflexi*-like organisms. One duplicate gene was assigned to an organophosphate pesticide hydrolase (similarity 44%, coverage 76%) of *Pseudomonas* sp. Ag1. Overall, the analysis suggested a broad range of putative source organisms such as, next to *Stenotrophomonas*, *Chloroflexi*, *Actinobacteria*, *Firmicutes* and *Cyanobacteria*.

Fosmid 53D1

A total of 34 ORFs, with positive and negative transcription frames, and one gap, were identified in the 35.4 kb-insert fosmid 53D1 sequence (Figure 3 - 53D1, Table S4). The overall G+C content was 54.6%. The minimum gene size was 143 bp (CDS6, transposase IS66) and the maximum size was 2,309 bp (CDS31, putative protein kinase-transcriptional regulator). Putative genes for hypothetical proteins represented 32% of the 53D1 sequence. One gene potentially encoding a chitinase (CDS20, 1,190 bp) was identified. The best BLAST hit of this latter gene was with a gene from an "uncultured bacterium" (48% similarity, 94% coverage), followed by one from *Stenotrophomonas maltophilia* AU12-09 (45% similarity, 87% coverage) and *Ktedonobacter racemifer* DSM 44963 (41% similarity, 93% coverage). Moreover, genes encoding predicted *N*-acetyl-glucosamine transporter and sugar ABC transporter proteins, affiliated with genes from *Ktedonobacter racemifer* DSM 44963, were

found downstream of the chitinase gene. Similarly, for the CDSs corresponding to transcriptional regulators and to hypothetical proteins flanking the chitinase gene, a *Chloroflexus*-type source organism was predicted (Table S4). Overall, 35% of the CDSs annotated as *Chloroflexus*-associated genes had, as close homologues, similar genes from the recently described *Nitrolancetus hollandicus* (41).

Fosmid 101F8

Fosmid 101F8 was found to contain a contiguous insert sequence of 37,907 bp. In total, 43 ORFs were identified, with positive and negative transcription frames (Figure 3 - 101F8; Table S5). The sizes of these ORFs varied from 230 (CDS4, hypothetical protein) to 1,580 bp (CDS33, putative sensory transduction protein). The overall G+C content was 59.6%. Two different putative chitinase genes and one putative chitin deacetylase gene were found. CDS3 (1,121 bp), with best BLAST hit (100% identity and coverage) with a gene region from the *Niastella koreensis* GR20-10 genome and CDS25 (404 bp), with best BLAST hit (77% similarity, 82% coverage) with a gene from *Streptomyces avermitilis* MA-4680, were annotated as putative chitinase genes. CDS5 (869 bp), with best BLAST hit (100% identity and coverage) to a gene from *Thermodesulfatator indicus* DSM15286, was assigned as belonging to a polysaccharide deacetylase protein family. The fosmid 101F8 sequence also contained the aforementioned genes for chitinase, transcriptional regulator, *N*-acetyl-glucosamine, sugar ABC transporters and carbohydrate metabolism. With the exception of the genes for the chitinases, putative deacetylase and one antiporter protein (CDS26), all CDSs were affiliated, with high similarity and coverage value, to genes from *Aeromonas*. Specifically, 72% of the CDSs were similar to genes from *Aeromonas veronii* (Table S5).

Genes and regions of similarity between fosmids

Although the genomic organisations were unique per fosmid insert, a number of common features was identified between the inserts. Fosmids 22G3 and 101F8 revealed the presence of a chitinase gene close to the 5'-end of the insert. All fosmids contained transcriptional regulators of the *LuxR, LitR* or *LysR* types, as well as sugar ABC transporter genes. Fosmids 22G3, 28C5 and 53D1 revealed one ORF encoding a putative *N*-acetyl-glucosamine transporter, downstream of the ORF for the predicted GH18 chitinase. Nucleotide frequency analyses indicated the presence of overlapping tetranucleotides between fosmids 28C5 and

53D1 (40.3% similarity), 28C5 and 22G3 (12.07% similarity) and 14A and 22G3 (7.26%). Progressive Mauve-based alignment of the ORF nucleic acid sequences showed the existence of 13 regions of significant similarity between fosmids 28C5 and 53D1 (Figure 4). Fosmids 14A and 101F8 revealed lower similarities when compared within the group of fosmids and they were considered to have unique sequences.

Putative genes for chitinases and selection of a candidate gene for expression analyses

Phylogenetic analyses of the putative chitinases (based on predicted protein sequences) showed some high-similarity clusterings with reference chitinases of the ChiA class and distant ones from the outgroup sequence of *E. coli P12b* cellulase (Figure 5). In the light of the annotation of the putative chitinase gene of fosmid 53D1 (homologous to a sequence from an uncultured bacterium, followed by best hits to regions of *Ktetodonobacter racemifer* DSM 44963 and *Stenotrophomonas maltophilia* AU12-09), we placed a focus on this sequence. The 53D1 chitinase ORF (G+C content 63.03%) was retrieved together with 200 nucleotides located upstream of the identified start codon, which was GTG (encoding Val). Then, essential genetic regions (promoter, RNA polymerase interaction site, Shine-Dalgarno sequence, start and stop codons), that are potentially important for expression in a heterologous recombination system, were identified (Figure 6). The sequence of the *53D1* ORF falls in the 14% of bacterial genes with an unusual promoter region and codon start (42, 43). The "-35…-10" region revealed the atypical sequence "ATGACT…CGGGAT", while the Shine-Dalgarno sequence was the universal AGGA. Overall, the rare reported transcriptional elements suggested that chitinase 53D1 had a weak promoter. The predicted protein was 396 amino acids long (having 44.7 kDa estimated molecular mass and an isoelectric point of 5.07). It belongs to the family-18 glycoside hydrolases, on the basis of the consensus sequence FDGIDIDWE, which confirmed the existence of a putative conserved active site within the catalytic domain (Figure 6). This sequence was further used for gene expression and protein characterisation studies.

Expression, purification and characterisation of the 53D1 chitinase

The amplified *53D1* gene region was cloned either into the pET24b(+) expression plasmid in *E. coli* BL21 Star^{TM}(DE3) cells under the control of the IPTG-inducible *T7* promoter, or into plasmid pColdI, a system based on a low-temperature-expression gene (cold shock gene),

which was specifically designed to improve the solubility of heterologous proteins in *E. coli* (44). Gene expression trials indicated that, in both expression systems, most of the recombinant protein (>80%) accumulated in insoluble cellular fractions, being some protein detectable in the soluble fraction (Supplementary Material and Figures S1, S2 and S3). Fluorimetric chitinase activity assays revealed that only the cytoplasmic soluble protein was active (maximal detectable activity of ca. 6 U per g of wet cells in optimised conditions, *i.e.* early exponentially growing *E. coli* BL21 Star^{TM}(DE3)/pET24b(+)::*53D1* in LB induced by 0.5 mM IPTG and harvested after overnight growth at 25°C and 200 rpm, Figure S2), whereas the accumulated insoluble form appeared inactive. Interestingly, when the *53D1* gene was cloned under the control of its native promoter, no protein was detectable (neither by chitinase assay nor by immunoblotting), confirming that the promoter activity was too weak to drive heterologous expression in *E. coli.* The 53D1 chitinase protein was purified from 3 g of cell paste obtained from 550 ml culture of *E. coli* BL21 Star^{TM}(DE3)/pET24b(+)::*53D1* grown as reported above. The protein was purified from the cytoplasmic soluble fraction by affinity chromatography on a HiTrap chelating column followed by gel filtration on PD10 Sephadex. The yield was 0.638 mg/l culture (0.117 mg/g cells). SDS-PAGE analysis showed that the protein migrated as a single band of 44.7 kDa and was >80% pure (Figure 7A). Zymogram on carboxy methyl chitin confirmed the chitinolytic activity of the protein (Figure 7B). Using three different-length analogues of natural chitooligosaccharides, 53D1 chitinase protein was found to have prevalent chitobiosidase activity (45.19 U/mg), weaker endochitinase activity (21.19 U/mg) and no β-*N*-acetyl-glucosaminidase activity. The enzyme was also capable of hydrolysing colloidal chitin, with an estimated activity of 2.28 U/mg.

Chitinase activity was then assayed in a pH range of 3.0-9.0 and a temperature range of 5.0-70.0°C using 4-MUF-$(GlcNAc)_2$ as substrate. The optimum pH for protein 53D1 activity was 5.0; more than 60 and 30% of the chitinolytic activities were maintained at pH 6.0 and 3.0-4.0, respectively (Figure 8A). At pH values exceeding 6.0, the activity drastically decreased. The optimum temperature for enzyme activity was between 35 and 40°C. However, more than 30% of the activity was retained even below 15°C, and more than 20 and 10 % at 50 and 70°C, respectively (Figure 8B). The effect of several compounds on 53D1 activity on 4-MU-$(GlcNAc)_2$ was then evaluated (Table 4). Among the metal ions tested, the presence of Mg^{2+} and Co^{2+} as well as the monovalent cation NH^{4+} did not significantly affect the hydrolytic activity of 53D1, while Cu^{2+}, Fe^{2+}, Mn^{2+} and Zn^{2+} effectively reduced it, with the

strongest inhibition being due to Cu^{2+} and Fe^{2+}. In contrast, Ca^{2+}, K^{+} and Ni^{2+} slightly increased the chitinolytic activity of the enzyme. Incubating the enzyme with the chelating agent EDTA inhibited its activity, suggesting that 53D1 is a metalloenzyme and that metal ions are needed for its catalytic activity. The enzyme inhibitors β-mercaptoethanol and DL-dithiothreitol (DTT) strongly reduced the activity of the 53D1 protein. The influence of a variety of detergents on 53D1 is shown in Table 4. Sodium dodecyl sulphate (SDS), sodium deoxycholate (DOC) and *N*-lauroylsarcosine (NLS) showed inhibitory effects, while other detergents (Triton X-100, Tween-20 and Nonidet P-40) had no effect or even slightly increased activity. The stability of the 53D1 protein activity was then evaluated using a panel of organic solvents. All solvents, *i.e*. ethanol, methanol, isopropanol and dimethyl sulfoxide (DMSO), significantly reduced the 53D1 protein activity, with an average residual activity in the range 45-65%. The activity of the 53D1 protein was slightly inhibited in the presence of 10 mM chitobiose and its activity increased slightly in the presence of 10 mM *N*-acetyl-D-glucosamine (NAG). Very interestingly, the 53D1 protein was resistant to, or even dependent on, high NaCl concentrations: its catalytic activity increased in the presence of NaCl, up to 2 M final concentration.

Discussion

Considering the prevalence of prokaryotic organisms in soil and their average estimated genome sizes (about 5 Mb; 45), the metagenomic library produced from the chitin-amended soil represented microbial community DNA equivalent to approximately 1,200 prokaryotic genomes. Genetic screening of this metagenome produced from the chitin-amended soil had, as the main objective, the identification of genes for novel proteins that belong to the functional group of chitin-active enzymes. The applied (genetic) screening strategy was based on the use of the highly diverse *chiA* gene (employment of conserved outer sequences as primer annealing sites) as the proxy for putative chitinases. One may argue that this screening strategy is contentious as it is limited to just one gene class, however it allowed us to screen the huge sequence space (10) around the *chiA* genes that are currently known. The strategy indeed proved to be successful in recovering several genomic fragments, in fosmids, containing putative active chitinases. In total, five sequences of chiA-like novel chitinases

were found in the same number of fosmids, next to a considerable number of (flanking) sequences related to the metabolic pathway of carbohydrate degradation, cellular transport and excretion systems, in addition to regulation of transcription. Given the estimated proportion of *chiA* genes in soil bacteria (roughly 1-5%), the frequency of recovery of *chiA* positive clones was consistent with the predicted one, corroborating that found in other reports (Wellington *et al*., in preparation).

The origins and source organisms of the genes for the putative chitin-active proteins were found to be diverse, indicating that genes for chitinases were spread across several dominating organisms. All five identified fosmid inserts had different predicted origins, but two of them, 28C5 and 53D1, were rather closely related. The majority of the putative genes on the fosmids revealed homologies to regions of the genomes of *Burkholderia*, *Actinobacteria*, *Stenotrophomonas*, *Acidobacterium* and *Aeromonas.* Particularly, fosmids 28C5 and 53D1 comprised genes which were to a rather great extent similar to those recently described in *Nitrolancetus hollandicus* (41) and in *Ktedonobacter racemifer*. Both organisms are members of the phylum *Chloroflexi*. The identification of a *Chloroflexus*-like chitinase gene in a metagenome from a chitin-enriched habitat may indicate another asset of the remarkable physiology of *N. hollandicus* like organisms. In fact, *N. hollandicus* has been described as a nitrite oxidiser, being the only one that is not affiliated with the *Proteobacteria*. The features of the 53D1 ChiA protein and the gene encoding it are remarkable. First, we did not detect 53D1 activity in *Escherichia coli* as driven from its own genetic background. This indicates that either the expression signals of the 53D1 *chiA* gene do not function properly in the *E. coli* host, or it is just too weakly expressed to enable detection. Moreover, the protein showed several features that characterise it as a typical temperate-climate-soil enzyme. First, it is apparently active under conditions of moderate temperatures as well as pH. Such conditions reign most of the time in many soils in temperate climate zones. Second, it revealed some sensitivity to organic solvents, indicating it has not been selected to withstand selective pressure from such sources. Third and remarkably, it showed quite high resistance to elevated levels of NaCl and even increased its activity at the highest salt levels, an unusual behaviour for a bacterial chitinase. In fact, few halo-tolerant bacterial chitinases, *i.e.* enzymes that can tolerate high saline concentrations despite having their maximum activity in the absence of salts, have already been characterised (examples can be found in (46, 47)). However, to our knowledge, among the

bacterial chitinases, only the two forms of chitinase C, Chi-I and Chi-II, from the halophilic bacterium *Salinivibrio costicola* (48) and ChiL from *Bacillus pumilus* SG2 (49) showed a similar behaviour to 53D1. The first two chitinases have, in fact, a salinity optimum at 1-2% NaCl, a residual activity of more than 80% and 50% in the presence of 3-5% and 14% NaCl respectively, and 95% activity without salt. Similarly, ChiL exhibits highest activity in the presence of 0.5 M NaCl. Halophilicity is more common among archeal chitinases: some of them are active even in the absence of salt but show maximum activities in high salinity conditions (50), others are not only adapted to tolerate high concentrations of salt, but also need a variable amount of NaCl for their correct folding (51, 52). This feature of 53D1 is truly interesting, as it (1) points to an *in-situ* activity whose level may depend on the presence of salt, and (2) may play a role in the soil in microhabitats where salt accumulates, *i.e*. in soils under drought stress.

Fourth and biotechnologically relevant, the 53D1 ChiA protein is active on colloidal chitin and not only on the chito-oligosaccharide analogues that are commonly used for chitinolytic enzyme detection. Consistently, the catalytic domain of the 53D1 ChiA protein was shown to contain a 62-residue chitin insertion domain (CID). This region, associated with the TIM-barrel structure typical of glycosyl hydrolase family 18, provides a deep substrate binding cleft, thus enhancing its exo-type activity on long-chain substrates (53). The 53D1 ChiA CID contains both of these sequence motifs (the N-terminal YxR and three separate [E/D]xx[V/I] motifs located in the central region) that are suggested to facilitate the access to recalcitrant substrates as chitin. These aspects, together with the remarkable salt tolerance of the 53D1 ChiA protein, are key properties of this enzyme, which make it an interesting candidate for the treatment of seafood waste such as shrimp carapace.

Finally, the fact that the 53D1 ChiA enzyme was most similar to a chitin-active protein from the *Chloroflexus* species *Ktedonobacter racemica* or *Nitrolancetus hollandicus,* next to its occurrence on another fosmid 28C5), indicated that as-yet-uncultured organisms that are affiliated with the mentioned (*Chloroflexus* types may play important roles in soils in which a substrate like chitin (that feeds them with respect to their carbon as well as nitrogen needs) is prevalent. Possibly, their value for the inferred source bacteria lies in their potential activity under drought stress, which comes with enhanced levels of dissolved salts in the soil solution. Furthermore, one may conclude that, in their ecological functioning, these organisms, being parts of complex communities “in action” on the offered substrate, may be

involved in horizontal gene transfers, given conditions of elevated cell-to-cell proximities and cellular activities. The remarkable differences in the genetic backgrounds of the very similar *chiA* homologs found on the two fosmids 53D1 and 28C5 appear to indicate the occurrence of horizontal transfers/recombinations involving the *chiA* gene in soil.

Acknowledgments: This study was funded by the EU-FP7 METAEXPLORE project awarded to J.D. van Elsas. We acknowledge the MIUR fellowship and contribution from Consorzio Interuniversitario per le Biotecnologie (CIB) to Francesca Berini. We also thank Elizabeth Wellington and Sara Sjöling for advice in data analysis and selection of candidate genes for further applications.

References

1. Henrissat B, Davies GJ. 2000. Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. *Plant Physiol.* **124**:1515-1519.
2. Van Scheltinga TAC, Kalk KH, Beintema JJ, Dijkstra BW. 1994. Crystal structures of hevamine, a plant defence protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure.* **2**:1181–1189.
3. Delpin M, Goodman AE. 2009. Nutrient regime regulates complex transcriptional start site usage within a Pseudoalteromonas chitinase gene cluster. *ISME J.* **3**:1053-1063.
4. Poulsen PH, Moller J, Magid J. 2008. Determination of a relationship between chitinase activity and microbial diversity in chitin amended compost. *Bioresour. Technol.* **99**:4355-4359.
5. Gooday GW. 1990. Physiology of microbial degradation of chitin and chitosan. *Biodegradation.* **1:**177–190.
6. Metcalfe AC, Krsek M, Gooday GW, Prosser JI, Wellington EMH. 2002. Molecular analysis of a bacterial chitinolytic community in an upland pasture. A*ppl. Environ. Microbiol.* **68**:5042-5050.
7. Manucharova NA, Vlasenko AN, Stepanov AL. 2007. Temperature as an autoecological factor of chitinolytic microbial complex formation in soils. *Biol. Bull.* **34**:163-169.
8. Bonfante P, Anca IA. 2009. Plants, mycorrhizal fungi, and bacteria: a network of interactions. *Annu. Rev. Microbiol.* **63**:363-383.
9. Korthals GW, de Boer M, Visser JHM, Molendijk LPG. 2010. Bodemgezondheid binnen bedrijfssytemen. *Mededelingeblad van de Koninklijke Nederlandse Plantenziektekundige Vereninging.* **41**:281-284.
10. Kielak AM, Cretoiu MS, Semenov AV, Sørensen SJ, van Elsas JD. 2013. Bacterial chitinolytic communities respond to chitin and pH alteration in soil. *Appl. Environ. Microbiol.* **79**:263-272.
11. Cretoiu MS, Kielak AM, Abu Al-Soud W, Sørensen SJ, van Elsas JD. 2012. Mining of unexplored habitats for novel chitinases--chiA as a helper gene proxy in metagenomics. *Appl. Microbiol. Biotechnol.* **94**:1347-1358.
12. Ferrer M, Martínez-Abarca F, Golyshin PN. 2005. Mining genomes and 'metagenomes' for novel catalysts. *Curr. Opin. Biotechnol.* **16**:588-593.
13. Simon C, Daniel R. 2009. Achievements and new knowledge unraveled by metagenomic approaches. *Appl. Microbiol. Biotechnol.* **85**:265-276.
14. Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R. 2011. Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. *FEMS. Microbiol. Ecol.* **78**:188-201.
15. van Elsas JD, Speksnijder AJ, van Overbeek LS. 2008. A procedure for the metagenomics exploration of disease-suppressive soils. *J. Microbiol. Methods.* **75**:515-522.
16. Ekkers DM, Cretoiu MS, Kielak AM, van Elsas JD. 2012. The great screen anomaly - a new frontier in product discovery through functional metagenomics. *Appl. Microbiol. Biotechnol.* **93**:1005-1020.
17. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH. 2012. Novel enzymes for the degradation of cellulose. *Biotechnology for Biofuels.* **5**:45.
18. Williamson N, Brian P, Wellington EMH. 2000. Molecular detection of bacterial and streptomycete chitinases in the environment. *Antonie Van Leeuwenhoek.* **78**:315-321.
19. Israel DI. 1993. A PCR-based method for high stringency screening of DNA libraries. *Nucleic Acid Res.* **21**:2627-2631.
20. Peterson DG, JP Tomkins, DA Frisch, RA Wing, AH Paterson. 2002. Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide (http:www.mgel.msstate.edu/newbac.htm).
21. Field D, Tiwari B, Booth T, Houten S, Swan D, Bertrand N, Thurston M. 2006. Open Software for biologists: from famine to feast. *Nature Biotechnol.* **24**:801-803.

22. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acid. Symp. Ser.* **41**:95-98.
23. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**:4636-4641.
24. Noguchi H, Park J, Takagi T. 2006. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Res.* **34**:5623-5630.
25. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Bioinformatics.* **9**:386.
26. Rost B. 1999. Twilight zone of protein sequence alignments. *Protein Eng.* **12**: 85-94.
27. Raghava GP, Barton GJ. 2006. Quantification of the variation in percentage identity for protein sequence alignments. *BMC Bioinformatics.* **7**:415.
28. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**(Web Server issue):W116-20.
29. Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**:725-738.
30. Salis HM, Mirsky EA, Voigt CA. 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**:946-50.
31. Shultzaberger RK, Bucheimer RE, Rudd KE, Schneider TD. 2001. Anatomy of *Escherichia coli* ribosome binding sites. *J. Mol. Biol.* **313**:215-228.
32. Stewart CR, Gaslightwala I, Hinata K, Krolikowski KA, Needleman DS, Peng AS, Peterman MA, Tobias A, Wei P. 1998. Genes and regulatory sites of the "host-takeover module" in the terminal redundancy of *Bacillus subtilis* bacteriophage SPO1. *Virology.* **246**:329-40.
33. Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC. 2010. CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiol.* **20**:1574-1584.
34. Raes J, Korbel JO, Lercher MJ, von Mering C, Bork P. 2007. Prediction of effective genome size in metagenomic samples. *Genome Biol.* **8**:R10.
35. Darling AE, Mau B, Perna NT. 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PlosOne* **5(6)**:e11147.
36. Schagger H, van Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**:368-379.
37. Hjort K, Presti I, Elvang A, Marinelli F, Sjöling S. 2014. Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics. *Appl. Microbiol. Biotechnol.* **98**:2819-2828.
38. Hsu SC, Lockwood JL. 1975. Powdered chitin agar as a selective medium for enumeration of actinomycetes in water and soil. *Appl. Microbiol.* **29**:422-426.
39. Anthon GE, Barrett DM. 2002. Determination of reducing sugars with 3-methyl-2-benzothiazolinonehydrazone. *Anal. Biochem* **305**:287-289.
40. Ichikawa N, Oguchi A, Ikeda H, Ishikawa J, Kitani S, Watanabe Y, Nakamura S, Katano Y, Kishi E, Sasagawa M, Ankai A, Fukui S, Hashimoto Y, Kamata S, Otoguro M, Tanikawa S, Nihira T, Horinouchi S, Ohnishi Y, Hayakawa M, Kuzuyama T, Arisawa A, Nomoto F, Miura H, Takahashi Y, Fujita N. 2010. Genome sequence *of Kitasatospora setae* NBRC 14216T: an evolutionary snapshot of the family *Streptomycetaceae*. *DNA Res.* **17**:393-406.
41. Sorokin DY, Lucker S, Vejmelkova D, Kostrikina NA, Kleerebezem R, Rijpstra WI, Damste JS, Le Paslier D, Muyzer G, Wagner M, van Loosdrecht MC, Daims H. 2012. Nitrification expanded: discovery, physiology and genomics of a nitrite-oxidizing bacterium from the phylum *Chloroflexi*. *ISME J.* **6**:2245-2256.

42. Tikole S, Sankararamakrishnan R. 2006. A survey of mRNA sequences with a non-AUG start codon in RefSeq database. *J. Biomol. Struct. Dyn.* **24:**33-42.
43. Nakamoto T. 2009. Evolution and the universality of the mechanism of initiation of protein synthesis. *Gene.* **432**: 1–6.
44. Qing G, Ma LC, Khorchid A, Swapna GVT, Mal TK, Takayama MM, Xia B, Phadtare S, Ke H, Acton T, Montelione GT, Ikura M, Inouye M. 2004. Cold-shock induced high-yield protein production in *Escherichia coli*. *Nat. Biotech.* **22**:877-882.
45. Hardeman F, Sjöling S. 2007. Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. *FEMS Microbiol. Ecol.* **59**:524-534.
46. Konagaya Y, Tsuchiya C, Sugita H 2006. Purification and characterization of chitinases from *Clostridium* sp. E-16 isolated from the intestinal tract of the South American sea lion (*Otaria flavescens*). *Lett Appl Microbiol.* **43**:187-193.
47. LeClair GR, Buchnan A, Maurer J, Moran MA, Hollibaugh JT 2007. Comparison of chitinolytic enzymes from an alkalin hypersaline lake and an estuary. *Environ Microbiol*. **9**:197-205.
48. Aunpad R, Rice DW, Sedelnikova S, Panbangred W 2007. Biochemical characterization of two forms of halo- and thermo-tolerant chitinase C of *Salinivibrio costicola* expressed in *Escherichia coli*. *Ann Microbiol*. **57**:249-257.
49. Vahed M, Motalebi E, Rigi G, Akbari Noghabi K, Soudi MR, Sadeghi M, Ahmadian G 2013. Improving the chitinolytic activity of *Bacillus pumilus* SG2 by random mutagenesis. *J Microbiol Biotechnol*. **23**:1519-1528.
50. Garcia- Fraga B, da Silva AF, Lopez-Seijas J, Sieiro C 2014. Functional expression and characterization of a chitinase from the marine archaeon *Halobacterium salinarum* CECT 395 in *Escherichia coli*. *Appl Microbial Biotechnol*. **98**:2133-2143.
51. Gomes J, Steiner W 2004. The biocatalytic potential of extremophiles and extremozymes. *Food Technol Biotechnol*. **42**:223-253.
52. Litchfield CD 2011. Potential for industrial products from the halopilic *Archea*. *J Ind Microbiol*. **38**:375-380.
53. Li H, Greene LH. 2010. Sequence and structural analysis of the chitinase insertion domain reveals two conserved motifs involved in chitin binding. *PLoS ONE* **5**:e8654.

Figure legends

Figure 1. Neighbour-joining tree of *chiA* amplicon sequences. Reference sequences represent nucleic sequences of characterised chitinases retrieved from GenBank and CAZy. Number near nodes indicate bootstrap values (only values ≥25).

Figure 2. Nucleotide frequency matrix along the whole insert fosmid sequences. Similarity (%) of overlap dotpoints is indicated in regions identified as significantly similar.

Figure 3. ORF orientation and position of selected genes.

Figure 4. Comparison of full length fosmid insert nucleotide sequences using progressive Mauve global and local alignment algorithms.

Figure 5. Maximum Likelihood phylogenetic analysis of chitinase protein sequences obtained in this study (marked) and 60 sequences of representative chitinases retrieved from CAZy. Substitution model Jone-Taylor-Thorton, uniform rates, partial deletion and site coverage cutoff 95%. Bootstrap values >25 are indicated.

Figure 6. Chitinase 53D1 ORF regions. Marked: RNA polymerase interaction region (grey shadow), -35 and -10 regions, mRNA start site, Shine-Dalgarno box, GH18 consensus sequence (indicator of active catalytic site), start and stop codons.

Figure 7. Purification of 53D1 from *E. coli* BL21 Star^{TM}(DE3)/pET24b(+)::*53D1* cells. (A) SDS-PAGE analysis of chromatography fractions. CE: crude extract; 1: flow-through; 2 and 3: fractions eluted at 125 and 500 mM imidazole, respectively. 53D1 protein spot is indicated by the arrow. (B) Zymogram analysis of purified 53D1 with CM chitin-RBV as substrate.

Figure 8. Enzymatic properties of the purified 53D1 chitinase, using 4-MU-$(GlcNAc)_2$ as substrate. Enzymatic activities are expressed as relative to the maximal recorded activity and the values represent the mean of three independent experiments (mean ± standard error) (A) pH profile of 53D1. (B) Temperature influence on chitobiosidase activity.

Tables

Table 1: Functional annotation criteria. (A) BLASTP settings according to protein size (amino acid residues) (B) BLASTP criteria used to validate the affiliation of ORFs to specific proteins

A.

Blast settings	*Query size (amino acids residues)*			
	>85 residues	*50-85 residues*	*35-50 residues*	*<35 residues*
Weight matrix	BLOSUM* 62	BLOSUM 80	PAM*70	PAM30
Gap cost	11,1	10,1	10,1	9,1
Low complexity filter	on	off	off	off
E value	10	10	1000	1000
Word size	3	3	3 or 2	2

*BLOSUM= Bloks Substitution Matrix. Calculates sequence comparison with less than 62% (default) or 80% identity.

*PAM=Point Accepted Mutation. Calculates sequence comparison based closely related proteins.

B.

Query coverage (%)	*Maximum identity (%)*	*Aligment scores (total & maximum)*	*E-value*
100%	20-100%	best hit	$<10^{-15}$
50%	48-100%	best hit	$<10^{-10}$
30%	75-100%	best hit	$<10^{-5}$

Table 2. General characteristics of genetic fragments recovered from chitin-amended soil fosmid library. (A) Sequence information (B) Fosmid length, gaps, GC%, ORFs.

A.

Fosmid	*Consensus length (bp)*	*Total read count*	*Single reads*	*Reads in pairs*	*Average coverage*
14A	22632	7547953	342043	7205910	5645,37
22G3	21255	8725007	436681	8288326	6428
28C5	31983	6426283	296701	6129582	47985,18
53D1	35473	9639420	29260	9610160	13743
101F8	37907	7863506	14887	7535920	1367,57

B.

Fosmid	*Length (bp)*	*Gaps*	*GC%*	*ORFs*			
				Total	*Non-hypothetical*	*Hypothetical*	*Unknown*
14A	22632	3	52,7	18	12	6	0
22G3	21255	1	58,8	19	11	7	1
28C5	31983	0	65,2	35	24	11	0
53D1	35473	1	54,6	34	23	11	1
101F8	37907	0	59,5	43	33	11	0

Table 3. Distribution of selected proteins and groups of cellular functions among fosmids.

Fosmid	Chitinase	Chitin deacetylase	Carbohydrate metabolism[a]	Hydrolase[b]	Transferase	Aldolase	Oxidoreductase	Dehydrogenase	Decarboxylase	ABC sugar-transporter[c]	Efflux transporters	Chemotaxis transducer	Transcriptional regulator	DNA replication	Periplasmic nucleotide binding protein	Membrane protein	Chaperonin	DNA repair	Non-ribosomal peptide synthetase	Anaerobic metabolism	Antibiotic resistance	Degradation of organic compounds	Plasmid-partitioning protein	Transposase/integrase	Hypothetical	Unknown
14A	1	0	2	0	0	0	0	0	0	1	2	0	0	2	0	0	0	0	0	0	0	0	3	1	6	0
22G3	1	0	0	0	0	0	0	0	0	1	2	0	2	0	0	2	1	1	2	0	0	0	0	0	7	1
28C5	1	0	0	1	2	1	1	0	0	3	2	1	4	1	1	2	1	0	0	0	0	2	0	1	11	0
53D1	1	0	0	1	0	0	1	1	0	3	2	0	4	1	3	0	1	0	1	0	1	0	0	2	11	1
101F8	2	1	2	1	4	0	0	0	1	0	6	0	4	3	1	2	0	1	2	1	1	0	0	0	11	0

a= other than chitinases & chitinase-like

b= other than carbohydrate hydrolases

c= N-acetyl-D-glucosamine ABC transport system & ABC-type sugar transport component

Table 4. Stability of 53D1 in the presence of different classes of compounds. The activity was measured on 4-MU-$(GlcNAc)_2$ as substrate, at 37°C in 100 mM sodium acetate pH 5.0. The values represent the mean of three independent experiments (mean ± std error).

Compounds	Final concentration	Relative activity (%)
Control		100
Metal ions	**mM**	
Ca^{2+}	20	126.02 ± 4.25
Cu^{2+}	20	7.62 ± 0.12
Fe^{2+}	20	15.99 ± 1.98
K^{+}	20	116.03 ± 0.59
Mg^{2+}	20	98.33 ± 1.26
Mn^{2+}	20	65.75 ± 2.52
Ni^{2+}	20	131.74 ± 1.06
NH_4^{+}	20	99.77 ± 2.91
Zn^{2+}	20	30.34 ± 2.01
Co^{2+}	20	90.43 ± 0.90
EDTA	20	65.20 ± 1.35
Enzyme inhibitors	**% (v/v)**	
β-mercaptoethanol	5	10.27 ± 6.32
DTT	5	11.80 ± 4.41
Detergents	**% (w/v)**	
SDS	1	0.65 ± 0.06
Triton X-100	1	118.59 ± 1.82
Tween-20	1	116.40 ± 2.55
DOC	1	40.61 ± 4.94
Nonidet P-40	1	115.42 ± 1.42
NLS	1	76.62 ± 1.83
Sugars	**mM**	
NAG	10	120.31 ± 3.34
Chitobiose	10	92.00 ± 1.38
Organic solvents	**% (v/v)**	
Ethanol	10	48.93 ± 2.58
Methanol	10	64.82 ± 3.87
Isopropanol	10	61.01 ± 0.48
DMSO	10	50.88 ± 4.39
Salts	**M**	
NaCl	0.1	108.11 ± 2.56
NaCl	0.25	133.42 ± 4.25
NaCl	0.5	134.89 ± 0.25
NaCl	1	149.89 ± 1.48
NaCl	2	169.25 ± 0.79

Figures

Figure 1

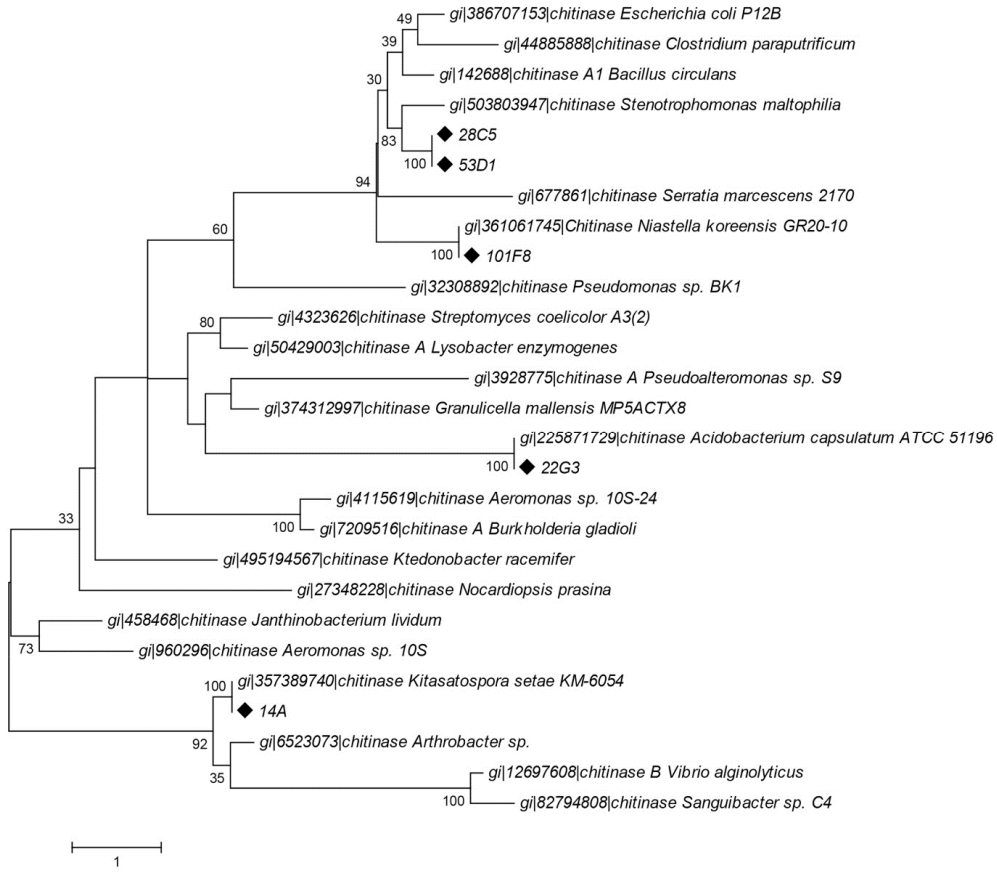


Figure 2

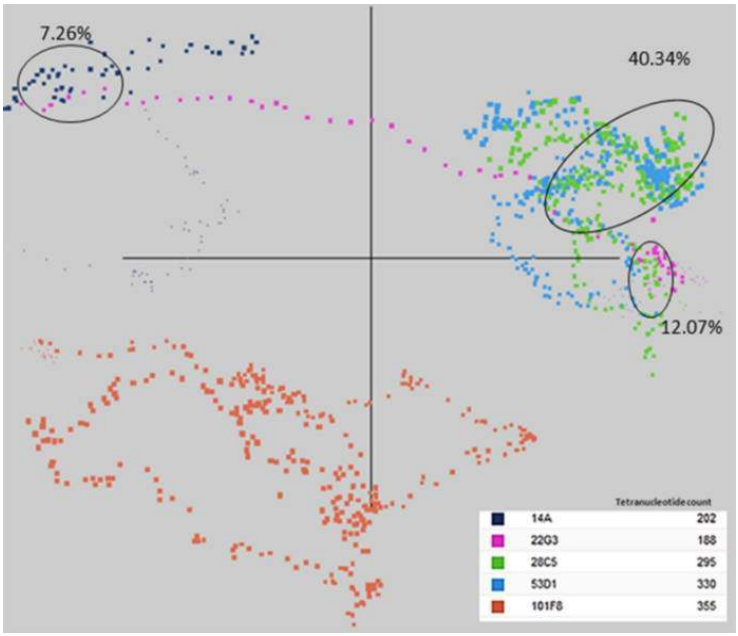


Figure 3

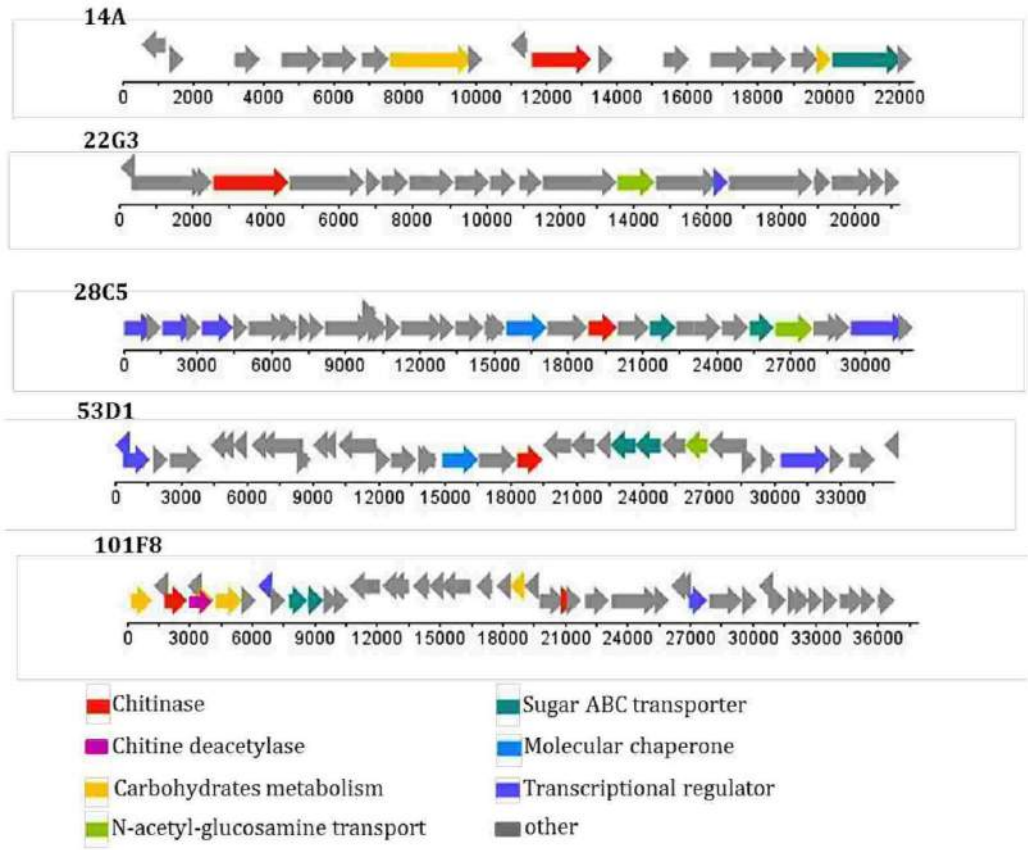


Figure 4

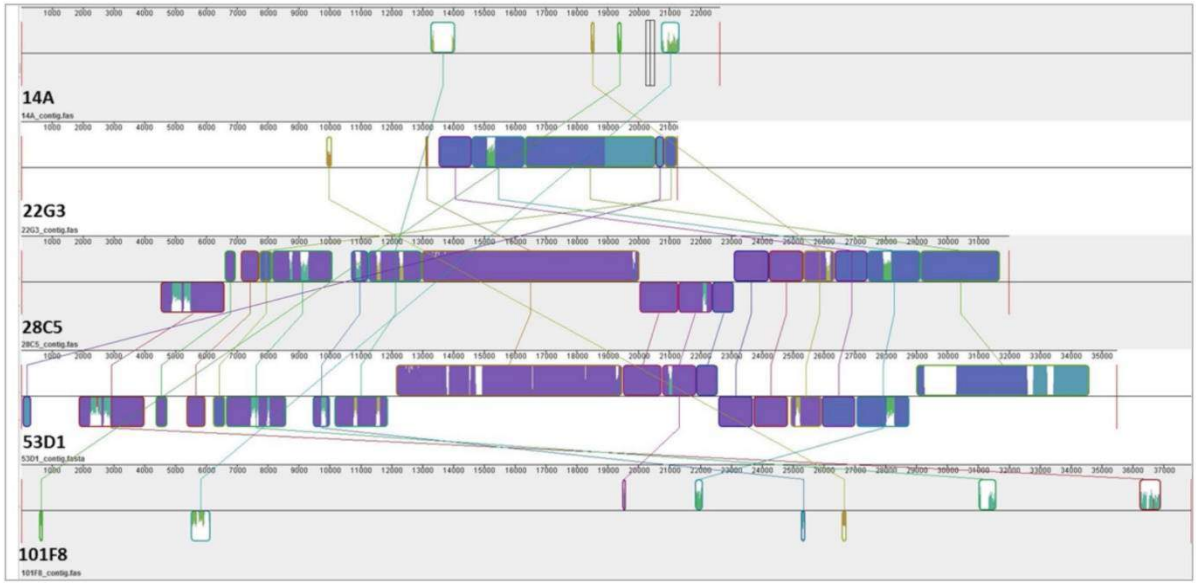


Figure 5

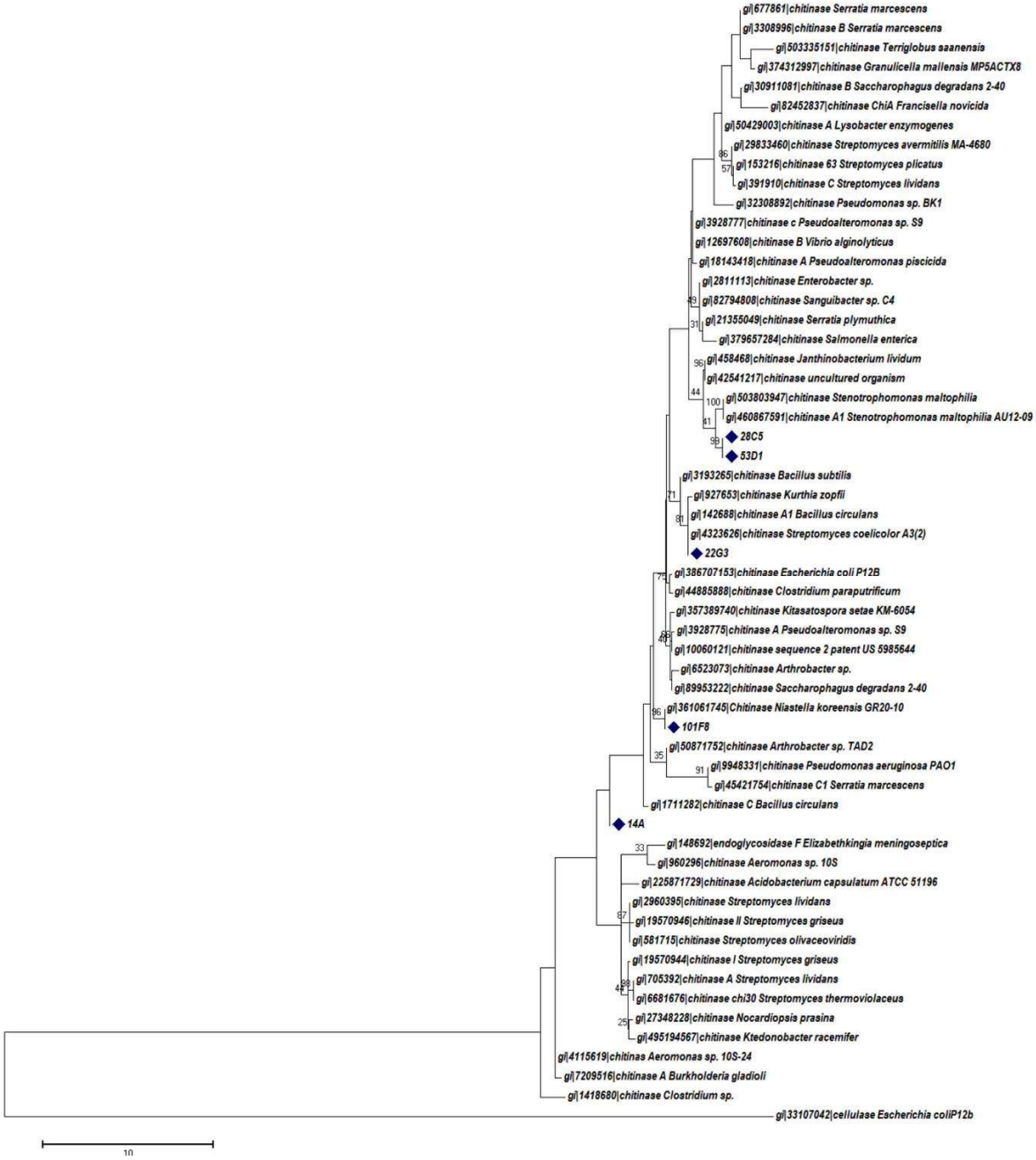

gi|677861|chitinase Serratia marcescens
gi|3308996|chitinase B Serratia marcescens
gi|5033351151|chitinase Terriglobus saanensis
gi|374312997|chitinase Granulicella mallensis MP5ACTX8
gi|30911081|chitinase B Saccharophagus degradans 2-40
gi|82452837|chitinase ChiA Francisella novicida
gi|50429003|chitinase A Lysobacter enzymogenes
gi|29833460|chitinase Streptomyces avermitilis MA-4680
gi|153216|chitinase 63 Streptomyces plicatus
gi|391910|chitinase C Streptomyces lividans
gi|32308892|chitinase Pseudomonas sp. BK1
gi|3928777|chitinase c Pseudoalteromonas sp. S9
gi|12697608|chitinase B Vibrio alginolyticus
gi|18143418|chitinase A Pseudoalteromonas piscicida
gi|2811113|chitinase Enterobacter sp.
gi|82794808|chitinase Sanguibacter sp. C4
gi|21355049|chitinase Serratia plymuthica
gi|379657284|chitinase Salmonella enterica
gi|458468|chitinase Janthinobacterium lividum
gi|42541217|chitinase uncultured organism
gi|503803947|chitinase Stenotrophomonas maltophilia
gi|460867591|chitinase A1 Stenotrophomonas maltophilia AU12-09
28C5
53D1
gi|3193265|chitinase Bacillus subtilis
gi|927653|chitinase Kurthia zopfii
gi|142688|chitinase A1 Bacillus circulans
gi|4323626|chitinase Streptomyces coelicolor A3(2)
22G3
gi|386707153|chitinase Escherichia coli P12B
gi|44885888|chitinase Clostridium paraputrificum
gi|357389740|chitinase Kitasatospora setae KM-6054
gi|3928775|chitinase A Pseudoalteromonas sp. S9
gi|10060121|chitinase sequence 2 patent US 5985644
gi|6523073|chitinase Arthrobacter sp.
gi|89953222|chitinase Saccharophagus degradans 2-40
gi|361061745|Chitinase Niastella koreensis GR20-10
101F8
gi|50871752|chitinase Arthrobacter sp. TAD2
gi|9948331|chitinase Pseudomonas aeruginosa PAO1
gi|45421754|chitinase C1 Serratia marcescens
gi|1711282|chitinase C Bacillus circulans
14A
gi|148692|endoglycosidase F Elizabethkingia meningoseptica
gi|960296|chitinase Aeromonas sp. 10S
gi|225871729|chitinase Acidobacterium capsulatum ATCC 51196
gi|2960395|chitinase Streptomyces lividans
gi|19570946|chitinase II Streptomyces griseus
gi|581715|chitinase Streptomyces olivaceoviridis
gi|19570944|chitinase I Streptomyces griseus
gi|705392|chitinase A Streptomyces lividans
gi|6681676|chitinase chi30 Streptomyces thermoviolaceus
gi|27348228|chitinase Nocardiopsis prasina
gi|495194567|chitinase Ktedonobacter racemifer
gi|4115619|chitinas Aeromonas sp. 10S-24
gi|7209516|chitinase A Burkholderia gladioli
gi|1418680|chitinase Clostridium sp.
gi|33107042|cellulase Escherichia coliP12b
10

Figure 6

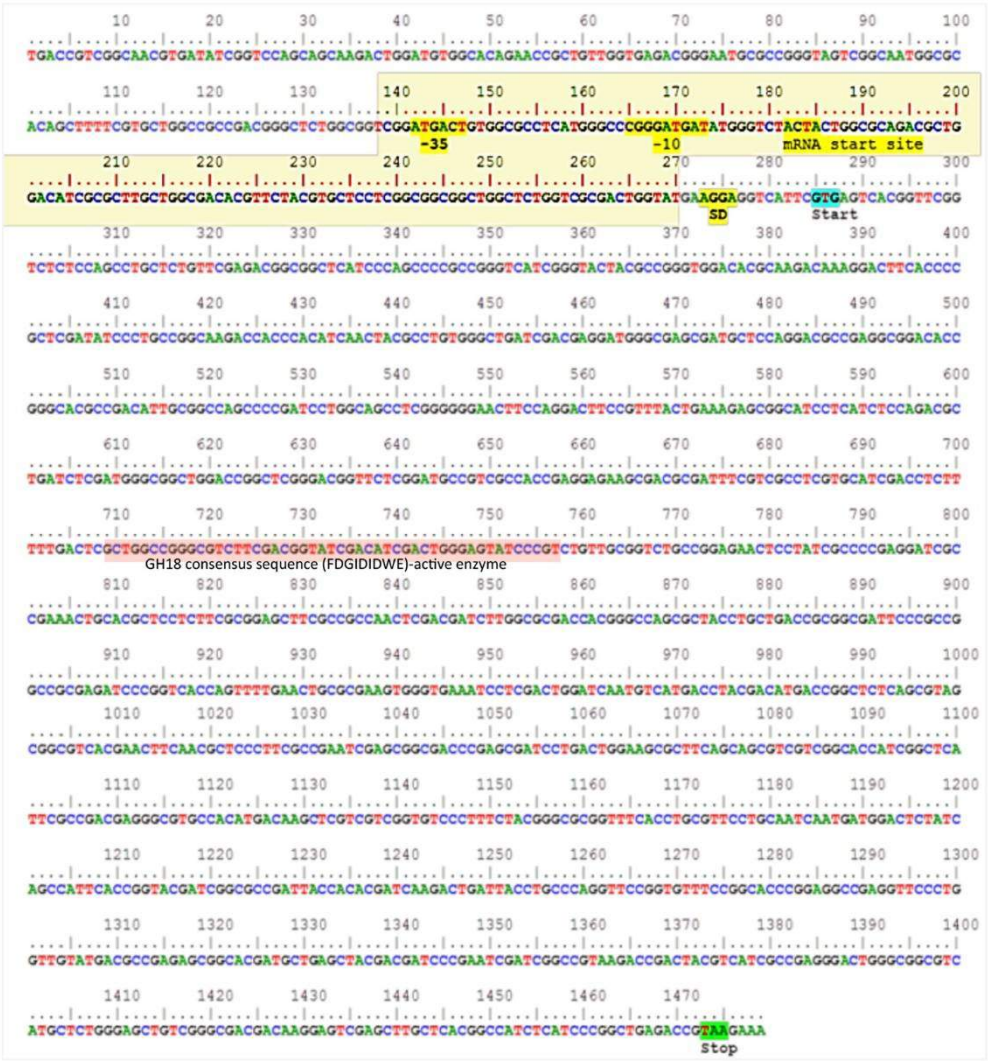

-35
-10
mRNA start site
SD
Start
GH18 consensus sequence (FDGIDIDWE)-active enzyme
Stop

Figure 7

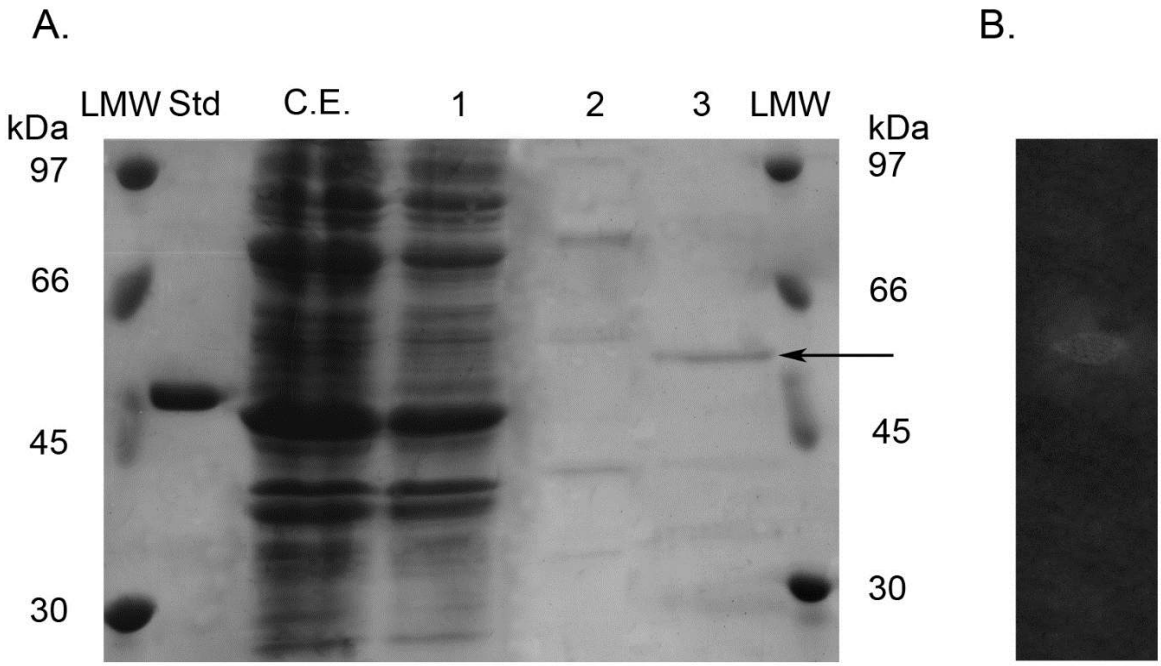
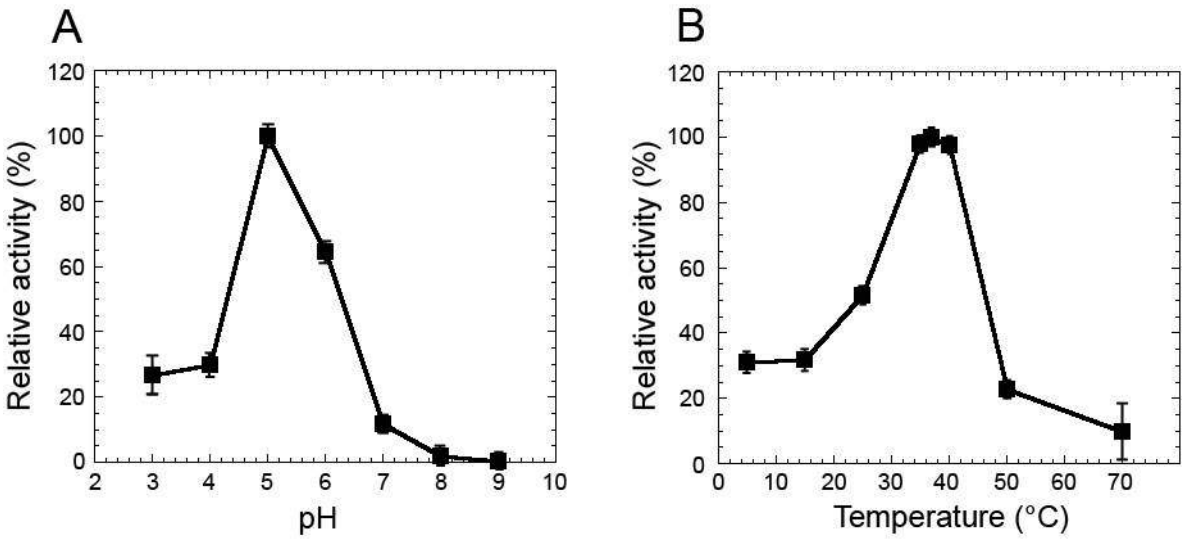


Figure 8



Supplementary material

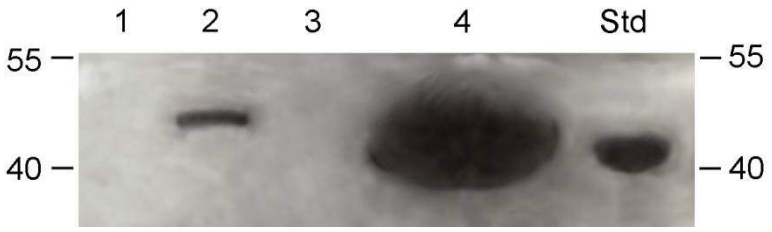
Expression of the *53D1* chitinase gene

Methods

Protein expression experiments were conducted in Luria Bertani (LB) and Terrific Broth (TB), both purchased from Sigma-Aldrich, St Louis, USA. Starter cultures were prepared from a single recombinant *E. coli* colony in 10 ml LB medium grown overnight at 37°C and 200 rpm. Baffled 500 ml Erlenmeyer flasks containing 100 ml of cultivation medium were inoculated with the starter culture (initial OD_{600nm} = 0.1) and further incubated as before. For the *E. coli* BL21 Star™(DE3) cells transformed with either pColdI::*53D1* or pET24b(+)::*53D1* vectors, protein expression was induced adding 0.5 mM IPTG to cells in the early-exponential growth phase (OD_{600nm} ~ 0.6 in LB, ~ 1.0 in TB). After induction, *E. coli* BL21 Star™(DE3)/pColdI::*53D1* cells were cultured at 15°C and 200 rpm according to manufacturer's instructions, while *E. coli* BL21 Star™(DE3)/pET24b(+)::*53D1* cells were incubated at 37°C or 25°C and 200 rpm. Cells were harvested at regular time intervals by centrifugation (1,900 x *g* for 30 minutes) at 4°C. Supernatants (*i.e.* the cell-free fermentation broths) were treated with 10% v/v trichloroacetic acid. Cell pellets were instead sonicated on ice (3-5 cycles of 30 s each, with a 30-s interval, using a Branson Sonifier 250, Danbury, USA) in Phosphate Buffer Saline (PBS) pH 7.3 containing 10 µg/ml deoxyribonuclease (DNase, Sigma-Aldrich, St Louis, USA), 0.19 mg/ml phenylmethylsulfonylfluoride (PMSF, Sigma-Aldrich, St Louis, USA) and 0.7 mg/ml pepstatin (Sigma-Aldrich, St Louis, USA). Soluble and insoluble fractions were then separated by centrifugation (20,000 x *g* for 40 minutes) at 4°C. Insoluble fractions (containing membrane and inclusion bodies) were re-suspended in a volume of PBS equal to the corresponding cytoplasmic soluble fraction (2-3 ml per gram of cells) for successive analyses. Protein concentration was determined by Biuret assay (Gornall *et al*., 1949). Chitinase production was estimated through densitometric analysis of SDS gel bands with the software Quantity One (Bio-Rad Laboratories, Hercules, USA) and His_6-glycin oxidase (His_6-GO) from *Bacillus subtilis* gently provided by Loredano Pollegioni, University of Insubria (Job *et al*., 2002) as standard. Chitinase activity was measured by the fluorimetric assay on 4-MU-$(GlcNAc)_2$ (Hjort *et al*., 2014).

Results

Basal expression was performed with *E. coli* BL21 Star™(DE3)/pET24b(+)::*53D1* in LB cultures, adding 0.5 mM IPTG when OD_{600nm} reached 0.6 and collecting cells after an additional 2 h of growth at 37°C. Using these conditions, most of the recombinant protein, corresponding to a band of 44.7 kDa, accumulated in insoluble fractions, being a low amount of protein detectable in the soluble fraction (Supplementary Fig. S1). Protein was not found in medium fractions, nor in fractions collected from control *E. coli* BL21 Star™(DE3) cells carrying the empty vector.

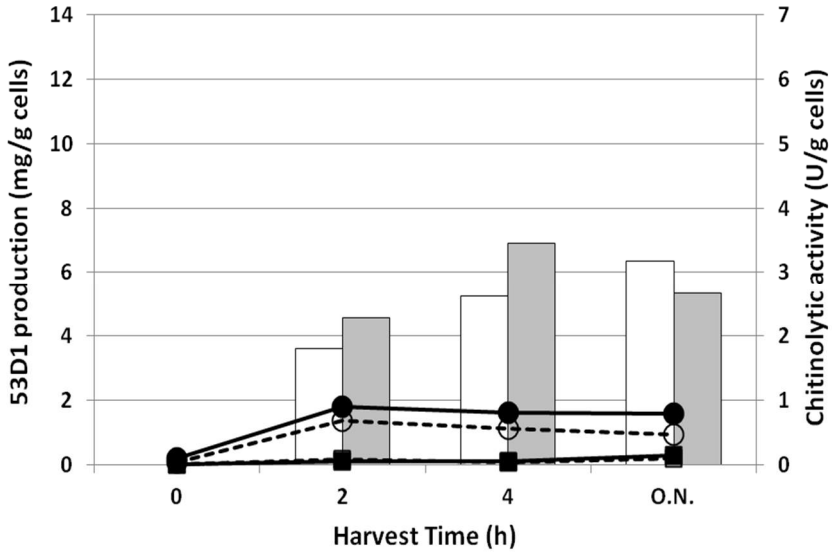


Supplementary Figure S1. Western blot analysis of *E. coli* BL21 Star™(DE3) cells carrying pET24b(+) or pET24b(+)::*53D1* plasmids, grown in LB medium for 2 h at 37°C after induction. From *E. coli* BL21 Star™(DE3)/pET24b(+): soluble (lane 1) and insoluble (lane 3) fractions; from *E. coli* BL21 Star™(DE3)/pET24b(+)::*53D1*: soluble (lane 2) and insoluble (lane 4) fractions. In each lane, samples corresponding to 2 ml of cell culture were loaded. Std reference protein: His_6-GO from *Bacillus subtilis* (10 μg, 42.66 kDa) gently provided by Loredano Pollegioni, University of Insubria (Job *et al*., 2002).

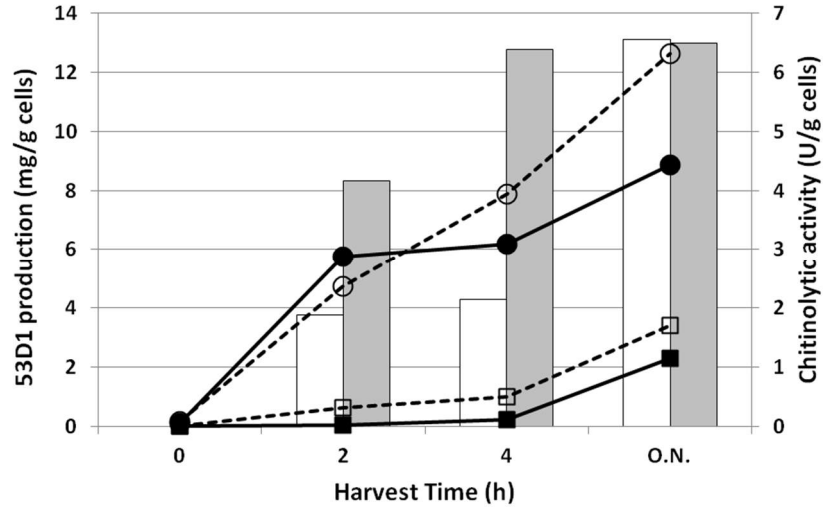
Chitinase activity assays revealed that the low amount of soluble protein in the above conditions was active (ca. 1 U per g of wet cells), whereas the insoluble form appeared inactive. Supplementary Figure S2A and B report the results from experiments conducted to increase the yield of active soluble 53D1 chitinase. Replacement of LB with the richer TB medium did not increase yield, whereas optimisation of the expression conditions (*i.e*. incubating cells overnight at 25°C after IPTG addition) yielded up to 6 U of activity per g cells, even if >80% remained insoluble and inactive (Supplementary Figure S2). When *E. coli* BL21 Star™(DE3)/pColdI::*53D1* was used with the aim of increasing soluble protein at low-temperature, a protein corresponding to the expected molecular mass was detectable in the insoluble fractions and chitinase activity (ca.4 U per g cells) was associated with the soluble

form of the recombinant protein as reported for the pET24b(+) plasmid (Supplementary Figure S3).

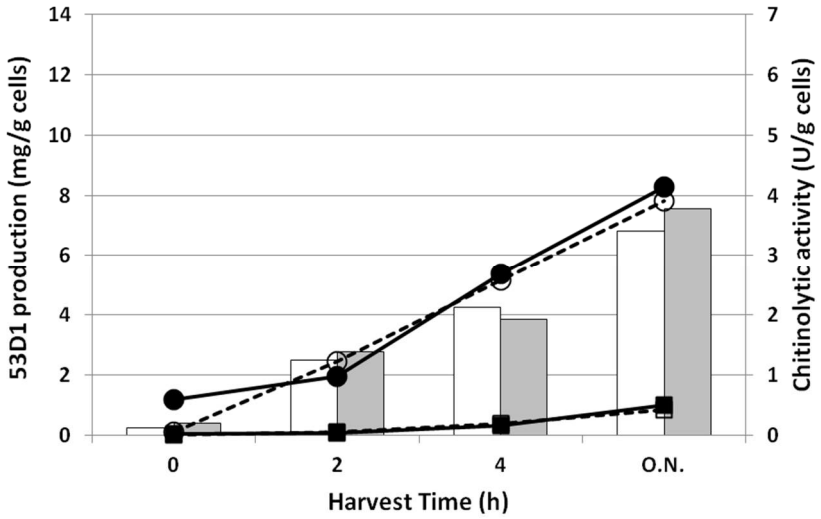
A



B



Supplementary Figure S2. 53D1 expression levels in *E. coli* BL21 Star™(DE3)/pET24b(+)::*53D1* recombinant strain. The recombinant strain was incubated at 37°C (A) or 25°C (B) after induction with 0.5 mM IPTG and cells were harvested after 0, 2, 4 hours and overnight (O.N.) from induction. 53D1 production into insoluble fractions (expressed as milligram of protein per gram of cells in wet weight) was determined by Western blot analysis using an anti-His-tag monoclonal antibody in cells grown in LB (empty bars) or TB (grey bars). Chitinase activity was measured by fluorimetric assay on 4-MU-$(GlcNAc)_2$, either in the soluble cytoplasmic fractions of cells grown in LB (○, dashed line) or TB (●, solid line) or in the insoluble fractions (membrane and inclusion bodies) of cells grown in LB (□, dashed line) or TB (■, solid line).



Supplementary Figure S3. 53D1 expression levels in *E. coli* BL21 Star™(DE3)/pCOLDI::*53D1* recombinant strain. The recombinant strain was incubated at 15°C after induction with 0.5 mM IPTG and cells were harvested after 0, 2, 4 hours and overnight (O.N.) from induction. 53D1 production into insoluble fractions (expressed as milligram of protein per gram of cells in wet weight) was determined by Western blot analysis using anti-His-tag monoclonal antibody in cells grown in LB (empty bars) or TB (grey bars). Chitinase activity was measured by fluorimetric assay on 4-MU-$(GlcNAc)_2$, either in the soluble cytoplasmic fractions of cells grown in LB (○, dashed line) or TB (●, solid line) or in the insoluble fractions (membrane and inclusion bodies) of cells grown in LB (□, dashed line) or TB (■, solid line).

References

Gornall AG, Bardawill CJ, David MM. 1949. Determination of serum proteins by means of the biuret reaction. *J Biol Chem*. **177(2):**751-766.

Job V, Marcone GL, Pilone MS, Pollegioni L. 2002. Glycine oxidase from Bacillus subtilis. Characterization of a new flavoprotein. *J Biol Chem*. **277(9):**6985-6993.

Hjort K, Presti I, Elväng A, Marinelli F, Sjöling S. 2014. Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics. *Appl Microbiol Biotechnol.* **98:**2819-2828.

Fosmid annotation

Table S1

Fosmid 14A

CDS	Locus	Gene lenght	Description	Best match accession no.	Source organism	Query cover (%)	E-value	Identity (%)
1	1203-544	659	hypothetical protein	ZP_02433069	*Clostridium scindens* ATCC 35704	43%	0.22	42%
2	1422-1769	347	hypothetical protein / phage integrase	ZP_06661375	*Escherichia coli* i B088	84%	2.00E-47	30%
3	3164-3919	755	replication initiation protein	NP_061424	*Escherichia coli* K-12	47%	7.00E-05	55%
4	4507-5673	1166	plasmid-partitioning protein ParA	NC_010804	*Burkholderia multivorans* ATCC 17616	76%	0.002	44%
5	5673-6644	971	plasmid-partitioning protein ParB	NP_061426	*Burkholderia pseudomallei* 1710b	50%	0.95	36%
6	6774-7559	785	efflux ABC transporter, permease protein	ZP_02505309	*Burkholderia pseudomallei* BCC215	94%	2.00E-45	27%
7	7559-9952	2393	beta-D-galactosidase	YP_002928937	*Escherichia coli* BW2952	39%	7.00E-12	63%
8	10028-10216	188	hypothetical protein UUU_32860	ZP_14290605	*Klebsiella pneumoniae* DSM 30104	93%	0.18	50%
9	11480-11019	461	hypothetical protein BACUNI_00159	ZP_02068759	*Bacteroides uniformis* ATCC 8492	36%	1.3	38%
10	**11606-13303**	**1697**	**putative endochitinase**	**YP_004905998**	***Kitasatospora setae* KM-6054**	**99%**	**0.0**	**99%**
11	13580-13927	347	phage integrase family protein	CP001408	*Burkholderia pseudomallei* MSHR346	91%	8.00E-26	46%
12	15322-16077	755	replication initiation protein	AM747720	*Burkholderia cenocepacia* J2315	98%	5.00E-11	44%
13	16665-17831	1166	plasmid-partitioning protein ParA	NC_016590	*Burkholderia* sp. YI23	99%	5.00E-41	39%
14	17831-18802	971	hypotethical protein BPSL1549	NP_061426	*Burkholderia pseudomallei* K96243	17%	9.8	82%
15	18932-19717	785	efflux ABC transporter, permease protein	NC_007952	*Burkholderia xenovorans* LB400	99%	1.00E-63	36%
16	19717-20103	386	beta-mannosidasee	NC_010943	*Stenotrophomonas maltophilia* K279a	94%	2.00E-81	68%
17	20111-22108	1997	putative sugar ABC transporter ATP-binding	AM747720	*Burkholderia cenocepacia* J2315	97%	9.00E-17	54%
18	22184-22372	188	hypothetical protein / phage integrase	ZP_06661375	*Escherichia coli* B088	39%	7.00E-12	63%

Table S2

Fosmid 22G3

CDS	Locus	Gene lenght	Description	Best match accession no.	Source organism	Query cover (%)	E-value	Identity (%)
1	330-43	287	hypothetical protein	NC_008536	Candidatus *Solibacter usitatus*	86%	6.00E-05	29%
2	334-2412	2078	ABC transporter	YP_002977039	Candidatus *Koribacter versatilis* Ellin345	43%	4.00E-80	50%
3	2413-2550	137	hypothetical protein, partial	WP_000469023	*Shigella flexneri*	100%	5.00E-22	100%
4	**2551-4638**	**2087**	**chitinase**	**YP_002753183**	***Acidobacterium capsulatum* ATCC 51196**	**100%**	**0**	**100%**
6	6727-7137	410	hypothetical protein Z1010	NP_286554	*Shigella flexneri*	100%	7.00E-95	100%
7	7138-7899	761	hypothetical protein Z1009	NP_286553	*Escherichia coli* O157:H7 str. EDL933	100%	0	100%
8	7900-9141	1241	Cardiolipin synthetase	NP_286552	*Escherichia coli*	100%	0.0	100%
9	9142-10098	956	conserved hypothetical protein	NC_012483	*Acidobacterium capsulatum* ATCC 51196	98%	6.00E-105	38%
10	10099-10812	713	ABC transporter inner membrane protein	YP_593488	Candidatus *Koribacter versatilis* Ellin345	79%	5.00E-19	42%
11	10873-11517	644	membrane protein	YP_002754337	*Acidobacterium capsulatum* ATCC 51196	98%	6.00E-105	48%
12	11518-13539	2021	excinuclease ABC subunit C	YP_005058066	*Granulicella mallensis* MP5ACTX8	59%	1.3	28%
13	13540-14595	1055	N-Acetyl-D-glucosamine ABC transport system	NC_012483	*Acidobacterium capsulatum* ATCC 51196	86%	4.00E-101	51%
14	14596-16314	1718	putative bacterial extracellular solute-binding protein	NZ_ACCL02000054	*Bryantella formatexigens* DSM 14469	89%	2.00E-85	41%
15	16315-16581	266	universal stress protein family protein	YP_002753605	*Acidobacterium capsulatum* ATCC 51196	97%	5.00E-28	32%
16	16582-18891	2309	LuxR family transcriptional regulator	YP_004218938	*Granulicella tundricola* MP5ACTX9	78%	5.00E-38	33%
17	18892-19359	467	unknown protein	AEU38485	*Granulicella mallensis* MP5ACTX8	26%	1.00E-08	37%
18	19360-20541	1181	hypothetical protein Sthe_2283	YP_003320526	*Sphaerobacter thermophilus* DSM 20745	79%	5.00E-19	42%
19	20542-20829	287	hypothetical protein OSG_eHP23_00230	AFH22375	Environmental *Halophage* eHP-23	17%	9.2	82%
20	20830-21255	425	non-ribosomal peptide synthetase	YP_007362971	*Myxococcus stipitatus* DSM 14675	36%	0.060	50%

Table S3

Fosmid 28C5

CDS	Locus	Gene lenght	Description	Best match accession no.	Source organism	Query cover (%)	E-value	Identity (%)
1	1-1308	1307	transcriptional regulator, LuxR family	ZP_06974075	*Ktedonobacter racemifer* DSM 44963	98%	3.00E-31	41%
2	1309-1596	287	hypothetical protein OSG_eHP23_00230	AFH22375	Environmental *Halophage* eHP-23	17%	9.8	82%
3	1597-2904	1307	transcriptional regulator, LuxR family	ZP_10243899	*Nitrolancetus hollandicus* Lb	99%	9.00E-54	36%
4	2905-3192	287	hypothetical protein OSG_eHP23_00230	AFH22375	Environmental *Halophage* eHP-23	17%	9.8	82%
5	3193-4500	1307	transcriptional regulator, LuxR family	ZP_10243899	*Nitrolancetus hollandicus* Lb	99%	1.00E-63	36%
6	4501-5079	578	conserved hypothetical protein	ZP_10243899	*Nitrolancetus hollandicus* Lb	94%	2.00E-81	68%
7	5080-6594	1514	transposase IS4 family protein	ZP_06965924	*Ktedonobacter racemifer* DSM 44963	97%	9.00E-176	54%
8	6595-6966	371	hypothetical protein NITHO_490006	ZP_10246048	*Nitrolancetus hollandicus* Lb	39%	7.00E-12	63%
9	6967-7110	143	hypothetical protein NITHO_490005	ZP_10246047	*Nitrolancetus hollandicus* Lb	93%	0.18	50%
10	7111-7716	605	transaldolase	ZP_09746516	*Saccharomonospora cyanea* NA-134	36%	1.3	38%
11	7717-8142	425	methyl-accepting chemotaxis sensory transducer	YP_001584157	*Burkholderia multivorans* ATCC 17616	43%	0.22	42%
12	8143-10098	1955	hypothetical protein Psta_2746	YP_003371273	*Pirellula staleyi* DSM 6068	84%	2.00E-47	30%
14	10273-10530	257	hypothetical protein MODMU_3678	YP_006367565	*Modestobacter marinus*	47%	7.00E-05	55%
15	10531-10710	179	organophosphate pesticide hydrolase	ZP_10474859	*Pseudomonas* sp. Ag1	76%	0.002	44%
16	10717-11241	524	hypothetical protein Xen7305	ZP_21054152	*Xenococcus* sp. PCC 7305	50%	0.95	36%
17	11242-12987	1745	hypothetical protein	WP_001567847	*Escherichia coli*	94%	2.00E-45	27%
18	12988-13428	440	dehydrogenase/reductase family oxidoreductase	ZP_14469092	*Pseudomonas stutzeri* TS44	91%	8.00E-26	46%
19	13429-14631	1202	dihydroorotase	YP_003861823	*Maribacter* sp. HTCC2170	98%	5.00E-113	44%
20	14632-15312	680	Zn-dependent hydrolase, glyoxylase	ZP_18907840	*Leptolyngbya* sp. PCC 7375	99%	5.00E-41	39%

CDS	Locus	Gene lenght	Description	Best match accession no.	Source organism	Query cover (%)	E-value	Identity (%)
21	15313-15492	179	organophosphate pesticide hydrolase	ZP_10474859	*Pseudomonas* sp. Ag1	76%	0.002	44%
22	15493-17127	1634	molecular chaperone GroEL	YP_001124352	*Geobacillus thermodenitrificans* NG80-2	95%	2.00E-80	36%
23	17128-18831	1703	two component regulator propeller domain-containing protein	ZP_24013577	*Clostridium termitidis* CT1112	99%	1.00E-15	24%
24	**18832-20022**	**1190**	**chitinase**	**ZP_23766528**	***Stenotrophomonas maltophilia* AU12-09**	**88%**	**2.00E-89**	**45%**
25	20023-21288	1265	hypothetical protein Tter_1563	YP_003323291	*Thermobaculum terrenum* ATCC BAA-798	92%	5.00E-45	36%
26	21289-22386	1097	ABC-type sugar transport system, ATP-binding component	ZP_10244076	*Nitrolancetus hollandicus* Lb	99%	1.00E-93	44%
27	22387-23088	701	putative acetyltransferase	YP_004591376	*Enterobacter aerogenes* KCTC 2190	39%	3.60E+00	31%
28	23089-24216	1127	ABC transporter	YP_003322712	*Thermobaculum terrenum* ATCC BAA-798	98%	3.00E-110	48%
29	24217-25341	1124	ABC transporter B family member 1	NP_181228	*Arabidopsis thaliana*	58%	7.00E-23	31%
30	25342-26364	1022	ABC-type sugar transport system, permease component	ZP_14747165	*Rhizobium* sp. CF080	97%	2.00E-49	24%
31	26365-27420	1055	N-Acetyl-D-glucosamine ABC transport system	YP_004961886	*Streptomyces bingchenggensis* BCW-1	85%	8.00E-55	25%
32	27478-29139	1661	putative bacterial extracellular solute-binding protein	ZP_05344232	*Bryantella formatexigens* DSM 14469	64%	8.00E-06	26%
33	29140-29406	266	conserved hypothetical protein	ZP_10244518	*Nitrolancetus hollandicus* Lb	78%	0.019	35%
34	29407-31716	2309	LuxR family transcriptional regulator	ZP_21980543	*Rhodococcus triatomae* BKS 15-14	98%	7.00E-105	37%
35	31717-31983	266	putative flavin-nucleotide-binding protein	ZP_10012406	*Saccharomonospora glauca* K62	53%	2.00E-04	49%

Fosmid 53D1

CDS	Locus	Gene lenght	Description	Best match accession No.	Source organism	Query cover (%)	E-value	Identity (%)
1	330-43	288	transcriptional regulator, LuxR family	ZP_06974075	*Ktedonobacter racemifer* DSM 44963	96%	1.00E-17	54%
2	334-1641	1308	putative Protein kinase/transcriptional regulator, LuxR family	ZP_10243899	*Nitrolancetus hollandicus* Lb	99%	9.00E-54	36%
3	1858-2436	579	conserved hypothetical protein	ZP_10245859	*Nitrolancetus hollandicus* Lb	94%	3.00E-69	68%
4	2485-3999	1514	transposase IS4 family protein	ZP_06965924	*Ktedonobacter racemifer* DSM 44963	97%	3.00E-171	53%
5	4733-4362	371	unknown function	YP_007149953	*Cylindrospermum stagnale* PCC 7417	100%	2.00E-16	38%
6	4935-4792	143	transposase IS66	YP_578978	*Nitrobacter hamburgensis* X14	97%	1.0	43%
7	5968-5363	605	DNAprimase	WP_009117856	*Neisseria shayegani*	41%	0.58	31%
8	6636-6211	425	non-ribosomal peptide synthetase	YP_007362971	*Myxococcus stipitatus* DSM 14675	30%	0.89	50%
9	8591-6636	1955	conserved hypothetical protein	WP_008478050	*Nitrolancetus hollandicus* Lb	82%	9.00E-35	29%
10	8783-8956	173	major facilitator transporter	YP_003336588	*Streptosporangium roseum* DSM 43021	91%	6.0	37%
11	9306-9049	257	hypothetical protein MODMU_3678	YP_006367565	*Modestobacter marinus*	46%	6.00E-05	55%
12	9991-9461	530	succinic semialdehyde dehydrogenase	WP_006508576	*Xenococcus* sp. PCC 7305	12%	0.01	41%
13	11889-10144	1745	hypothetical protein YO5_01076	WP_008071442	*Novosphingobium nitrogenifigens*	95%	5.00E-45	28%
14	12140-12580	440	short chain dehydrogenase/reductase family oxidoreductase	YP_001174360	*Pseudomonas stutzeri* A1501	90%	2.00E-27	45%
15	12580-13782	1202	hypothetical protein HMPREF9473_05085	WP_006783073	*Clostridium hathewayi* WAL-18680	97%	1.00E-134	51%
16	13838-14518	680	beta-lactamase domaincontaining protein	YP_822748	Candidatus *Solibacter usitatus*	94%	5.00E-08	38%
17	14546-14725	179	methyl parathion hydrolase	NP_899941	*Chromobacterium violaceum*	86%	0.032	44%
18	14918-16552	1634	molecular chaperone GroEL	YP_001124352	*Bacillus clausii* KSM-K16	95%	2.00E-80	36%
19	16569-18272	1703	hypothetical protein CLDAP_07830	YP_005440720	*Caldilinea aerophila* DSM 14535	94%	1.00E-21	27%
				BAF02588	**uncultured bacterium**	**94%**	**2.00E-109**	**48%**
20	**18285-19475**	**1190**	**chitinase**	**ZP_23766528**	***Stenotrophomonas maltophilia* AU12-09**	**87%**	**1.00E-89**	**45%**
				ZP_06970122	***Ktedonobacter racemifer* DSM 44963**	**93%**	**6.00E-81**	**41%**

CDS	Locus	Gene lenght	Description	Best match accession No.	Source organism	Query cover (%)	E-value	Identity (%)
21	20756-19491	1265	hypothetical protein Tter_1563	YP_003323291	*Thermobaculum terrenum* ATCC BAA-798	92%	5.00E-45	36%
22	21866-20769	1097	ABC transporter	YP_003323291	*Thermobaculum terrenum* ATCC BAA-799	98%	3.00E-97	45%
23	22572-21871	701	hypothetical protein Sinac_4169	YP_007204071	*Singulisphaera acidiphila* DSM 18658	59%1	1.3	28%
24	23705-22578	1127	sugar ABC transporter ATP-binding protein	YP_001637023	*Thermomicrobium roseum* DSM 5159	86%	4.00E-101	51%
25	24841-23717	1124	sugar ABC transporter ATP-binding protein	YP_002250459	*Dictyoglomus thermophilum* H-6-12	89%	2.00E-85	41%
26	25939-24917	1022	binding-protein-dependent transport systems inner membrane component	ZP_06975261	*Ktedonobacter racemifer* DSM 44963	97%	5.00E-28	32%
27	26997-25942	1055	N acetyl-glucosamine transport	ZP_06975262	*Ktedonobacter racemifer* DSM 44964	78%	5.00E-38	33%
28	28770-27052	1718	extracellular solute-binding protein	YP_001635626	*Chloroflexus aurantiacus* J-10-fl	26%	1.00E-08	37%
29	28986-29252	266	putative flavin-nucleotide-binding protein	WP_005462993	*Saccharomonospora glauca* K62	53%	2.00E-04	49%
30	29932-30129	197	hypothetical protein Namu_4482	YP_003203754	*Nakamurella multipartita* DSM 44233	95%	8.9	35%
31	30282-32591	2309	putative Protein kinase/transcriptional regulator, LuxR family	ZP_10243899	*Nitrolancetus hollandicus* Lb	94%	4.00E-70	35%
32	32770-33237	467	hypothetical protein	Q8ZY16	*Pyrobaculum aerophilum* str. IM2	65%	7.3	26%
33	33432-34613	1181	NHL repeat containing protein	YP_005009929	*Niastella koreensis* GR20-10	94%	1.00E-44	33%
34	35343-35023	320	conserved hypothetical protein	ZP_06588277	*Streptomyces roseosporus* NRRL	30%	5	48%

Table S5

Fosmid 101F8

CDS	Locus	Gene lenght	Description	Best match accession no.	Source organism	Query cover (%)	E-value	Identity (%)
1	208-1221	1013	UDP-glucose 4-epimerase	ZP_11085591	*Aeromonas veronii* AMC35	100%	0.0	99%
2	1792-1292	500	hypothetical protein HMPREF1168_01549	ZP_17191914	*Aeromonas veronii* AMC36	100%	4.00E-101	92%
3	**1801-2922**	**1121**	**chitinase**	**YP_005010140**	***Niastella koreensis* GR20-10**	**100%**	**0.0**	**100%**
4	3115-2885	230	hypothetical protein HMPREF1170_03800	ZP_11085592	*Aeromonas veronii* AMC35	84%	5.00E-37	100%
5	**3283-4152**	**869**	**polysaccharide deacetylase family protein**	**YP_004625545**	***Thermodesulfatator indicus* DSM 15286**	**100%**	**0.0**	**100%**
6	4244-5614	1370	UDP-N-acetylmuramate ligase	ZP_11085593	*Aeromonas veronii* AMC35	100%	0.0	99%
7	5611-6222	611	3-octaprenyl-4-hydroxybenzoate decarboxylase UbiX	YP_004391328	*Aeromonas veronii* B565	100%	5.00E-144	97%
8	6846-6295	551	LitR, transcriptional regulator	YP_004391327	*Aeromonas veronii* B565	100%	1.00E-123	93%
9	7111-7641	530	hypoxanthine phosphoribosyltransferase	ZP_17187671	*Aeromonas veronii* AER39	100%	3.00E-123	98%
10	7753-8670	917	ABC-type multidrug transporter, ATP binding component	ZP_12960923	*Aeromonas salmonicida* 01-B526	100%	0.0	96%
11	8667-9440	773	ABC transporter permease	YP_004391324	*Aeromonas veronii* B565	100%	5.00E-179	99%
12	9694-10113	419	methylamine utilization protein-like protein	YP_004391321	*Aeromonas veronii* B565	100%	8.00E-75	96%
13	10118-10645	527	hypothetical protein B565_0668	YP_004391320	*Aeromonas veronii* B566	100%	5.00E-124	98%
14	12104-10725	1379	hypothetical protein HMPREF1169_03016	ZP_17197498	*Aeromonas veronii* AER397	98%	0.0	96%
15	12599-12249	350	antibiotic biosynthesis monooxygenase domain-containing protein	YP_004391318	*Aeromonas veronii* B565	100%	7.00E-80	99%
16	13481-12624	857	pantothenate synthetase	ZP_17187681	*Aeromonas veronii* AER39	100%	0.0	98%
17	14523-13729	794	3-methyl-2-oxobutanoate hydroxymethyltransferase	YP_004391316	*Aeromonas veronii* B565	100%	0.0	99%
18	15029-14520	509	putative pyrophosphokinase	YP_004391315	*Aeromonas veronii* B565	100%	5.00E-98	96%
19	16471-15041	1430	poly(A) polymerase	YP_858006	*Aeromonas hydrophila* ATCC 7966	95%	0.0	95%
20	17513-16719	794	glutamyl-queuosine tRNA(Asp) synthetase	ZP_11085610	*Aeromonas veronii* AMC35	100%	1.00E-176	93%

CDS	Locus	Gene lenght	Description	Best match accession no.	Source organism	Query cover (%)	E-value	Identity (%)
21	18188-17739	449	RNA polymerase-binding protein DksA	YP_858004	*Aeromonas hydrophila* ATCC 7966	100%	2.00E-100	95%
22	19051-18326	725	sugar fermentation stimulation protein	ZP_11085612	*Aeromonas veronii* AMC35	100%	3.00E-171	95%
23	19575-19048	527	2'-5' RNA ligase (polyribonucleotide synthase (ATP)	ZP_17191893	*Aeromonas veronii* AMC34	100%	2.00E-112	91%
24	19751-21049	1298	ATP-dependent helicase HrpB	ZP_11085614	*Aeromonas veronii* AMC35	92%	0.0	97%
25	**21080-21484**	**404**	**chitinase**	**NP_828094**	***Streptomyces avermitilis* MA-4680**	**82%**	**9.00E-48**	**77%**
26	21511-21786	275	Na+/H+ antiporter	ZP_10511762	*Bacillus mojavensis* RO-H-1	96%	2.00E-11	40%
27	21925-23136	1211	ATP-dependent helicase HrpB	YP_004391309	*Aeromonas veronii* B565	100%	0.0	96%
28	23210-25525	2315	penicillin-binding protein 1B	ZP_17187690	*Aeromonas veronii* AER39	100%	0.0	96%
29	25585-25986	401	Glyoxylase I family protein	YP_004391307	*Aeromonas veronii* B565	100%	1.00E-75	92%
30	26372-26052	320	SMR family multidrug efflux pump	ZP_08521667	*Aeromonas caviae* Ae398	100%	6.00E-45	87%
31	26761-26360	401	SMR family multidrug efflux pump	ZP_11388620	*Aeromonas aquariorum* AAK1	100%	8.00E-72	83%
32	26919-27800	881	LysR family transcriptional regulator	YP_001140693	*Aeromonas veronii* B565	100%	0.0	93%
33	27911-29491	1580	putative sensory transduction protein	YP_004391303	*Aeromonas veronii* B566	100%	0.0	96%
34	29698-30228	530	inorganic pyrophosphatase	YP_001140672	*Aeromonas salmonicida* A449	100%	1.00E-125	99%
35	30604-30257	347	hypothetical protein HMPREF1167_01280	ZP_17187697	*Aeromonas veronii* AER39	100%	1.00E-71	90%
36	30744-31634	890	hypothetical protein HMPREF1168_01518	ZP_17191883	*Aeromonas veronii* AMC34	100%	0.0	91%
37	31855-32397	542	hypothetical protein HMPREF1170_03832	ZP_11085624	*Aeromonas veronii* AMC35	100%	3.00E-123	97%
38	32442-32741	299	hypothetical protein HMPREF1167_01283	ZP_17187700	*Aeromonas veronii* AER39	100%	3.00E-59	89%
39	32773-33357	584	hypothetical protein AHA_3556	YP_858026	*Aeromonas hydrophila* ATCC 7966	100%	1.00E-97	86%
40	33362-34051	689	antimicrobial peptide ABC transporter	YP_004391296	*Aeromonas veronii* B565	100%	5.00E-158	97%

SECTION 2

"*Streptomyces* spp. as alternative heterologous expression systems"

Chapter 3

Binda *et al. BMC Biotechnology* 2013, **13**:24
http://www.biomedcentral.com/1472-6750/13/24

BMC Biotechnology

RESEARCH ARTICLE Open Access

Streptomyces spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance

Elisa Binda[1,2*], Giorgia Letizia Marcone[1,2], Francesca Berini[1,2], Loredano Pollegioni[1,2] and Flavia Marinelli[1,2]

Abstract

Background: $VanY_n$, encoded by the *dbv7* gene (also known as $vanY_n$) of the biosynthetic cluster devoted to A40926 production, is a novel protein involved in the mechanism of self-resistance in *Nonomuraea* sp. ATCC 39727. This filamentous actinomycete is an uncommon microorganism, difficult-to-handle but biotechnologically valuable since it produces the glycopeptide antibiotic A40926, which is the precursor of the second-generation dalbavancin in phase III of clinical development. In order to investigate $VanY_n$ role in glycopeptide resistance in the producer actinomycete an appropriate host-vector expression system is required.

Results: The cloning strategy of $vanY_n$ gene (G-C ratio 73.3%) in the expression vector pIJ86 yielded a recombinant protein with a tag encoding for a histidine hexamer added at the C-terminus (C-His_6-$vanY_n$) or at the N-terminus (N-His_6-$vanY_n$). These plasmids were used to transform three *Streptomyces* spp., which are genetically-treatable high G-C content Gram-positive bacteria taxonomically related to the homologous producer *Nonomuraea* sp.. Highest yield of protein expression and purification (12 mg of protein per liter of culture at 3 L bioreactor-scale) was achieved in *Streptomyces venezuelae* ATCC 10595, that is a fast growing streptomyces susceptible to glycopeptides. $VanY_n$ is a transmembrane protein which was easily detached and recovered from the cell wall fraction. Purified C-His_6-$VanY_n$ showed D,D-carboxypeptidase and D,D-dipeptidase activities on synthetic analogs of bacterial peptidoglycan (PG) precursors. C-His_6-$VanY_n$ over-expression conferred glycopeptide resistance to *S. venezuelae*. On the contrary, the addition of His_6-tag at the N-terminus of the protein abolished its biological activity either *in vitro* or *in vivo* assays.

Conclusions: Heterologous expression of $vanY_n$ from *Nonomuraea* sp. ATCC 39727 in *S. venezuelae* was successfully achieved and conferred the host an increased level of glycopeptide resistance. Cellular localization of recombinant $VanY_n$ together with its enzymatic activity as a D,D-peptidase/D,D-carboxypeptidase agree with its role in removing the last D-Ala from the pentapeptide PG precursors and reprogramming cell wall biosynthesis, as previously reported in glycopeptide resistant pathogens.

Keywords: *Streptomyces*, Heterologous protein production, D,D-carboxypeptidases, Glycopeptide production, Glycopeptide resistance, Dalbavancin

* Correspondence: elisa.binda@uninsubria.it
[1]Department of Biotechnology and Life Sciences, University of Insubria, via J. H. Dunant 3, Varese, Italy
[2]"The Protein Factory" Research Center, Politecnico of Milano, ICRM CNR Milano and University of Insubria, Varese, Italy

Binda et al. BMC Biotechnology 2013, 13:24
http://www.biomedcentral.com/1472-6750/13/24

Background

Filamentous actinomycetes are high G-C Gram-positive microorganisms commercially widely used as producers of natural products (in particular antibiotics) and industrial enzymes [1]. Genome sequencing of representative microbes belonging to this group has shown that they possess a vast array of genes devoted to the production and secretion of enzymes, due to the role they play in recycling organic material in the biosphere [2]. Genome annotation of *Streptomyces coelicolor* [3], which is the model system for this microbial group, revealed that it encodes 819 potentially secreted proteins including hydrolases, proteases/peptidases, chitinases/chitosanases, cellulases/endoglucanases, amylases and pectate lyases. Moreover, filamentous actinomycetes are rich of novel enzymatic functionalities since each genome hosts twenty-thirty biosynthetic gene clusters (20–100 kbp each) devoted to the production of chemically diverse bioactive metabolites [1]. Among filamentous actinomycetes, the most studied genus is the *Streptomyces* one, whose members produce two-thirds of the known antibiotics [4]. *Streptomyces* spp. are also used as host systems for the production of heterologous proteins and of whole biosynthetic clusters originating from less easy-to-handle actinomycetes, such as those belonging to *Nonomuraea*, *Actinoplanes*, *Planomonospora*, *Planobispora*, *Microbispora* genera [1]. These uncommon actinomycetes possess a still-untapped richness of metabolic pathways - and some of them are valuable producers of new drugs - but their exploitation is often limited by the lack of genetic manipulation tools [5,6].

Streptomyces spp. as heterologous hosts for gene expression and protein production offer some advantages in comparison to *Escherichia coli*, the preferred cell factory for industrial enzymes and therapeutic proteins [7,8]. Production and isolation of heterologous proteins from *E. coli* is often limited by insolubility, citotoxicity, uncorrect folding, aggregation in inclusion bodies and lack of secretion [2]. Secretion capability of *Streptomyces* spp. may prevent the local accumulation of the over-expressed recombinant proteins, reduce their toxicity to host cells, eventually aid correct folding and favour increased production and purification yields [9]. Heterologous expression is often facilitated when the selected host cells are phylogeneticaly related to the homologous producer. This is due to the similarity of codon usage, compatibility with translation machinery, molecular chaperons, and/or redox state of the cells [2].

We recently started studying the role of genes involved in the biosynthesis, regulation, transport and self-resistance of glycopeptide antibiotics in producing strains which belong to uncommon actinomycetes such as *Actinoplanes teichomyceticus* producing teicoplanin [10,11] and *Nonomuraea* sp. ATCC 39727, which produces A40926 [12]. A40926 is the precursor of the semi-synthetic derivative dalbavancin, which is a second generation glycopeptide currently in phase III clinical development for its improved activity, pharmacokinetics and pharmacodynamics [13]. The *dbv* cluster, which is a contiguous set of 37 ORFs devoted to the production, regulation and transport of A40926, contains the $vanY_n$ gene, whose function was proposed to confer glycopeptide resistance to the producing strain by reprogramming peptidoglycan cell wall biosynthesis [12,13].

The aim of this work was developing an appropriate heterologous expression system for $VanY_n$ characterization to help deciphering its role in glycopeptide resistant cells. Our attention was given to screening different *Streptomyces* spp. as hosts for recombinant $VanY_n$ secretion in biologically active form. Conditions for production (and purification) of functional $VanY_n$ were finally successfully settled in *Streptomyces venezuelae* ATCC 10595 at flask and at fermentor-scale.

Results

Heterologous expression of $VanY_n$ in *Streptomyces* spp.

$vanY_n$ encoding gene (CAD91202) was amplified by PCR using chromosomal DNA template from *Nonomuraea* sp. ATCC 39727 [12] and cloned in the multicopy vector pIJ86, under the control of the heterologous constitutive promoter *ermE** [14]. The gene was cloned in frame with a sequence encoding for a histidine hexamer added at the C-terminus (C-His_6-$vanY_n$) or at the N-terminus (N-His_6-$vanY_n$) of the protein product.

pIJ86ΩC-His_6-$vanY_n$, pIJ86ΩN-His_6-$vanY_n$, and the empty pIJ86 used as a control, were introduced in three different *Streptomyces* spp. by intergeneric conjugation from *E. coli*. *Streptomyces lividans* TK24 was selected as one of the hosts since it is used for heterologous protein production due to its proven excellence in secretion capacity and low extracellular protease activity [15]. *Streptomyces venezuelae* ATCC 10595 is a fast growing streptomyces naturally susceptible to glycopeptides [12]. *Streptomyces coelicolor* A3(2) represents the model system [3] and possesses a complete set of *vanRSHAX* genes conferring high resistance to vancomycin: consequently a glycopeptide susceptible mutant deleted in the two component regulatory system Δ*vanRS* was used in our experiments [16].

Growth curves at shake flask-scale of the ex-conjugants *S. coelicolor* Δ*vanRS*, *S. lividans* and *S. venezuelae* carrying pIJ86ΩC-His_6-$vanY_n$ or pIJ86ΩN-His_6-$vanY_n$ were compared with control strains containing the empty vector in two different cultivation conditions, i.e. by using the limpid YEME medium or the rich and complex BTSB medium (Figure 1 and Table 1). Figure 1A and 1B show the growth curves of *S. venezuelae* recombinant strains, which produced abundant biomass in both the media. Introduction

Binda *et al. BMC Biotechnology* 2013, **13**:24
http://www.biomedcentral.com/1472-6750/13/24

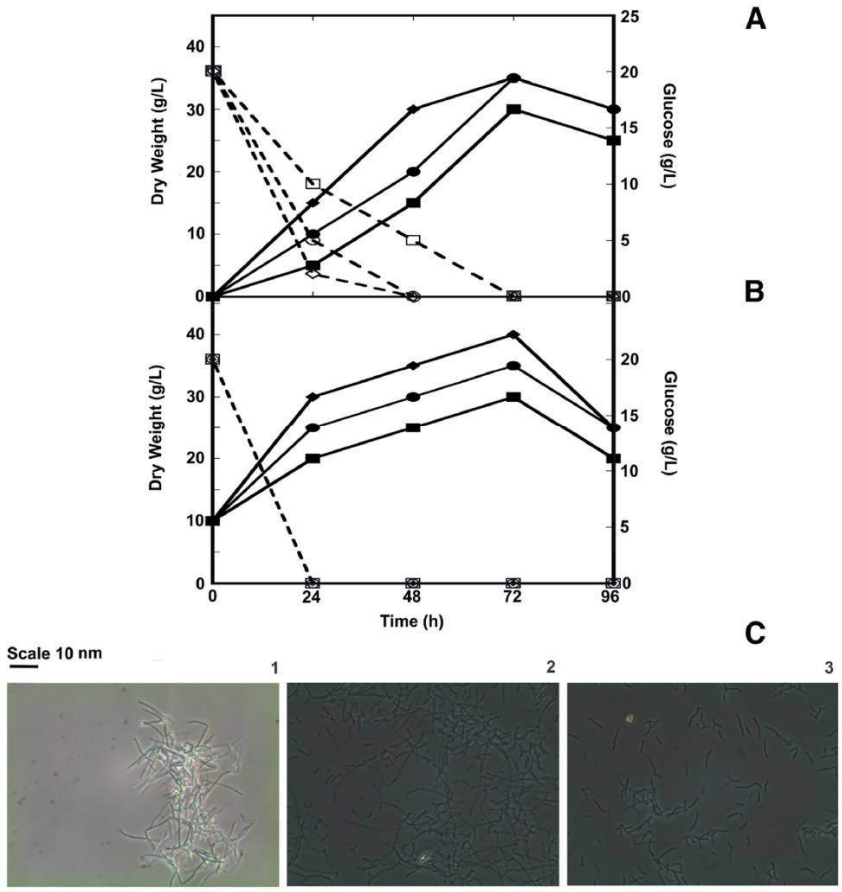


Figure 1 Growth of *S. venezuelae* pIJ86ΩC-His$_6$-*vanY$_n$*, *S. venezuelae* pIJ86ΩN-His$_6$-*vanY$_n$* and *S. venezuelae* pIJ86 in YEME (A) and in BTSB (B) at flask level. *S. venezuelae* pIJ86ΩC-His$_6$-*vanY$_n$* dry weight (■ , solid line) and glucose consumption (□, dashed line); *S. venezuelae* pIJ86ΩN-His$_6$-*vanY$_n$* dry weight (● , solid line) and glucose consumption (○, dashed line); *S. venezuelae* pIJ86 dry weight (♦, solid line) and glucose consumption (◊, dashed line). In (**C**), morphology of *S. venezuelae* pIJ86 (**1**), *S. venezuelae* pIJ86ΩN-His$_6$-*vanY$_n$* (**2**) and *S. venezuelae* pIJ86ΩC-His$_6$-*vanY$_n$* (**3**) grown in YEME for 72 hours at the optical microscopic observation (Zeiss Primo Star phase-contrast microscopy with 40X enlargement).

of *vanY$_n$* slightly affected biomass productivity. Growth kinetics was similar among *S. venezuelae* carrying either C-His$_6$-*vanY$_n$* or N-His$_6$-*vanY$_n$* constructs and the control strain. Maximum biomass production (≥ 30 g/L dry weight) was achieved after 72 hours from the inoculum in both YEME and BTSB media. Glucose was consumed with different kinetics between the two media, but it was anyhow completely depleted within 72 hours of growth. Observation at the optical microscope showed that mycelium in recombinant strains was more fragmented than in the control one (Figure 1C).

Data in Table 1 shows that *S. coelicolor* Δ*vanRS* recombinant strains grew better in BTSB than in YEME, whereas recombinant *S. lividans* strains did not grow in YEME medium. Different effects of medium composition on growth rate and morphology of suspension cultures (i.e. size of mycelium pellets) among diverse streptomyces hosts have been previously observed [14]. Production of His$_6$-VanY$_n$ was evaluated by Western blot analysis after 72 hours of growth in *S. venezuelae*, *S. coelicolor* Δ*vanRS* and *S. lividans* recombinant strains. In BTSB medium, recombinant VanY$_n$ volumetric and specific productivities ranged between 18 to more than 30 mg per liter of culture and from 0.8 to more than 1 mg per gram of cells, within the different recombinant strains. In YEME medium, the highest values of volumetric (about 30 mg of VanY$_n$ per liter of culture) and specific (slightly less than 1 mg of enzyme per gram of cells) productivities were achieved only for *S. venezuelae* recombinant strains. Considering that BTSB medium contains complex components which may later on interfere with protein purification, YEME medium was preferred for

Binda *et al. BMC Biotechnology* 2013, **13**:24
http://www.biomedcentral.com/1472-6750/13/24

Table 1 Biomass and $VanY_n$ production in recombinant *Streptomyces* spp.

Strain	Vector	Medium	Dry weight (g cell/L)	$VanY_n$ (mg/g cell)	$VanY_n$ (mg/L)
S. venezuelae	pIJ86ΩC-His_6-*vanY*$_n$	BTSB	30	0.94	28.4
S. venezuelae	pIJ86ΩC-His_6-*vanY*$_n$	YEME	30	0.97	30.7
S. venezuelae	pIJ86ΩN-His_6-*vanY*$_n$	BTSB	36	1.47	23.9
S. venezuelae	pIJ86ΩN-His_6-*vanY*$_n$	YEME	35	1.10	31.6
S. coelicolor Δ*vanRS*	pIJ86ΩC-His_6-*vanY*$_n$	BTSB	25	1.16	28.9
S. coelicolor Δ*vanRS*	pIJ86ΩC-His_6-*vanY*$_n$	YEME	20	0.62	12.4
S. coelicolor Δ*vanRS*	pIJ86ΩN-His_6-*vanY*$_n$	BTSB	30	1.06	31.8
S. coelicolor Δ*vanRS*	pIJ86ΩN-His_6-*vanY*$_n$	YEME	8	0.79	6.3
S. lividans	pIJ86ΩC-His_6-*vanY*$_n$	BTSB	20	1.02	20.3
S. lividans	pIJ86ΩN-His_6-*vanY*$_n$	BTSB	22	0.80	18

Streptomyces spp. recombinant strains carrying different constructs (C-His_6-*vanY*$_n$ or N-His_6-*vanY*$_n$) were grown in YEME or BTSB for 72 hours. $VanY_n$ was quantified by immunoblotting on crude extracts as reported in the Methods.

large-scale protein preparation. Accordingly, *S. venezuelae* was selected as the preferable host for protein production scaling-up and purification.

Cellular localization of His_6-$VanY_n$ was analyzed after 72 hours of growth in *S. venezuelae*, *S. coelicolor* Δ*vanRS* and *S. lividans* recombinant strains by Western blot analysis on different fractions (prepared as described in the Methods section). Figure 2A shows a ~ 25 kDa band corresponding to His_6-$VanY_n$ (predicted molecular mass of native $VanY_n$ is 22.1 kDa) in the insoluble and soluble

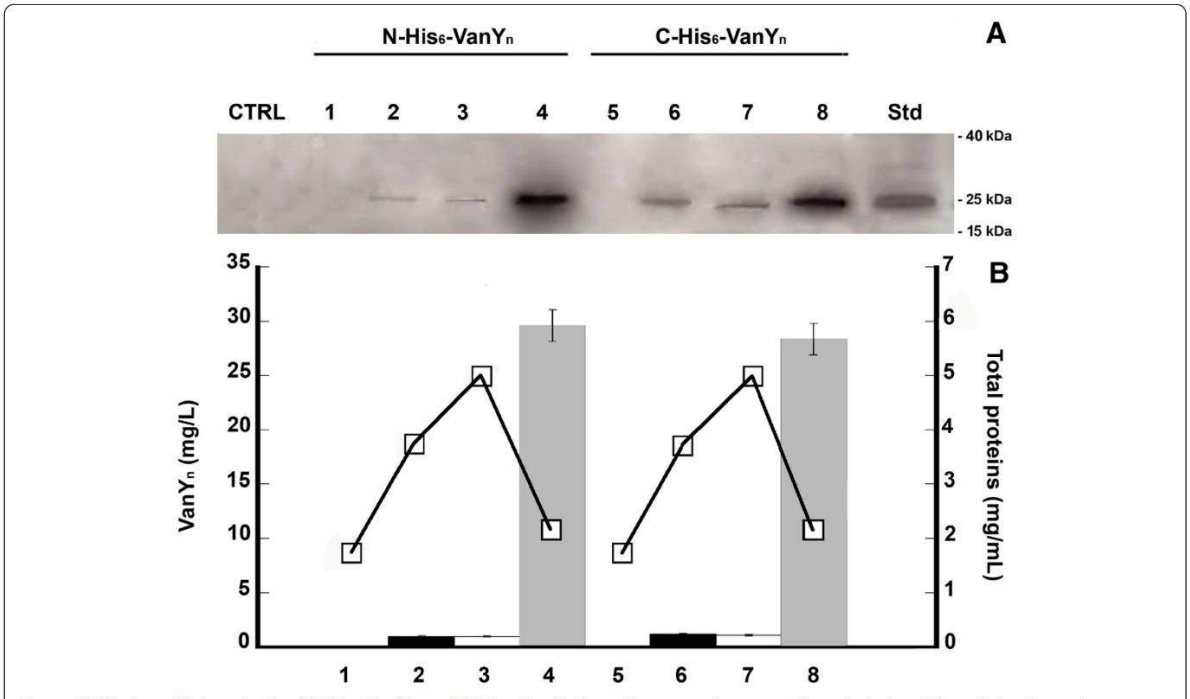


Figure 2 Western blot analysis of C-His_6-$VanY_n$ and N-His_6-$VanY_n$ from *S. venezuelae* recombinant strains. (**A**): analysis of samples corresponding to cellular and extracellular fractions from cultures of *S. venezuelae* recombinant strains grown in YEME for 72 hours. From *S. venezuelae* pIJ86ΩN-His_6-*vanY*$_n$: extracellular fraction (lane 1), soluble cell-free fraction (lane 2), insoluble cell-free fraction (lane 3) and cell wall fraction (lane 4); from *S. venezuelae* pIJ86ΩC-His_6-*vanY*$_n$: extracellular fraction (lane 5), soluble cell-free fraction (lane 6), insoluble cell-free fraction (lane 7) and cell wall fraction (lane 8). A crude extract from *S. venezuelae* cells carrying the empty vector was loaded as a negative control (CTRL). In each lane, samples corresponding to 100 μL of cell culture were loaded. Std reference protein: C-His_6-$VanY_n$ from *E. coli* (5 μg, 25 kDa). The corresponding SDS-PAGE is shown in the Additional file 1: Figure S1. (**B**): quantitative analysis of recombinant $VanY_n$ present in each fraction, performed using the Quantity One program (Bio-Rad Laboratories, Milan, Italy) and His_6-$VanY_n$ as comparative standard. Total protein content in each fraction was determined by Biuret method (□, solid line).

Binda *et al. BMC Biotechnology* 2013, **13**:24
http://www.biomedcentral.com/1472-6750/13/24

cell-free fractions and in the cell wall fractions from *S. venezuelae* recombinant strains grown in YEME medium. Recombinant $VanY_n$ was never found in the concentrated broths (extracellular fractions, lanes 1 and 5 of Figure 2A). Densitometric analysis demonstrated that His_6-$VanY_n$ preferentially accumulated (90%, lanes 4 and 8 of Figure 2B) in the cell wall fractions obtained by the step of spheroplast preparation, independently on the localization of the His_6-tag (at the C-terminus or at the N-terminus of the protein). Only 6% and 4% of recombinant protein was detected in insoluble and soluble cell-free fractions following spheroplast burst, respectively. For all recombinant strains grown in BTSB or YEME (Table 1), the heterologous protein distribution was exactly as for *S. venezuelae*, with most of $VanY_n$ recovered from the cell wall fraction (data not shown).

$VanY_n$ production by *S. venezuelae* in 3 L bench-bioreactor scale

Both the forms of recombinant His_6-$VanY_n$ were produced growing *S. venezuelae* recombinant strains in YEME medium at 3 L bench-bioreactor scale. Figures 3A, B and C show the time course of growth and $VanY_n$ production of *S. venezuelae* carrying pIJ86ΩC-His_6-*vanY*$_n$. Exponential growth was actually completed within 48 hours and gave a maximum biomass yield of 40 g/L dry weight. This growth phase was accompanied by a complete depletion of glucose from the medium and by a transient reduction of dissolved oxygen (pO_2). Medium pH slightly decreased during the exponential growth phase and then tended to increase during the stationary growth phase. An estimated maximum of production of ~ 30 mg $VanY_n$/L in the crude extract was determined at the peak of biomass production, i.e. at 48 hours from inoculum. An overlapping profile of growth curve and $VanY_n$ production was similarly obtained for *S. venezuelae* carrying pIJ86ΩN-His_6-*vanY*$_n$ (not shown).

$VanY_n$ purification from *S. venezuelae*

Purification of both the recombinant forms of His_6-$VanY_n$ from *S. venezuelae* cells grown in YEME medium, was attempted by means of metal-chelating chromatography on HiTrap chelating column of crude extracts obtained by whole cell sonication, as described in the Methods section. N-His_6-$VanY_n$ was eluted following standard procedures. In the same conditions, C-His_6-$VanY_n$ did not bind to the column, suggesting that the six histidines at C-terminus were masked. Accordingly, a denaturing agent such as urea was added to both the loading and equilibration buffers. The partially denatured C-His_6-$VanY_n$ protein interacted with the matrix and was subsequently re-folded directly on the column by a linear gradient of decreasing urea concentration. SDS-PAGE analysis (Figure 4) confirmed that both the tagged forms of His_6-$VanY_n$ migrated as a single band of 25 kDa and were > 90% pure. Final purification yield was comparable between the two forms: approximately 12 mg of C- or N-His_6-$VanY_n$ protein from one liter of culture were recovered.

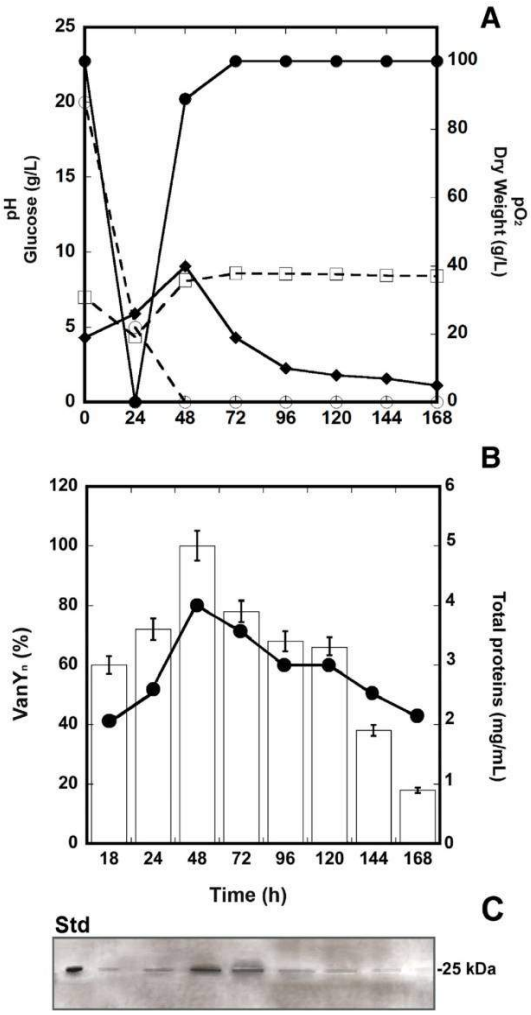


Figure 3 Growth curve and $VanY_n$ production in 3-L batch fermentations of *S. venezuelae* pIJ86ΩC-His_6-*vanY*$_n$ in YEME. (**A**) Time courses of pH (□, dashed line), pO_2 (●, solid line), glucose (○, dashed line), and growth curve measured as dry weight (♦, solid line). (**B**) Time course of recombinant $VanY_n$ production as determined by quantitative Western blot analysis of crude extracts from cell samples collected at different times of fermentation (see panel **C**). Maximum protein expression at 48 hours of growth was set as 100% (corresponding to ~ 30 mg/L of culture). Total protein content in each fermentation sample was determined by Biuret method (● , solid line). In (**C**), samples corresponding to 100 μL of cell culture were loaded in each lane; Std, standard reference protein: C-His_6-$VanY_n$ from *E. coli* (5 μg, 25 kDa). The corresponding SDS-PAGE is shown in the Additional file 2: Figure S2.

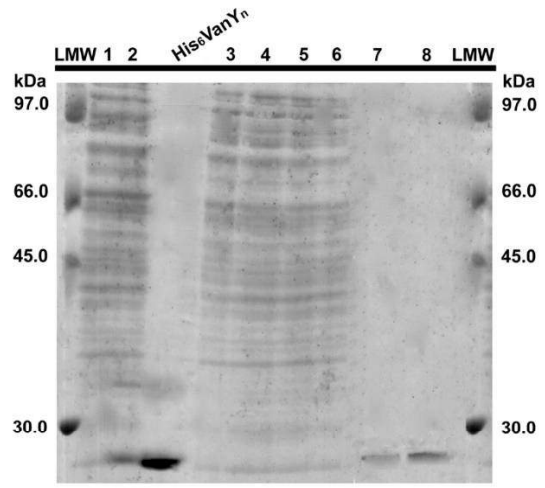


Figure 4 Purification of recombinant VanY$_n$. SDS-PAGE Analysis (staining with Coomassie brilliant blue) of protein samples collected during purification of VanY$_n$ from *S. venezuelae* carrying pIJ86ΩN-His$_6$-*vanY$_n$* or pIJ86ΩC-His$_6$-*vanY$_n$*. Crude fractions from *S. venezuelae* carrying pIJ86ΩN-His$_6$-*vanY$_n$* (lane 1) and *S. venezuelae* carrying pIJ86ΩC-His$_6$-*vanY$_n$* (lane 2). Fractions from HiTrap Chelating chromatography: N-His$_6$-VanY$_n$ flow through (lane 3), C-His$_6$-VanY$_n$ flow through at 6 M (lane 4), 4 M (lane 5) and 1 M (lane 6) urea. Purified N-His$_6$-VanY$_n$ (lane 7, 2 μg) and C-His$_6$-VanY$_n$ (lane 8, 4 μg) after desalting gel permeation. The purified VanY$_n$ migrated as a single band at a molecular mass of ~ 25 kDa and with > 90% purity. Standard reference protein: C-His$_6$-VanY$_n$ from *E. coli* BL21(DE3) Star (5 μg, 25 kDa). LMW, molecular mass standard proteins (Amersham, GE-Healthcare).

Enzymatic activity of pure recombinant VanY$_n$

Activities of purified C-His$_6$-VanY$_n$ and N-His$_6$-VanY$_n$ from *S. venezuelae* recombinant strains were assayed on commercially available surrogates of peptidoglycan precursors, in parallel with the previously characterized C-His$_6$-VanY$_n$ produced by recombinant *E. coli* [17]. These results were obtained by the D-amino acid oxidase/peroxidase colorimetric coupled reaction [18,19], and confirmed by a fluorimetric assay [20]. As shown in Table 2, C-His$_6$-VanY$_n$ from *S. venezuelae* cleaved the last D-Ala from the tripeptide N$_\alpha$,N$_\varepsilon$-diacetyl-L-Lys-D-Ala-D-Ala and acetyl-L-Lys-D-Ala-D-Ala, this activity being only slightly affected by the acetylation grade of the Lys. The activity was halved if the substrate was the D-Ala-D-Ala. Thus, C-His$_6$-VanY$_n$ showed a higher D,D-carboxypeptidase activity (VanY-like) than a D,D-peptidase activity (VanX-like). Surprisingly, N-His$_6$-VanY$_n$ from *S. venezuelae* did not show any activity on the three substrates (Table 2), suggesting that the tag position dramatically influenced enzyme competence.

These results were supported by the circular dichroism (CD) spectra of the recombinant proteins. CD spectrum of C-His$_6$-VanY$_n$ produced in *S. venezuelae* overlapped with that of the protein produced in *E. coli* (Figure 5): analysis of secondary structure indicates a predominance of β-sheets (~ 38%) and ~ 15% of α-helices. This structure content was altered for the N-His$_6$-VanY$_n$ (Figure 5), indicating that N-terminal tag interfered with the proper protein folding and secondary structure content of the recombinant enzyme.

Table 2 Substrate specificity of His$_6$-VanY$_n$ recombinant forms produced by different hosts

Substrate	*S. venezuelae*		*E. coli*
	C-His$_6$-VanY$_n$ (U/mg)	N-His$_6$-VanY$_n$ (U/mg)	C-His$_6$-VanY$_n$ (U/mg)
D-Ala-D-Ala	18 ± 5.0	0	19 ± 3
N$_\alpha$,N$_\varepsilon$-diacetyl-L-Lys-D-Ala-D-Ala	38 ± 3.8	0	36 ± 4.0
Acetyl-L-Lys-D-Ala-D-Ala	40 ± 5.6	0	40 ± 3.0

Specific activity (U/mg protein) was determined using 40 μg of His$_6$-VanY$_n$ added to 10 mM solutions of the indicated compounds. The activity was assayed at 25°C as described in the Methods section. Results are the average of three independent experiments.

Resistance phenotype in *S. venezuelae* recombinant strains

Role of different constructs of *vanY$_n$*, when over-expressed in a glycopeptide-susceptible heterologous host such as *S. venezuelae*, was investigated *in vivo* by determining the D,D-peptidase/D,D-carboxypeptidase activity of alkaline cell extracts and the glycopeptide resistance phenotype of recombinant strains. As shown in Table 3, D,D-peptidase/D,D-carboxypeptidase activity was detectable only in *S. venezuelae* pIJ86ΩC-His$_6$-*vanY$_n$* and *S. venezuelae* pIJ86Ω*vanY$_n$* (a strain with *vanY$_n$* cloned

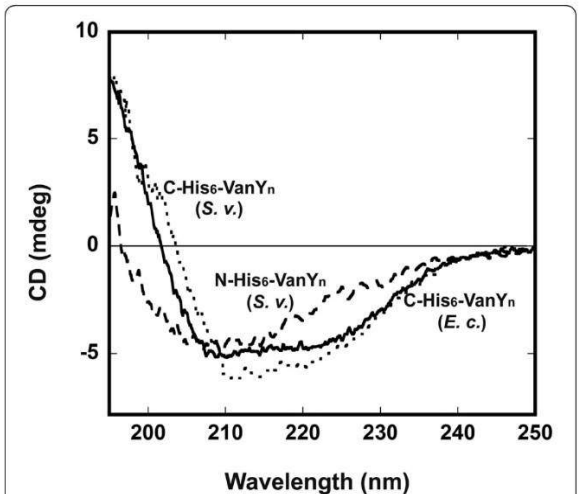


Figure 5 Far-UV CD spectra of recombinant VanY$_n$. C-His$_6$-VanY$_n$ (dotted line) and N-His$_6$-VanY$_n$ (dashed line) from recombinant *S. venezuelae* (*S. v.*); C-His$_6$-VanY$_n$ from recombinant *E. coli* (*E. c.*) (continuous line). Protein concentration was 0.1 mg/mL, 15°C.

Binda *et al. BMC Biotechnology* 2013, **13**:24
http://www.biomedcentral.com/1472-6750/13/24

Table 3 Resistance and enzyme activity profiles of *S. venezuelae* strains

	MIC		Enzyme activity	
Strain	Vancomycin (μg/mL)	Teicoplanin (μg/mL)	D,D-carboxypeptidase (U/mg)	D,D-peptidase (U/mg)
wild-type	0.2	0.5	0	0
pIJ86	0.2	0.5	0	0
pIJ86ΩC-His$_6$-*vanY$_n$*	0.4	0.7	189 ± 5.5	100 ± 3.9
pIJ86ΩN-His$_6$-*vanY$_n$*	0.2	0.5	0	0
pIJ86Ω*vanY$_n$*	0.4	0.7	172 ± 4.3	80 ± 3.2

MICs of glycopeptides were determined by plating 10^6 cfu/mL of each strain on agar medium added with different antibiotic concentrations. Results are the average of three independent experiments in which the standard deviation was less than 5%. Specific activity (U/mg protein) was determined in alkaline cell extracts added to 10 mM solutions of acetyl-L-Lys-D-Ala-D-Ala or D-Ala-D-Ala. The activity was assayed at 25°C as described in the Methods section [17]. Results are the average of three independent experiments.

without any tag, see [12]), but not in *S. venezuelae* pIJ86ΩN-His$_6$-*vanY$_n$* neither in *S. venezualae* wild-type or in *S. venezualae* containing the empty vector. Table 3 also reports the minimal inhibitory concentrations (MICs) of vancomycin and teicoplanin to the recombinant strains in comparison with the wild-type. Accordingly to the expression of the D,D-peptidase/D,D-carboxypeptidase activity, *S. venezuelae* expressing C-His$_6$-VanY$_n$ and *S. venezuelae* producing VanY$_n$ without any tag showed an increased resistance level against both glycopeptides. Resistance phenotype of *S. venezuelae* producing the N-His$_6$-VanY$_n$ was the same as in the control strain with empty vector or in the wild-type.

These data confirm the role of VanY$_n$ in conferring glycopeptide resistance to a susceptible host and indicate that C-terminal His$_6$-VanY$_n$ behaved *in vivo* as the native VanY$_n$, whereas the addition of His$_6$ at the N-terminus of the protein abolished its biological activity.

Discussion

C-His$_6$-VanY$_n$ was previously produced in BL21(DE3) Star *E. coli* cells as a cytoplasmatic soluble protein [13]. In that case, codon usage optimization of the synthetic gene was essential since *E. coli* is a low G-C ratio (~ 50%) Gram-negative bacterium whereas *Nonomuraea* sp. is a high G-C Gram-positive actinomycete. The G-C ratio of *vanY$_n$* is 73.3% which is similar to the genomic ratio of the reference *S. coelicolor* A3(2) (72%) and *S. venezuelae* ETH14630 (71%) [14]. The G-C ratio of the *dbv* cluster is 70.4%. Notwithstanding our efforts in optimizing cultivation, induction and purification conditions, the best productivity achieved in recombinant *E. coli* was 4.6 mg/L culture (0.13 mg/g cells) [17]. Similar level of expression was obtained by other authors who cloned *vanXY$_C$* from *Enterococcus gallinarum* BM4174 in *E. coli* JM83 [21], *vanX* from *Enterococcus faecium* BM4147 in *E. coli* W3110 [22] and *vanY* from *Enterococcus faecium* BM4147 in *E. coli* JM83 [20]. VanXY$_C$, VanX and VanY are D,D-carboxypeptidases/D,D-peptidases homologous to VanY$_n$ involved in conferring glycopeptide resistance to enterococci, which are Gram-positive pathogens with low G-C content [23,24].

By cloning *vanY$_n$* from *Nonomuraea* sp. ATCC 39727 in the taxonomically closely related *S. venezuelae*, pure protein production was at least three folds higher (12 mg/L culture) than in *E. coli*. Since estimated volumetric productivity in the harvested cells was around 30 mg of protein per liter of culture, there is still room to improve the final yield by further optimization of the purification procedure. Estimated specific productivity of around 1 mg VanY$_n$ per gram of cells in streptomycetes was much higher than the one achieved in *E. coli* (0.13 mg/g cells) confirming that streptomycetes are preferable hosts for VanY$_n$ production. Comparable biomass and protein production were achieved at flask and at 3 L bioreactor-scale: maximum values being one day anticipated in bioreactor runs. Fermentation and downstream technologies are well developed for streptomycetes, which are industrially relevant microorganisms producing many valuable antibiotics and other bioactive metabolites [1]; thus they can be successfully further applied to improving and scaling-up VanY$_n$ production.

An important difference between heterologous production of VanY$_n$ in recombinant streptomycetes and *E. coli* is the preferential subcellular localization of the protein. VanY$_n$ (196 amino acids) contains three different putative domains: a cytoplasmatic domain at the N-terminus (the first 20 amino acids), an hydrophobic trans-membrane portion (20 amino acids), followed by the C-terminal domain exposed on the external face of cytoplasmatic membrane. This extracellular domain contains conserved motifs (SxHxxGxAxD and ExxH) for the coordination of zinc ions and the catalytic active site [23,25,26]. In *E. coli* recombinant cells, C-His$_6$-tagged-VanY$_n$ was accumulated in the cytoplasm, whereas in streptomycetes the most (90%) of the tagged enzyme was recovered from the cell wall fraction. The localization of VanY$_n$ confirms its role in the extracellular phases of cell wall biosynthesis [23,25]. The addition of detergents did not improve purification from the whole cell extract, confirming that the protein is easily

Binda *et al. BMC Biotechnology* 2013, **13**:24
http://www.biomedcentral.com/1472-6750/13/24

detached from membranes. VanY D,D-carboxypeptidases from glycopeptide resistant enterococci are in fact involved in the extracytoplasmatic hydrolysis of the last D-Ala from the UDP-pentapeptide PG precursors [20].

The enzymatic activity and the spectral investigation (by CD spectroscopy) confirmed the identity of C-His_6-$VanY_n$ produced by recombinant cells of either *S. venezuelae* or *E. coli*. On substrates that mimic peptidoglycan precursors, $VanY_n$ showed D,D-carboxypeptidase and D,D-dipeptidase activity, resembling more $VanXY_C$ from VanC-type *Enterococcus gallinarum* [21] than typical VanY D,D-carboxypeptidases and VanX D,D-peptidases firstly characterized in *Enterococcus faecium* BM4147 [20,27,28]. On the other hand, the introduction of an histidine hexamer at the N-terminus of $VanY_n$ abolishes the enzymatic activity and alters protein secondary structure, suggesting that it interferes with the folding of the active protein. Till now, VanY family of D,D-carboxypeptidases have not been structurally investigated. The possibility to produce discrete amounts of pure $VanY_n$ by using *S. venezuelae* as a cell factory open the way to better investigate the peculiar bifunctional activity of this D,D-dipeptidase/D,D-carboxypeptidase and its interaction with substrates and inhibitors.

Finally, the functional study of recombinant $VanY_n$ activity in cellular extracts confirmed the role of this protein in conferring glycopeptide resistance in a susceptible host such as *S. venezuelae*, which lacks those *vanRSHAX* genes commonly considered essential to confer glycopeptide resistance [24]. Our data show that when $vanY_n$ gene was produced in the biologically active form (C-His_6-$VanY_n$ or $VanY_n$ without any tag), its expression conferred resistance in the absence of *vanRSHAX* genes, albeit at a reduced level, as previously demonstrated in the homologous producer *Nonomuraea* sp. ATCC 39727 [12]. As proof of this, expression of the inactive N-His_6-$VanY_n$ form did not increase the level of resistance towards two glycopeptides (vancomycin and teicoplanin) of the recombinant hosts, which remains the same as in the wild-type and as in the control recombinant strain containing only the empty vector.

Conclusions

Streptomyces spp. were demonstrated valuable hosts for the production of specific peptidases, involved in cell wall turnover of glycopeptide resistant microbial cells. The novel D,D-dipeptidase/D,D-carboxypeptidase $VanY_n$ from the glycopeptide producer *Nonomuraea* sp. ATCC 39727 was successfully produced in a good yield and in the biologically active form conferring resistance to the glycopeptide susceptible *S. venezuelae* strain. Developing such streptomyces cell factory system for $VanY_n$ production opens the way to a further characterization of the enzyme, to a better comprehension of its role in glycopeptide resistance, and to its use as novel biocatalyst.

Methods

Strains and media

Nonomuraea sp. ATCC 39727 was maintained and cultivated to prepare genomic DNA according to [12]. *Streptomyces coelicolor* Δ*vanRS* was gently donated by Hee-Jeon Hong, University of Cambridge, UK [16], *Streptomyces lividans* TK24 and *Streptomyces venezuelae* ATCC 10595 were a gift from Mervyn Bibb, John Innes Institute, Norwich, UK [14]. *Streptomyces* spp. strains were maintained as spores in 10% (v/v) glycerol and propagated in MYM and SFM agar media [14]. For growing ex-conjugants containing pIJ86, pIJ86ΩC-His_6-$vanY_n$, pIJ86ΩN-$vanY_n$ and pIJ86Ω$vanY_n$, agar plates were added with 50 μg/mL of apramycin (Sigma-Aldrich, Milan, Italy). Agar plates were incubated at 28°C. Liquid media for streptomycetes were YEME -containing in (w/v) 0.3% yeast extract, 0.5% bacto-peptone, 0.3% oxoid malt extract, 1% glucose in deionized water, pH 7.0- and BTSB -containing in (w/v) 10% sucrose, 1% yeast extract, 1% glucose, 0.5% NaCl, 0.5% soybean meal, 1.7% tryptone and 0.25% K_2HPO_4 in deionized water, pH 7.0. All medium components were from Sigma-Aldrich (Milan, Italy), unless otherwise stated. Colonies were picked up from agar plates and inoculated into 300 mL baffled flasks containing 50 mL of YEME or BTSB. Flask cultures were incubated on a rotary shaker at 200 rpm and 28°C. Media and culture conditions for *E. coli* were described in [12]. *E. coli* DH5α was used as host for plasmid construction. *Escherichia coli* ET12567/pUZ8002 [5] was used as non-methylating plasmid donor strain for intergeneric conjugation with *Streptomyces* spp. Cells were propagated in Luria-Bertani (LB) (Sigma-Aldrich, Milan, Italy) broth at 37°C.

Plasmids

pIJ86 (gift from Mervyn Bibb) was used as a multicopy vector for heterologous expression in *Streptomyces* spp. [12]. Plasmids pIJ86ΩC-His_6-$vanY_n$ and pIJ86ΩN-His_6-$vanY_n$ were constructed as follow. Expand High Fidelity polymerase (Roche, Milan, Italy) was used to amplify $vanY_n$ using genomic DNA of *Nonomuraea* sp. ATCC 39727 as template with oligonucleotide primers *vanY*004*Fw* (5′-ATA**GGATCC**CCA GACTGGAGGAGAGGGATGAGGAGAAGCGAGGG TGAC-3′) and *vanY*004*Rev* (5′-GAT**AAGCTT**CTAG TGGTGGTGGTGGTGGTGGACCCGGCCCCCGTTC CGGCT-3′) that introduced a C-terminal tag of six histidine residues and the *HindIII* and *BamHI* (Roche, Milan, Italy) restriction sites, respectively, into the PCR product, allowing insertion into the multiple cloning site of the multicopy expression vector pIJ86. The

Binda *et al. BMC Biotechnology* 2013, **13**:24
http://www.biomedcentral.com/1472-6750/13/24

vanY$_n$ with a N-terminal tag of six histidines was produced using as oligonucleotide primers *vanY*003*Fw* (5′-ATATTT**GGATCC**ATGCACCACCACCACCACCACAG GAGAAGCGAGGGTGACGAC-3′) and *vanY*003*Rev* (5′-GAT**AAGCTT**CCCGTGCCCTAGCTAGACCCGGCCCC CGTTCCGGCT-3′) that introduced *HindIII* and *BamHI* restriction sites. The *vanY*$_n$ without any tag was produced using as oligonucleotide primers *vanY86Fw* (5′-AT**GGATCC**CAGACTGGAGGAGAGGGATG-3′) and *vanY86Rev* (5′-GAT**AAGCTT**CGATCCTGGAGTTCGTC TTC-3′) that introduced *BamHI* and *HindIII* restriction sites [12]. The PCR products were purified, digested with *HindIII* and *BamHI*, and ligated with pIJ86 vector that had similarly been digested, to produce pIJ86ΩC-His$_6$-*vanY*$_n$, pIJ86ΩN-His$_6$-*vanY*$_n$ and pIJ86Ω*vanY*$_n$. These vectors with *vanY*$_n$ transcribed from the strong constitutive *ermE** promoter [14] were used to transform *E. coli* ET12567/pUZ8002 cells.

Intergeneric conjugation

Intergeneric conjugation was performed according to a modified protocol from [5,14]. In brief, a culture of the donor *E. coli* ET12567/pUZ8002 containing the selected plasmid was grown in 10 mL LB supplemented with 50 μg/mL apramycin, 25 μg/mL chloramphenicol (Sigma-Aldrich, Milan, Italy) and 50 μg/mL kanamycin (Sigma-Aldrich, Milan, Italy) to an OD_{600nm} of 0.4. Cells were collected by centrifugation, washed twice with an equal volume of LB, and resuspended in 1 mL of LB. For each conjugation approximately 10^8 *Streptomyces* spp. spores, collected from agar plates in sterile glycerol, were added to 500 μL 2X YT broth [14], heat shocked at 50°C for 10 minutes and then allowed to cool. 500 μL of *E. coli* cells were added to 500 μL of heat-shocked spores and mixed briefly. Mixture was plated out on SFM or MYM agar added with 10 mM $MgCl_2$ and incubated for 20 hours at 30°C. Plates were overlaid with 1 mL of water containing 500 μg/mL nalidixic acid (Sigma-Aldrich, Milan, Italy) and 50 μg/mL apramycin and were incubated at 30°C until colonies appear.

Colony PCR and sequencing

The presence of pIJ86Ω*vanY*$_n$, pIJ86ΩC-His$_6$-*vanY*$_n$ or pIJ86ΩN-His$_6$-*vanY*$_n$ was checked by colony PCR and DNA sequencing of ex-conjugants. Single colonies were transferred onto DNA medium [14]. Plates were incubated at 30°C overnight. Mycelium was scraped from the plates using a sterile toothpick and was introduced into 50 μL 100% v/v DMSO in a 1.5 mL tube. The tube was shaken vigorously for 1–2 hours and then centrifuged: 2.5 μL of the supernatant were used for PCR. For control, 1 μL of DNA samples were mixed with 1.5 μL DMSO. One initial step of 10 minutes at 95°C was included in the PCR program to ensure the complete cell lysis. The PCR was performed for 30 cycles as follows: 95°C for 45 seconds, 56°C for 30 seconds and 72°C for 1 minute. *vanY*004*Fw*, *vanY*004*Rev*, *vanY*003*Fw* and *vanY*003*Rev*, *vanY86Fw* and *vanY86Rev* were used as oligonucleotide primers.

VanY$_n$ expression

Recombinant *Streptomyces* spp. were grown aerobically in 100 mL YEME and BTSB media added with apramycin (50 μg/mL) in 500 mL Erlenmayer flasks for different time intervals (up to 196 h) at 28°C and 200 rpm. Cells were collected by centrifugation at 39000 × g for 15 minutes and washed three times with water. The supernatant (named "extracellular fraction") was collected and precipitated with the trichloroacetic acid (TCA) method. TCA precipitation was performed adding 1/10 (v/v) of 100% TCA (w/v) to an appropriate medium volume and vortexed for 15 seconds, placed on ice for 15 minutes and then centrifugated at 14,000 × *g* for 10 minutes. The supernatant was removed and discarded. The pellet was washed twice with 100 μL of pure acetone and then air dried for about 60 minutes.

Washed cells were suspended in 10 mL of 0.1 M Tris–HCl buffer (pH 7.6) containing 0.4 M sucrose and 6.7 mg of lysozyme (Sigma-Aldrich, Milan, Italy). After 6 hours of incubation at 37°C, the mixture was centrifuged at 12800 × *g* for 10 minutes, obtaining spheroplasts clearly distinguishable at the optical microscopic observation (40X, Zeiss Primo Star microscope, Arese, Italy). The supernatant thus obtained was named "cell wall fraction". The spheroplasts were washed with 0.4 M sucrose and then 5 mL of water was added to burst the spheroplasts. The suspension was centrifuged at 2000 × *g* for 30 minutes. The supernatant and the precipitate were named as "cell-free soluble extract" and "cell-free insoluble fraction", respectively.

Scale up in 3-L reactor

Flask cultures of recombinant *Streptomyces* spp. grown in YEME and apramycin (50 μg/mL) for 48–72 hours were used to inoculate - at 2.5% (v/v) - 3-L P-100 Applikon glass reactor (height 25 cm, diameter 13 cm) equipped with a AD1030 Biocontroller and AD1032 motor, containing 2 L YEME and apramycin (50 μg/mL). Cultivations in fermenter were carried out for 168 hours at 28°C, with stirring at 500 rpm (corresponding to 1.17 m/s of tip speed) and 2 L/min aeration rate. Dissolved oxygen (measured as % of the initial pO_2 value) was monitored using an Ingold polarographic oxygen electrode. The pH values of culture broths were monitored using a pH meter. Foam production was controlled by adding Antifoam SE-15 (Sigma-Aldrich, Milan, Italy) through an antifoam sensor.

Binda et al. BMC Biotechnology 2013, 13:24
http://www.biomedcentral.com/1472-6750/13/24

$VanY_n$ purification

For protein purification, cultures of recombinant *S. venezuelae* grown in YEME medium were harvested and centrifuged at 39000 × *g* for 15 min at 4°C. Cells were sonicated in a buffer solution 300 mM NaCl, 30 mM imidazole, 10 μg/mL DNAseI (Sigma-Aldrich, Milan, Italy), 0.19 mg/mL PMSF (Sigma-Aldrich, Milan, Italy) and 0.7 μg/mL pepstatine (Sigma-Aldrich, Milan, Italy) in 50 mM potassium phosphate buffer (pH 7.0), for 15 cycles of 30 seconds each on ice. After a centrifugation at 39000 × *g* for 60 min at 4°C, the supernatant was collected. N-His_6-$VanY_n$ was purified by affinity chromatography onto a HiTrap chelating affinity column (1.6 × 2.5 cm, 5 mL, GE Healthcare Sciences, Milan, Italy) equilibrated with 50 mM potassium phosphate buffer (pH 7.0) containing 30 mM imidazole and 300 mM NaCl, according to the manufacturer's instructions. After extensive washing, the bound protein was eluted with 50 mM potassium phosphate buffer (pH 7.0) containing 300 mM NaCl and 300 mM imidazole [17]. For C-His_6-tagged $VanY_n$, 6 M urea was added to the equilibration buffer as denaturant; in order to refold the protein bound to the column, a linear gradient from 6 to 0 M of urea (2 mL/min) was performed before starting the elution phase [29]. Fractions containing pure recombinant $VanY_n$ were loaded on PD10 Sephadex G25 column (Ge Healthcare Sciences, Milan, Italy) equilibrated with 50 mM potassium phosphate buffer (pH 7.0). Protein purity was checked by SDS-PAGE (using 15% polyacrylamide gels and staining with Coomassie brilliant blue) and Western blot analysis. Protein concentration was estimated using the extinction coefficient at 280 nm (45258 M^{-1} cm^{-1}) determined by urea denaturation and the theoretical extinction coefficient based on amino acid sequence.

Western blot analysis

Following electrophoresis of proteins from bacterial cell fractions (corresponding to 100 μL of culture) or affinity-purified fractions and the transfer to a nitrocellulose sheet (GE Healthcare Sciences, Milan, Italy), the membrane was incubated with 1:1000 (v/v) His•Tag® Antibody HRP Conjugate Kit (Novagen, Milan, Italy) in alkalin-casein solution 1% (v/v). The immunorecognition was visualized by ECL Detection Reagents (GE Healthcare Sciences, Milan, Italy). The quantitative analysis was performed using the bioinformatics program Quantity One (Bio-Rad Laboratories, Milan, Italy) and C-His_6-$VanY_n$ from *E. coli* as standard protein. For molecular mass determination, PageRuler™ Prestained Protein Ladder (Thermo Scientific, Milan, Italy) markers were used. Protein content in each fraction was assayed by Biuret method and SDS-PAGE analysis (Additional file 1: Figure S1 and Additional file 2: Figure S2).

Alkaline extraction of D,D-carboxypeptidase

All manipulations were carried out at 0 to 4°C. Cells at different growth phases were harvested and suspended in 2 ml per gram of cells of physiological solution (0.9% (v/v) NaCl). The mycelium was fragmented by mild sonication and cells were collected by centrifugation at 39,000 × *g* for 15 minutes. Alkaline extractions were carried out by suspending the cell suspension in ice-cold distilled water containing the proteinase inhibitors (0.19 mg/ml PMSF (Sigma-Aldrich, Milan Italy) and 0.7 μg/ml pepstatine (Sigma-Aldrich, Milan Italy) and then bringing the suspension to pH 12 by adding an appropriate volume of 2.5 N NaOH. After centrifugation (28,000 × *g*, 15 min, 4°C), the supernatants were neutralized to pH 7 by the addition of 0.5 M sodium acetate (pH 5.4) [30].

D,D-dipeptidase and D,D-carboxypeptidase assays

Enzyme activity was assayed by measuring the release of D-Ala from N_α,N_ε-diacetyl-L-Lys-D-Ala-D-Ala, acetyl-L-Lys-D-Ala-D-Ala and D-Ala-D-Ala by a D-amino acid oxidase/peroxidase coupled colorimetric assay (i) or by reaction with a fluorescent reagent (ii). All substrates were purchased from Sigma-Aldrich. One unit of D, D-carboxypeptidase activity was defined as the amount of enzyme that produced 1 μmol of D-Ala per min.

(i) D-Amino acid oxidase/peroxidase assay [17,18]. Reaction mixtures contained 10 mM of the substrate (N_α,N_ε-diacetyl-L-Lys-D-Ala-D-Ala, or acetyl-L-Lys-D-Ala-D-Ala or D-Ala-D-Ala), 5 mM of the peroxidase colorimetric substrate 4-aminoantipyrine (4-AAP, from Sigma-Aldrich, Milan, Italy), 3 U/mL *Rg*DAAO (D-amino acid oxidase from *Rhodotorula gracilis* [18]), 7.5 U/mL horseradish peroxidase (HRP from Sigma-Aldrich, Milan, Italy), 6 mM phenol in 50 mM 1,3-bis[tris(hydroxymethyl)methylamino] propane (pH 7.4) in a final volume of 1 mL. At 25°C, 40 μg of recombinant $VanY_n$ or the amount of cell wall alkaline extract correspondent to 50 mg of cells, or the soluble fraction or the supernatant, was added to the reaction mixture and the increase in absorbance (ΔAbs/min) at 510 nm was measured for the test sample as well as for the control to which no $VanY_n$ was added. A molar extinction coefficient for chinonemine of 6.58/mM cm was used.

(ii) Fluorimetric *o*–phthaldialdehyde (OPTA) method [20]. Reaction mixtures contained 10 mM N_α,N_ε-diacetyl-L-Lys-D-Ala-D-Ala, 40 μg of recombinant $VanY_n$ in 50 mM phosphate buffer (pH 7.0) in a final volume of 200 μL. After 10 minutes at 25°C, the reaction was stopped by adding 50 μL of 250 mM HCl followed by 750 μL of water. Enzymatically released D-Ala was detected by the

addition of 100 μL of fluoraldehyde (OPTA) solution (Sigma-Aldrich, Milan, Italy) to 100 μL of the reaction mix, followed by incubation at room temperature for 5 min. 800 μL of water was added and 200 μL was removed: the fluorescence intensity was measured (λ_{ex} = 340 nm; λ_{em} = 455 nm) in a fluorescence microplate reader (Tecan Infinite® 200 Pro, Milan, Italy). Assays were quantified from a standard curve prepared with known amounts of D-Ala.

Circular dichroism measurements

Far-UV CD spectra were recorded with a Jasco J-715 (Jasco Europe, Cremella, Italy) spectropolarimeter in the 195–250 nm wavelength range. Measurements were made in quartz cuvettes of 1 mm pathlength, employing protein solutions of 0.1 mg/mL, and were corrected for buffer contribution. Secondary structure fractions were calculated from deconvolution of the CD spectra using the program K2D2 (http://www.ogic.ca/projects/k2d2/) [31].

Determination of the biological activity: the minimum inhibitory concentration (MIC)

Minimal inhibitory concentrations (MICs) of teicoplanin and vancomycin (Sigma-Aldrich, Milan, Italy) to *S. venezuelae* containing pIJ86 or pIJ86ΩC-His$_6$-*vanY$_n$* or pIJ86ΩN-His$_6$-*vanY$_n$* was determined in MYM agar added with 50 μg/mL apramycin and increasing concentrations of glycopeptides. The inoculum was 10^6 cfu/mL (after mycelium sonication with Sonics Vibracell VCX 130 – power 130 Watt, 230 Volt, 50–60 Hertz, frequency 20 Hz 5 minutes of sonication, 20 seconds for each cycles with 90% of amplitude), and plates were incubated at 28°C until colonies appeared. MIC was the lowest concentration of the antibiotic that inhibits the visible growth of the recombinant *S. venezuelae* strains [32].

Additional files

Additional file 1: Figure S1. SDS-PAGE Analysis of C-His$_6$-VanY$_n$ and N-His$_6$-VanY$_n$ from *S. venezuelae* recombinant strains as in Figure 2 main text. Analysis of samples corresponding to cellular and extracellular fractions from cultures of *S. venezuelae* recombinant strains grown in YEME for 72 hours. From *S. venezuelae* pIJ86ΩN-His$_6$-*vanY$_n$*: extracellular fraction (lane 1), soluble cell-free fraction (lane 2), insoluble cell-free fraction (lane 3) and cell wall fraction (lane 4); from *S. venezuelae* pIJ86ΩC-His$_6$-*vanY$_n$*: extracellular fraction (lane 5), soluble cell-free fraction (lane 6), insoluble cell-free fraction (lane 7) and cell wall fraction (lane 8). In each lane, samples corresponding to 100 μL of cell culture were loaded. Std reference protein: C-His$_6$-VanY$_n$ from *E. coli* (5 μg, 25 kDa).

Additional file 2: Figure S2. SDS-PAGE Analysis of C-His$_6$-VanY$_n$ from *S. venezuelae* recombinant strain growth in 3-L batch fermentor as in Figure 3 main text. Crude extracts of cell samples collected at different times of fermentation: 18 (lane 1), 24 (lane 2), 48 (lane 3), 72 (lane 4), 96 (lane 5), 120 (lane 6), 144 (lane 7), 168 (lane 8) hours. In each lane, samples corresponding to 100 μL of cell culture were loaded. Std reference protein: C-His$_6$-VanY$_n$ from *E. coli* (5 μg, 25 kDa).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
FM conceived the project and wrote the paper. EB performed most of the experiments on protein expression, purification and biochemical characterization, and prepared figures and tables. GLM developed conjugation and molecular biology tools. FB focused on protein purification procedures. LP designed the experiments on the protein biochemistry. All authors have read and approved the final manuscript.

Acknowledgements
We acknowledge EUFP7/2007-2011 Metaexplore grant agreement No: 222625 to GLM and FM for funding our research on novel heterologous hosts for protein expression. This work was also supported by grants from Fondo di Ateneo per la Ricerca to FM, Progetto Cariplo: Promuovere Capitale Umano d'Eccellenza to EB and GLM, and by MIUR fellowships to GLM and FB. We appreciate the support from Consorzio Interuniversitario per le Biotecnologie (CIB).

Received: 26 November 2012 Accepted: 10 March 2013
Published: 16 March 2013

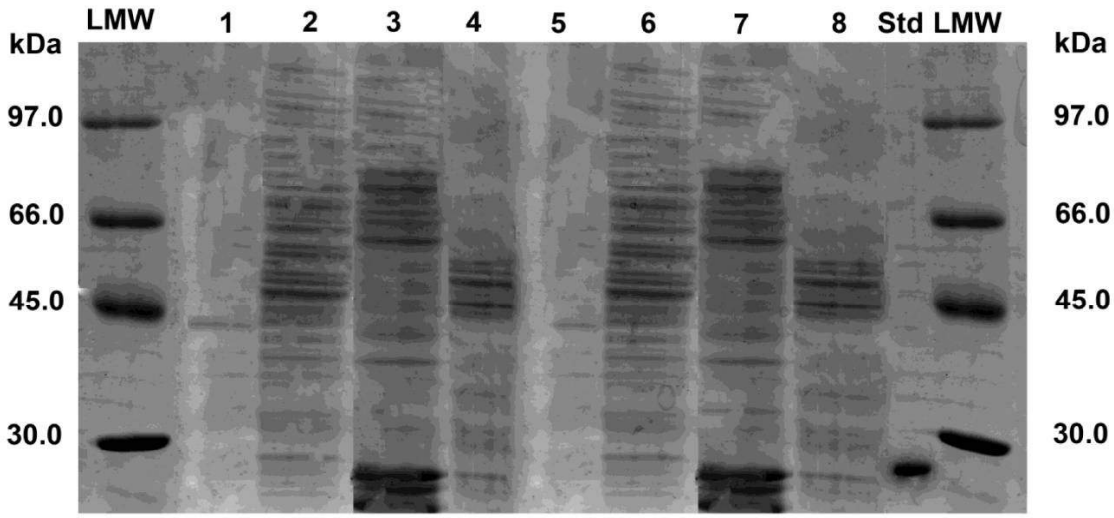
References

1. Marinelli F, Marcone G: **Small Molecules, Microbial Secondary Metabolites.** In *Comprehensive Biotechnology Second Edition volume 3.* Edited by Moo-Young M. Amsterdam, The Netherlands: Elsevier; 2011:285–297.
2. Nakashima N, Mitani Y, Tamura T: **Actinomycetes as host cells for production of recombinant proteins.** *Microb Cell Fact* 2005, **4**(1):7.
3. Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, *et al*: **Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2).** *Nature* 2002, **417**(6885):141–147.
4. Marinelli F: **Chapter 2. From microbial products to novel drugs that target a multitude of disease indications.** *Methods Enzymol* 2009, **458**:29–58.
5. Marcone GL, Foulston L, Binda E, Marinelli F, Bibb M, Beltrametti F: **Methods for the genetic manipulation of *Nonomuraea* sp. ATCC 39727.** *J Ind Microbiol Biotechnol* 2010, **37**(10):1097–1103.
6. Marcone GL, Carrano L, Marinelli F, Beltrametti F: **Protoplast preparation and reversion to the normal filamentous growth in antibiotic-producing uncommon actinomycetes.** *J Antibiot (Tokyo)* 2010, **63**(2):83–88.
7. Volontè F, Pollegioni L, Molla G, Frattini L, Marinelli F, Piubelli L: **Production of recombinant cholesterol oxidase containing covalently bound FAD in *Escherichia coli*.** *BMC Biotechnol* 2010, **10**:33.
8. Romano D, Molla G, Pollegioni L, Marinelli F: **Optimization of human D-amino acid oxidase expression in *Escherichia coli*.** *Protein Expr Purif* 2009, **68**(1):72–78.
9. Noda S, Ito Y, Shimizu N, Tanaka T, Ogino C, Kondo A: **Over-production of various secretory-form proteins in *Streptomyces lividans*.** *Protein Expr Purif* 2010, **73**(2):198–202.
10. Beltrametti F, Consolandi A, Carrano L, Bagatin F, Rossi R, Leoni L, Zennaro E, Selva E, Marinelli F: **Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by *van* gene homologue expression directing the synthesis of a modified cell wall peptidoglycan.** *Antimicrob Agents Chemother* 2007, **51**(4):1135–1141.
11. Taurino C, Frattini L, Marcone GL, Gastaldo L, Marinelli F: ***Actinoplanes teichomyceticus* ATCC 31121 as a cell factory for producing teicoplanin.** *Microb Cell Fact* 2011, **10**:82.
12. Marcone GL, Beltrametti F, Binda E, Carrano L, Foulston L, Hesketh A, Bibb M, Marinelli F: **Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727.** *Antimicrob Agents Chemother* 2010, **54**(6):2465–2472.
13. Van Bambeke F, Van Laethem Y, Courvalin P, Tulkens PM: **Glycopeptide antibiotics: from conventional molecules to new derivatives.** *Drugs* 2004, **64**(9):913–936.
14. Kieser T: *M. J. Bibb, M. J. Buttner, K. F.* Practical Streptomyces Genetics. The John Innes Foundation, Norwich: Chater, Hopwood DA; 2000.
15. Anné J, Maldonado B, Van Impe J, Van Mellaert L, Bernaerts K: **Recombinant protein production and streptomycetes.** *J Biotechnol* 2011, **158**(4):159–167.

16. Hutchings MI, Hong HJ, Buttner MJ: **The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*.** *Mol Microbiol* 2006, **59**(3):923–935.
17. Binda E, Marcone GL, Pollegioni L, Marinelli F: **Characterization of $VanY_n$, a novel D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727.** *FEBS J* 2012, **279**(17): 3203–3213.
18. Molla G, Vegezzi C, Pilone MS, Pollegioni L: **Overexpression in *Escherichia coli* of a recombinant chimeric *Rhodotorula gracilis* D-amino acid oxidase.** *Protein Expr Purif* 1998, **14**(2):289–294.
19. Granier B, Jamin M, Adam M, Galleni M, Lakaye B, Zorzi W, Grandchamps J, Wilkin JM, Fraipont C, Joris B: **Serine-type D-Ala-D-Ala peptidases and penicillin-binding proteins.** *Methods Enzymol* 1994, **244**:249–266.
20. Wright GD, Molinas C, Arthur M, Courvalin P, Walsh CT: **Characterization of *vanY*, a D,D-carboxypeptidase from vancomycin-resistant *Enterococcus faecium* BM4147.** *Antimicrob Agents Chemother* 1992, **36**(7):1514–1518.
21. Podmore AH, Reynolds PE: **Purification and characterization of VanXY(C), a D, D-dipeptidase/D, D-carboxypeptidase in vancomycin-resistant *Enterococcus gallinarum* BM4174.** *Eur J Biochem* 2002, **269**(11):2740–2746.
22. Wu Z, Wright GD, Walsh CT: **Overexpression, purification, and characterization of VanX, a D,D-dipeptidase which is essential for vancomycin resistance in *Enterococcus faecium BM4147*.** *Biochemistry* 1995, **34**(8):2455–2463.
23. Arthur M, Depardieu F, Cabanie L, Reynolds P, Courvalin P: **Requirement of the VanY and VanX D,D-peptidases for glycopeptide resistance in enterococci.** *Mol Microbiol* 1998, **30**(4):819–830.
24. Courvalin P: **Vancomycin resistance in gram-positive cocci.** *Clin Infect Dis* 2006, **42**(Suppl 1):S25–S34.
25. Arthur M, Depardieu F, Snaith HA, Reynolds PE, Courvalin P: **Contribution of VanY D,D-carboxypeptidase to glycopeptide resistance in *Enterococcus faecalis* by hydrolysis of peptidoglycan precursors.** *Antimicrob Agents Chemother* 1994, **38**(9):1899–1903.
26. Reynolds PE, Arias CA, Courvalin P: **Gene *vanXYC* encodes D,D-dipeptidase (VanX) and D,D-carboxypeptidase (VanY) activities in vancomycin-resistant *Enterococcus gallinarum* BM4174.** *Mol Microbiol* 1999, **34**(2):341–349.
27. Reynolds PE, Depardieu F, Dutka-Malen S, Arthur M, Courvalin P: **Glycopeptide resistance mediated by enterococcal transposon *Tn1546* requires production of VanX for hydrolysis of D-alanyl-D-alanine.** *Mol Microbiol* 1994, **13**(6):1065–1070.
28. McCafferty DG, Lessard IA, Walsh CT: **Mutational analysis of potential zinc-binding residues in the active site of the enterococcal D-Ala-D-Ala dipeptidase VanX.** *Biochemistry* 1997, **36**(34):10498–10505.
29. Molla G, Bernasconi M, Sacchi S, Pilone MS, Pollegioni L: **Expression in *Escherichia coli* and *in vitro* refolding of the human protein pLG72.** *Protein Expr Purif* 2006, **46**(1):150–155.
30. Kariyama R, Massidda O, Daneo-Moore L, Shockman GD: **Properties of cell wall-associated D,D-carboxypeptidase of *Enterococcus hirae* (*Streptococcus faecium*) ATCC 9790 extracted with alkali.** *J Bacteriol* 1990, **172**(7):3718–3724.
31. Perez-Iratxeta C, Andrade-Navarro MA: **K2D2: estimation of protein secondary structure from circular dichroism spectra.** *BMC Struct Biol* 2008, **8**:25.
32. Andrews JM: **Determination of minimum inhibitory concentrations.** *J Antimicrob Chemother* 2001, **48**(Suppl 1):5–16.

doi:10.1186/1472-6750-13-24
Cite this article as: Binda *et al.*: *Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance. *BMC Biotechnology* 2013 **13**:24.

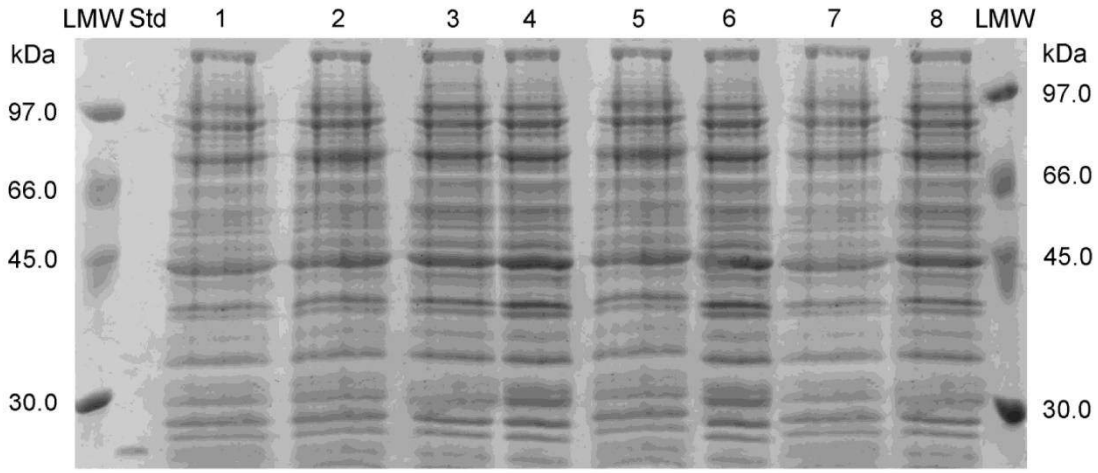
Additional file 1



Additional Figure 1

SDS-PAGE Analysis of C-His$_6$-VanY$_n$ and N-His$_6$-VanY$_n$ from *S. venezuelae* recombinant strains as in Fig. 2 main text. Analysis of samples corresponding to cellular and extracellular fractions from cultures of *S. venezuelae* recombinant strains grown in YEME for 72 hours. From *S. venezuelae* pIJ86ΩN-His$_6$-*vanY$_n$*: extracellular fraction (lane 1), soluble cell-free fraction (lane 2), insoluble cell-free fraction (lane 3) and cell wall fraction (lane 4); from *S. venezuelae* pIJ86ΩC-His$_6$-*vanY$_n$*: extracellular fraction (lane 5), soluble cell-free fraction (lane 6), insoluble cell-free fraction (lane 7) and cell wall fraction (lane 8). In each lane, samples corresponding to 100 µL of cell culture were loaded. Std reference protein: His$_6$-VanY$_n$ from *E. coli* (5 µg, 25 kDa).

Additional file 2



Additional Figure 2

SDS-PAGE Analysis of C-His_6-$VanY_n$ from *S. venezuelae* recombinant strain growth in 3-L batch fermentor as in Fig. 3 main text. Crude extracts of cell samples collected at different times of fermentation: 18 (lane 1), 24 (lane 2), 48 (lane 3), 72 (lane 4), 96 (lane 5), 120 (lane 6), 144 (lane 7), 168 (lane 8) hours. In each lane, samples corresponding to 100 µL of cell culture were loaded. Std reference protein: His_6-$VanY_n$ from *E. coli* (5 µg, 25 kDa).

Chapter 4

Streptomyces lividans as host for the heterologous expression of a metagenome-sourced chitobiosidase

Francesca Berini[1], Ilaria Presti[1§], Giorgia Letizia Marcone[1], Loredano Pollegioni[1], Flavia Marinelli[1*]

[1]Department of Biotechnology and Life Sciences, University of Insubria, Varese Italy and "The Protein Factory" Research Center, Politecnico di Milano, ICRM CNR Milano and University of Insubria, Varese Italy

*Corresponding author, flavia.marinelli@uninsubria.it

[§]Present address: Industriale Chimica, Saranno, Varese Italy

Abstract

Metagenomics is one of the most promising technologies for the discovery and characterisation of novel biocatalysts to be exploited at industrial level. Thanks to the availability of extended genetic toolkits, *Escherichia coli* has been employed as the preferential cloning host in most of the metagenomic studies performed thus far. However, since several drawbacks could be associated with the use of this microorganism as protein factory, there is the need to identify and develop alternative expression platforms with different codon usages and higher protein secretion capacity. In this work, the possibility to employ *Streptomyces lividans* as host for the heterologous expression of metagenome-sourced chitinases is evaluated. Different approaches to repress its endogenous chitinolytic system, which could interfere with heterologous chitinase production and detection, are investigated. Moreover, the results obtained from the cloning and expression of Chi18H8, a chitobiosidase previously isolated from a suppressive-soil metagenomic library, are reported.

Keywords: *Streptomyces*, chitinase, chitobiosidase, metagenome, heterologous expression

Introduction

Metagenomics is an innovative and powerful tool for the discovery of novel biocatalysts encrypted in natural microbial communities, with high potential for use in biotechnological or environmental applications. Screening of metagenomic libraries has been focused on the identification of several classes of enzymes, including lipolytic enzymes, oxidoreductases and dehydrogenases or polysaccharide degrading enzymes (amylases, cellulases, chitinases, xylanases) [1, 2]. However, the complete characterisation and subsequent application at industrial level of these candidate proteins has often been limited by problems associated with their over-expression. Indeed, protein production in microbial hosts frequently represents the real bottleneck for the exploitation of metagenome-sourced biocatalysts. The Gram-negative bacterium *Escherichia coli* is certainly the most employed expression platform, thanks to its unparalleled fast growth kinetics also in inexpensive media and the availability of many molecular tools and manipulation protocols. However, it has been estimated that only 30-40% of bacterial genes could be efficiently expressed in *E. coli*, a value dropping to 7% for high G+C DNA [3]. This might be due to a plethora of factors, such as codon usage differences, improper promoter recognition, citotoxicity, incorrect folding, inclusion bodies formation and inability to secrete the translated proteins [4, 5]. Hence, there is the need to develop alternative cloning hosts, with different codon usages and higher protein secretion capacity than *E. coli*. *Streptomyces* spp., Gram-positive high G+C content bacteria, belonging to the Actinomycetales and naturally living in the soil, are promising candidates [6, 7]. Their innate secretion capacity reduces the risk of local accumulation of the over-expressed recombinant proteins, helps the correct folding and simplifies the purification procedures; moreover, it limits contamination with host proteins, nucleic acids and endotoxins. The ability to grow on defined media, the presence of natural mechanisms of genetic exchange and the increasing availability of expression vectors support the employment of these filamentous bacteria as protein factories. Among streptomycetes, *S. lividans* is one of the preferred hosts thanks to the limited restriction-modification system and the low endogenous protease activity [8]. Novel enzymes,

identified in soil and marine meta-genomic screens, have already been successfully produced to high levels in *S. lividans* [9-11]. The main limitation in using *Streptomyces* spp. as systems for the heterologous expression of metagenome-derived chitinases is the presence of an endogenous chitinolytic system, which can interfere with the production and detection of the recombinant enzyme. As typical degraders of biomasses in soil habitats, in fact, streptomycetes have multiple and diverse chitinase genes with different specificities and characteristics, which are thought to act synergistically in chitin degradation [12, 13]. *S. coelicolor* A3(2)'s genome sequencing revealed the presence of thirteen different genes coding for chitinases, which are not organised in clusters but are randomly scattered on the chromosome [14, 15]. Similarly, multiple chitinolytic enzymes have been identified also in other streptomycetes, including *S. lividans* and *S. olivaceoviridis* [16]. Since the knock-out of all these genes is not feasible, alternative approaches for the repression of the endogenous chitinolytic systems should be followed. In this work we report on the investigation of *S. lividans* TK24 as alternative cloning host for the production of the chitobiosidase Chi18H8, previously isolated from a metagenomic library of a suppressive soil for clubroot disease of cabbage [17].

Material and methods

Strains and cultivation media

Streptomyces lividans TK24 was kindly donated by M.J. Bibb, John Innes Centre, Norwich UK. The strain was maintained as spores in 10% (v/v) glycerol and propagated in soy flour mannitol (SFM) agar medium [18]. *Streptomyces lividans ΔdasR* was a gift from S. Rigali, University of Liège, Liège Belgium. Ex-conjugants of *S. lividans* TK24 containing pIJ86, pIJ86::C-His_6-*chi18H8* and pIJ86::N-His_6-*chi18H8*, and *S. lividans ΔdasR* were grown at 28 °C on SFM agar plates supplemented with 50 µg/mL apramycin. Colonies were picked up from agar plates and inoculated into 100 mL Erlenmeyer flasks with 20 mL liquid medium and 50 µg/mL apramycin. Flask cultures were incubated on a rotary shaker at 200 revolutions per minute (rpm) and 28 °C. Liquid media for streptomycetes were: yeast extract – malt extract medium (YEME), containing in (w/v) 0.3% yeast extract, 0.5% bacto-peptone, 0.3% malt extract, 1.0% glucose in deionised water, pH 7.0; bacto tryptic soy broth (BTSB), containing

in (w/v) 10% sucrose, 1.0% glucose, 1.0% yeast extract, 0.5% NaCl, 0.5% soybean meal, 1.7% tryptone and 0.25% K_2HPO_4 in deionised water, pH 7.2; medium V (MV), containing in (w/v) 2.4% soluble starch, 0.1% glucose, 0.3% meat extract, 0.5% yeast extract, 0.5% tryptose in deionised water, pH 7.2; maltose yeast extract medium (MYM), containing in (w/v) 0.4% maltose, 0.4% yeast extract, 1.0% malt extract and 2 mL/L trace element solution (TES), in deionised water. The TES was composed of 40 mg/mL $ZnCl_2$, 200 mg/mL $FeCl_2$ x 6 H_2O, 10 mg/mL $CuCl_2$ x 2 H_2O, 10 mg/mL $MnCl_2$ x 4 H_2O, 10 mg/mL $Na_2B_4O_7$ x 10 H_2O, 10 mg/mL $(NH_4)_6Mo_7O_{24}$ x 4 H_2O, in deionised water. For the detection of chitinolytic activity, *S. lividans ΔdasR* was grown on chitin agar medium (CHA) plates, composed of (in w/v) 0.4% colloidal chitin, 0.07% K_2HPO_4, 0.03% KH_2PO_4, 0.05% $MgSO_4$, 0.001% $FeSO_4$, 0.0001% $ZnSO_4$, 0.0001% $MnCl_2$, 2.0% agar. Colloidal chitin was prepared by the method of Hsu and Lockwood, 1975 [19], starting from chitin flakes from shrimp shells. *Escherichia coli* DH5α (Invitrogen-Life Technology, Carlsbad, USA) was used as host for cloning procedures. *E. coli* ET12567/pUZ8002 [20] was employed as non-methylating plasmid donor strain for intergeneric conjugation with *S. lividans* TK24. *E. coli* strains were maintained on Luria-Bertani (LB) agar plates. All reagents were from Sigma-Aldrich, St Louis USA, unless otherwise stated.

Plasmid construction

The multi-copy expression vector pIJ86 employed for heterologous expression in *S. lividans* TK24 was a gift of M.J. Bibb [7, 20]. Plasmids pIJ86::C-His_6-*chi18H8* and pIJ86::N-His_6-*chi18H8* were constructed as follows. Expand High Fidelity polymerase (Roche Italia, Milan Italy) was employed to amplify *chi18H8* using fosmid DNA as template [17]. In both constructs, *Hind*III and *Bgl*II restriction sites were introduced into the primer sequence (underlined) for insertion into the multiple cloning site of the pIJ86 vector. A C-terminal tag of six histidine residues was introduced by amplification with the primers *chiHisC_Fw* (5'-ATAAAAAGCTTGATGCGCCAGCTCACGCTTCTCCTTGCGACCGCTGC-3') and *chiHisC_Rev* (5'-ATAAAAGATCTTCAGTGGTGGTGGTGGTGGTGATTGCCCCTTGCAGACTGG-3'). The histidine tag was added at the N-terminus of *chi18H8* with the primers *chiHisN_Fw* (5'-ATAAAAAGCTTAATGCACCACCACCACCACCACCGCCAGCTCACGCTTC-3') and *chiHisN_Rev* (5'-ATAAAAGATCTTCAATTGCCCCTATGCAGACTGGCGGTGATCGCTCGC-3'). PCR was performed for 30 cycles as follows: 95 °C for 45 seconds, 74 °C for 45 seconds and 72 °C for 90 seconds.

The PCR products were purified, digested with *Hind*III and *Bgl*II and ligated with T4-DNA ligase (Roche Italia, Milan Italy) into the double-digested pIJ86. The resulting plasmids, pIJ86::C-His_6-*chi18H8* and pIJ86::N-His_6-*chi18H8*, were transformed into *E. coli* DH5α and controlled by DNA sequencing (BMR Genomics, Padua Italy). The plasmids were finally cloned into *E. coli* ET12567/pUZ8002 cells.

Intergeneric conjugation

Intergeneric conjugation was performed according to the modified protocol described in Binda *et al.*, 2013 [7]. The correct transformation of the recombinant *S. lividans* strains, carrying pIJ86::C-His_6-*chi18H8,* pIJ86::N-His_6-*chi18H8* or the empty pIJ86 plasmid, was checked by colony PCR. Single colonies were transferred onto difco nutrient agar (DNA) medium [18] and allowed to grow at 28 °C over night (O.N.). Mycelium was scraped from the plates using a sterile tip and re-suspended in 50 µL 100% (v/v) dimethyl sulfoxide (DMSO). The samples were vigorously shaken for 90-120 minutes at room temperature and then centrifuged at 14000 x *g* for 3 minutes. 2.5 µL of the supernatant were used for PCR. For control, 1 µL of DNA sample was mixed with 1.5 µL 100% (v/v) DMSO. PCR was performed for 26 cycles as follows: 94 °C for 45 seconds, 55 °C for 45 seconds and 72 °C for 30 seconds. One initial step of 10 minutes at 94 °C was included in the program to ensure a complete cell lysis. *chiHisC_Fw, chiHisC_Rev, chiHisN_Fw* and *chiHisN_Rev* were employed as oligonucleotide primers to check the presence of pIJ86::C-His_6-*chi18H8* and pIJ86::N-His_6-*chi18H8*, respectively. The presence of the empty pIJ86 plasmid was verified with the primers *pIJ86_Fw* (5'-TGCACGCGGTCGATCTTGAC-3') and *pIJ86_Rev* (5'-TCATGGTCGGTCTCCTGGTG-3'), annealing to regions of the vector around the multiple cloning site.

***S. lividans* strains growth**

Wild type and recombinant *S. lividans* TK24 strains, as well as *S. lividans ΔdasR*, previously re-activated in 100 mL Erlenmeyer flasks for 72 hours, were inoculated at 10% (v/v) in baffled 500 mL Erlenmeyer flasks with 100 mL liquid medium supplemented, if necessary, with 50 µg/mL apramycin. Flasks were incubated at 28 °C and 200 rpm for different time intervals (up to 240 hours) and regularly sampled. For the construction of growth curves, 5 mL were centrifuged at 1900 x *g* for 10 minutes at room temperature. pH and residual

glucose were measured with pH Test Strips 4.5-10.0 (Sigma-Aldrich, St Louis USA) and Diastix strips (Bayer, Leverkusen Germany), respectively. Biomass production was calculated as wet weight. For chitinase activity measurements and SDS-PAGE analysis, 10 mL were instead centrifuged at 1900 x *g* for 20 minutes at 4 °C. Part of the supernatants, *i.e.* the cell-free culture broth, was concentrated with the trichloracetic acid (TCA) method. Briefly, 1/10 (v/v) of 100% (w/v) TCA was added to 5 mL sample, vortexed for 15 seconds and placed on ice for 15 minutes. After centrifuging at 14000 x *g* for 10 minutes, the supernatant was removed and discarded. The pellet was washed twice with 500 µL of pure acetone and then air dried for about 90 minutes. Finally, the pellet was re-suspended in 100 µL phosphate buffer saline (PBS) pH 7.3 (140 mM NaCl, 2.7 mM KCl, 10.1 mM Na_2HPO_4, 1.8 mM KH_2PO_4) and 100 µL of SDS-sample buffer 4x (200 mM Tris-HCl pH 6.8, 33.3% (v/v) glycerol, 8.4% (v/v) β-mercaptoethanol, 6.66% (w/v) sodium dodecyl sulphate (SDS), 0.01% (w/v) bromophenol blue), heat-shocked for 3 minutes at 85 °C.

Celle pellets were sonicated on ice (10-15 cycles of 30 seconds each, with a 30-second interval, using a Branson Sonifier 250, Dansbury USA) in PBS 1x containing 10 µg/mL deoxyribonuclease (DNase), 0.19 mg/mL phenylmethylsulfonylfluoride (PMSF) and 0.7 mg/mL pepstatin. Soluble and insoluble fractions were then separated by centrifugation at 20000 x *g* for 1 h at 4 °C. Insoluble fractions were re-suspended in a volume of PBS equal to the corresponding cytoplasmic soluble fraction (5 mL/g_{cells}) for successive analyses. Protein concentration was determined by the Biuret assay [21].

Chitinase activity measurement

Extracellular and intracellular chitinase activity was assayed with the fluorimetric chitooligosaccharide analogues 4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide (4-MU-GlcNAc), 4-methylumbelliferyl *N,N'*-diacetyl-β-D-chitobioside (4-MU-$(GlcNAc)_2$) and 4-methylumbelliferyl *N,N',N''*-triacetyl-β-D-chitotrioside (4-MU-$(GlcNAc)_3$) as substrates as previously reported [17]. One unit (U) of chitinase activity was defined as the amount of enzyme required for the release of 1 µmole of 4-MU per minute at 37 °C [17, 22]

SDS-PAGE and zymogram analysis

To study Chi18H8 expression, protein samples of recombinant *S. lividans* TK24 strains were separated by sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS-PAGE) with 12% (w/v) polyacrylamide gels, using a Tris-Glicine system and Comassie brilliant blue R-250 staining as described in [23]. Reference standard protein was His_6-glycine oxidase (His_6-GO) from *Bacillus subtilis* gently provided by L. Pollegioni, University of Insubria, Varese Italy [24]. The molecular weight markers were from GE-Healthcare Sciences, Little-Chalfont UK. Chitinolytic activity was detected through zymogram analysis using 10% (w/v) separating polyacrylamide gels containing 0.7 mg/mL carboxymethyl-chitin-Remazol brilliant violet (CM-chitin-RBV, Loewe Biochemica, Sauerlach Germany) as reported in [17]. Reference standard protein was *Trichoderma viride* chitinase (Sigma-Aldrich, St Louis USA).

Chi18H8 purification

For Chi18H8 purification, *S. lividans* TK24/pIJ86::C-His_6-*chi18H8* was grown for 96-120 hours in YEME medium, supplemented with 2.0% (w/v) glucose and 50 µg/mL apramycin. The cell-free culture broth was precipitated by slowly adding a sufficient amount of ammonium sulphate to reach 70% (w/v) saturation, then incubated for 3 hours at 4 °C and centrifuged (12000 x *g* at 4 °C for 30 minutes). The pellet was re-suspended in 1/10 (v/v) 20 mM potassium phosphate buffer (KPi) pH 6.7. The recombinant protein was purified by loading onto a 5-mL Ni^{2+}-Hitrap chelating affinity column (1.6 x 2.5 cm; GE Healthcare Sciences, Little Chalfont UK) equilibrated with 20 mM KPi pH 6.7, 30 mM NaCl and 5 mM imidazole, according to manufacturer's instructions. After extensive washing, the protein was eluted with an increasing concentration of elution buffer (20 mM KPi pH 6.7, 300 mM NaCl and 250 mM imidazole).

Results

Production and repression of endogenous chitinolytic activity

S. lividans TK24 chitinase production was at first assayed in two rich media, the limpid YEME and the complex BTSB, both commonly employed for the growth of streptomycetes and which do not contain chitin. According to the original recipe, the media were supplemented

with 1.0% (w/v) glucose. As shown in Figure 1, panels A&B, the maximum biomass production was reached after 48 hours-incubation and corresponded to a wet weight of 94 g/L and 139 g/L for YEME and BTSB, respectively. In the first medium, the glucose was totally consumed within 48 hours and the pH values were almost constant around 7.0 for the entire growth (Figure 1, panel A). In BTSB glucose depletion was completed in 24 hours and the pH constantly grew from neutral to slightly basic values (Figure 1, panel B). Maximum extracellular chitinase activity in BTSB (about 0.52 U/g_{cells}, corresponding to 60 U/L) was measured after 72 hours of growth, whereas the production in YEME reached 1.3 U/g_{cells} (32 U/L) after 96 hours. In both media, chitinase activity was detected when the cells enter into the stationary phase of growth, having completely consumed the available glucose, suggesting that its production is under the control of catabolite repression, as observed also in other streptomycetes [16, 25]. No significant chitinolytic activity was detected in the intracellular fractions obtained after cell sonication (data not shown). *S. lividans* TK24's extracellular chitinases, produced in rich media when glucose is depleted and also in the absence of chitin as inducer, can therefore interfere with the expression and purification of heterologous chitinases.

Chitinolytic activity production was then investigated in *S. lividans ΔdasR* that is a recombinant strain deleted in the regulatory gene *dasR*, gently provided by S. Rigali from Liège University. DasR is a known transcriptional activator in *S. coelicolor* A3(2), where it recognises a 12-bp direct repeat sequence, called *dre* (DasR-responsive element), present in the promoter region of chitinase genes [26]. In this streptomycete, *dasR* knock out drastically reduces the expression of chitinolytic enzymes [15, 27, 28]. Since *S. lividans* is closely related to *S .coelicolor* A3(2) [29], it is conceivable that *dasR* knocking out in the former might as well reduce its endogenous chitinolytic activity. Indeed, replacing *dasR* by an apramcyin resistance cassette resulted in a non-sporulating bald phenotype on SFM agar (Figure 2, panel A), as demonstrated by Rigali and co-workers also for *S. coelicolor ΔdasR* [26]. However, differently from what observed in *S. coelicolor* knock-out mutant, *S. lividans ΔdasR* was able to degrade chitin when grown on chitin agar medium (CHA) for 10 days of incubation at 28 °C (Figure 2, panel B). As in the case of *S. coelicolor* A3(2) and *S. lividans* wild types, the chitinolytic activity is repressed if CHA is supplemented with 1% (w/v) glucose or 1% (w/v) *N*-acetylglucosamine (Figure 2, panel B).

Production of chitinolytic activity in *S. lividans ΔdasR* was assayed also in liquid culture, employing YEME and BTSB media (Figure 1, panels C&D). A biphasic growth was observed in YEME with a maximum of biomass production of 55 g/L after 48 hours from the inoculum (same timing as in the wild type, Figure 1, panel A), but with a second later peak corresponding to 192 g/L after 240 hours. Differently from the wild type, glucose was not consumed before 72 hours of growth, and then was slowly depleted within 240 hours, indicating that the first growth phase is supported by other carbon sources than glucose. The pH remained constant at 5.0 until 168 hours, then increasing to 9.0 in the very last phase of fermentation (Figure 1, panel C). Maximum biomass production in BTSB was higher than in the wild type (255 g/L), but the glucose consumption profile and the pH trend were similar between the recombinant strain and its parental one (Figure 1, panel D). In YEME, extracellular chitinase activity was detected only after 168 hours, when glucose was almost completely depleted, thus confirming that it may act as repressor. The chitinase production was delayed but it was twice more than in the wild type, *i.e.* 2.9 U/g_{cells} corresponding to 95 U/L volumetric productivity (Figure 1, panel C). In BTSB (Figure 1, panel D), production of the extracellular chitinase activity showed a similar trend than in the wild type, but even in this case it was almost two-fold what previously observed. In both media, only traces of intracellular chitinase activity were recorded, as in the wild type strain (data not shown). Taken together, the above experiments demonstrate that, despite the deletion of the *dasR* gene, *S. lividans ΔdasR* produces more endogenous chitinolytic activity than the wild type and hence it is not a suitable alternative host for heterologous chitinase expression.

To verify if homologous chitinase production could be effectively repressed by increasing glucose concentration, *S. lividans* TK24 was grown as above in YEME and BTSB media, but increasing the sugar concentration from 1.0% to 2.0% (w/v). In YEME plus glucose, biomass production was reduced but the stationary phase of growth was prolonged (with a maximum wet weight of 58 g/L between 24 and 72 hours), glucose depletion lasted 96 hours and the pH dropped below 5.0 after 48 hours from inoculum (Figure 1, panel E). Since streptomycetes prefer neutral-to-alkaline pH (the optimum growth being registered between 6.5 and 8.0 [30]), the acid environment could explain why higher glucose concentrations did not support a parallel increase in biomass production. For BTSB plus glucose (Figure 1, panel F), the free glucose was rapidly consumed within 48 hours and both the growth curve and the chitinolytic activity profile were similar to the ones reported in

Figure 1, panel B. Differently from BTSB plus glucose, the slower sugar consumption occurring in YEME plus glucose allowed appreciating the repressive effect exerted by the monosaccharide on chitinase activity, which never exceeded 0.17 U/g_{cells} (Figure 1, panel E). Glucose concentrations higher than 2.0% (w/v) did not improve the carbon catabolite repression effect (data not shown). Hence, it can be concluded that growth in YEME medium supplemented with 2.0% (w/v) glucose represents the preferential cultivation condition to be used for the expression of heterologous chitinases.

Chi18H8 heterologous expression

In order to clone *chi18H8* in *S. lividans* TK24, the gene was amplified by PCR and introduced into the multicopy plasmid pIJ86, under the control of the strong and constitutive promoter *ermE** [31]. A His_6-Tag sequence was introduced either at the N- or C-terminus of the protein, to facilitate its purification and to verify the influence of the Tag position on Chi18H8 activity. pIJ86::C-His_6-*chi18H8*, pIJ86::N-His_6-*chi18H8*, and the empty pIJ86 plasmid used as control, were introduced into *S. lividans* TK24 by intergeneric conjugation from an *E. coli* non-methylating plasmid donor strain. The correct transformation was confirmed by colony PCR and DNA sequencing.

The three ex-conjugants were thus grown in YEME plus glucose and BTSB plus glucose and their growth and chitinase production profiles compared with the wild type. The growth curves of the control recombinant strain, transformed with the empty multi-copy vector, were similar to the ones of the wild type (data not shown). Interestingly, they were almost overlapping also with those of the recombinant strain carrying pIJ86::N-His_6-*chi18H8* that are shown in Figure 1, panels G&H. Insertion of pIJ86 vectors did not affect maximum biomass productivity, but reduced growth speed, probably due to the metabolic burden associated with the heterologous plasmid maintenance. In BTSB plus glucose, maximum biomass production (around 150 g/L) and glucose depletion were reached after 72 hours of incubation (Figure 1, panels H&J). In YEME plus glucose, the sugar was not consumed for the first 72-96 hours of incubation, being after that depleted within 120 or 240 hours in the recombinant strains carrying pIJ86::N-His_6-*chi18H8* (Figure 1, panel G) or pIJ86::C-His_6-*chi18H8* (Figure 1, panel I), respectively. Maximum biomass production in the former (135 g/L) was measured after 120 hours, whereas in the latter constant values (around 95 g/L) were recorded during the entire stationary phase, from 72 to 168 hours of incubation.

In BTSB plus glucose, chitinase activity never exceeded 0.95 U/g_{cells}, neither in the recombinant microorganisms carrying pIJ86::*chi18H8* plasmids, nor in the control strain. The maximum of chitinase activity was detected when glucose was depleted. Substantial differences were visible when growing the microorganisms in YEME plus glucose (Figure 1, panels G&I). In *S. lividans* TK24/pIJ86::C-His_6-*chi18H8*, in fact, chitinase activity levels were significantly higher than in the other recombinant strains, reaching 2.4 U/g_{cells} and 200 U/L between 96 and 216 hours of incubation in the presence of significant level of residual glucose.

Chi18H8 production by *S. lividans* TK24/pIJ86::C-His_6-*chi18H8* was assayed also in other media frequently employed for heterologous protein expression in *Streptomyces* spp.: the rich but limpid MV and the semi-defined MYM. All culture broths were supplemented with 2.0% (w/v) glucose. However, the chitinase activity measured in these additional trials was overall lower than in YEME plus glucose, which was therefore selected as the best medium for recombinant Chi18H8 production.

Chi18H8 purification and characterisation

For Chi18H8 purification, *S. lividans* TK24/pIJ86::C-His_6-*chi18H8* was grown in YEME plus glucose and harvested after 96-120 hours from the inoculum. The cell-free culture broth was concentrated with ammonium sulphate precipitation and loaded onto a HiTrap chelating column. As shown in Figure 3, panels A&B, chitinase activities were recorded both in the first chromatographic fractions (corresponding to not-bounded proteins), both in fraction 8, corresponding to the elution pick. Nevertheless, the zymogram analysis (Figure 3, panel C) clearly demonstrated the efficacy of affinity chromatography in separating the His_6-tagged Chi18H8 from the other chitinolytic enzymes. In fact, a protein band migrating at the same height of the recombinant Chi18H8 previously purified from *E. coli* [Berini *et al*., manuscript in preparation] was detectable only in the crude extract loaded onto the column and in the elution pick fraction. Instead, the first chromatographic fractions showed only bands at different heights, corresponding to the same proteins seen in the negative control (*i.e. S. lividans* TK24 carrying the empty pIJ86 vector) and hence probably due to endogenous chitinolytic enzymes, which did not bind to the column. The successful purification procedure allowed the recovery from *S. lividans* broth of 216 μg of protein per litre of culture, corresponding to 16.9 μg/g_{cells}.

Similarly to the results obtained for Chi18H8 in *E. coli* [17], activity assays on fluorimetric chitooligosaccharides analogues demonstrated that the purified enzyme had a prevalent chitobiosidase activity, a weaker endochitinase activity (*ca.* 20% of the value recorded for the chitobiosidase one) and no β-*N*-acetylglucosaminidase activity.

Discussion

Chi18H8 coding sequence was identified in a fosmid metagenomic library of a suppressive soil and initially cloned in *E. coli*, in the expression vectors pGEX-6P-3 [17] and pET24b(+) [Berini *et al.*, manuscript preparation]. The biochemical and functional characterisation of this acidophilic chitobiosidase revealed interesting features, like antifungal activity, long-term stability and solvent-tolerance, which make this enzyme a promising candidate for biotechnological and environmental applications. However, high-yield purification from these heterologous systems was hampered by its accumulation in a mainly inactive form in the inclusion bodies. Hence, in the present work *chi18H8* cloning and expression in an alternative host was evaluated, with the aim to simplify its production and recovery.

The chosen expression platform was *S. lividans* TK24, a Gram-positive filamentous bacterium well-known for its proven excellence in secretion capacity and low extracellular protease activity, recently employed by our group for the successful expression of another protein [7]. Additionally, according to SignalP 4.1 server [32], Chi18H8 possess a 55-aminoacid long N-terminal signal peptide sequence, which has the probability to be correctly recognised by Gram positive hosts.

As member of *Streptomyces* genus, also *S. lividans* TK24 possess its own chitinolytic system, which may interfere with the activity of the recombinant chitinase. For this reason, before proceeding with *chi18H8* expression, different approaches for the repression of this complex system were evaluated. It is known that in these bacteria chitinase production is induced by chitin and repressed by readily utilizable carbon sources, such as glucose or *N*-acetylglucosamine, the monomer subunit of chitin [16, 25, 33]. We have hereby demonstrated that *S. lividans* TK24 chitinolytic system is actively expressed also in non-inducing media (*i.e.* not supplemented with chitin), but it can be efficiently repressed by glucose at 1.0-2.0% (w/v) concentrations, thus providing a cleaner background for the

successive heterologous expression of Chi18H8. The repression effect was more evident in YEME, where the glucose was more slowly consumed than in BTSB; this is a significant advantage, since heterologous protein purification from limpid media such as YEME is usually less difficult than from flour-rich media as BTSB.

A second evaluated approach was based on the selective deletion of *dasR*, a pleiotropic factor involved in chitinase gene regulation [26]. DasR is a known transcriptional activator for chitinase genes in *S. coelicolor* A3(2) [15, 28], whereas it acts as repressor for the genes of the sugar phosphotransferase system PTS for the uptake of *N*-acetylglucosamine [28, 34], and for *accII-4* and *redZ* genes involved in actinorhodin and undecylprodigiosin production, respectively [26-28]. DasR is also essential for the development of *S. coelicolor* and *S. griseus*, where it controls the *dasABC* transporter operon related to glucose-dependent morphogenesis [35]. The disruption of *S. lividans' dasR* by substitution of the gene sequence with apramycin resistance cassette resulted in a non-sporulating bald phenotype on SFM agar, as previously demonstrated for *S. coelicolor ΔdasR* [26]. However, the mutant strain still showed considerable chitinolytic activity both in liquid and solid culture, which, similarly to the wild type strain, was inhibited by glucose and by *N*-acetylglucosamine. The transcriptional regulation of chitinase production in streptomycetes involves multiple regulation systems [15]: beside DasR, the two-component system ChiS/ChiR initially identified for the regulation of *chiC* in *S. coelicolor* A3(2) [36], the Cpb1 DNA-binding protein for *chiA* [37] and Reg1 [38] both of *S. lividans*. Moreover, real-time PCR analysis conducted by Nazari and co-workers in 2011 [15] showed that, after *dasR* deletion from *S. coelicolor* M145, the induction levels of some chitinase genes were only partially reduced and the expression of *chiA* and *chiF* was even higher in the mutant than in the wild type strain. Hence, it is possible that also in *S. lividans* several and multilevel regulatory systems are involved in chitinase gene induction and that, for this reason, the disruption of the sole *dasR* is not sufficient to completely repress the endogenous chitinolytic activity. We can therefore conclude that the use of *S. lividans ΔdasR* strain, at least in the evaluated conditions, does not represent a substantial improvement for overcoming the problem of repressing endogenous chitinolytic activity when using this host for the heterologous expression of chitinase genes.

In the last part of this work we reported on *chi18H8* cloning in *S. lividans* TK24. The metagenome-sourced chitinase was successfully expressed in this host, with the best results

obtained in YEME medium. Interestingly, the position of the His_6-Tag at the N-terminus of the protein seemed to interfere with the recombinant protein production and/or activity: *S. lividans* clone possessing the His_6-Tag in fusion with the C-terminus of the protein reached higher chitinase activity levels than the other recombinant strain carrying the pIJ86::N-His_6-*chi18H8* vector. In addition, Chi18H8 production in *S. lividans* TK24/pIJ86::C-His_6-*chi18H8* recombinant strain occurred in the presence of glucose, consistently with the cloning of the heterologous gene under the strong constitutive promoter *ermE**, that is insensitive to glucose repression. Chi18H8 was produced only in the extracellular fractions, thus confirming that its signal peptide was correctly recognised by the host secretion system. We suppose that most of the chitinolytic activity detected in the *S. lividans* TK24/pIJ86::N-His_6-*chi18H8* strain was due to endogenous chitinase production. It is possible that the fusion of the His_6-Tag at the N-terminal of the heterologous protein hampers its correct processing and secretion, as it has been already demonstrated for the D,D-peptidase/D,D-carboxypeptidase $VanY_n$ [7].

Protein secretion into the culture broth has generally three major advantages over intracellular accumulation: secreted proteins are usually natively folded, they can be produced at similar or even higher levels than intracellular proteins and they can be more easily purified [8]. Indeed, recombinant Chi18H8 secretion simplified its purification procedure: culture broth precipitation with ammonium sulphate and one-step chromatography were sufficient to obtain a good purity protein. However, the purification yield from *S. lividans* (16.9 $\mu g/g_{cells}$), even if comparable to the one obtained with Chi18H8 purification from *E. coli* soluble fraction (21 $\mu g/g_{cells}$) [17], is significantly lower to what achieved with enzyme solubilisation from inclusion bodies (Berini *et al.,* manuscript in preparation). The limited production in streptomycetes could be due to different factors, among which problems with protein maturation and processing before or during its secretion, or a non-optimal codon usage. Chi18H8 G+C content, in fact, is 64.47%, which is similar but not identical to the genomic ratio of *Streptomyces* spp. (72.1% for *S. coelicolor* A3(2) [14]). Nonetheless, the protein secretion into culture medium and the easy purification procedure, as well as the possibility to significantly repress the endogenous chitinolytic system by simply adding glucose to the medium, make *S. lividans* TK24 an interesting and valuable candidate for the heterologous expression of metagenome-sourced chitinases, worthy of further exploration.

Acknowledgments

This research activity was funded by the EU FP7/2007-2011 Metaexplore project, grant agreement No. 222625 to Flavia Marinelli. Miur fellowship and CIB contribution to Francesca Berini are also acknowledged.

References

1. Lorenz P, Eck J: **Metagenomics and industrial applications**. *Nat Rev Microbiol* 2005, **3**(6):510-516.
2. Felczykowska A, Bloch SK, Nejman-Faleńczyk B, Barańska S: **Metagenomic approach in the investigation of new bioactive compounds in the marine environment**. *Acta Biochim Pol* 2012, **59**(4):501-505.
3. Gabor EM, Alkema WB, Janssen DB: **Quantifying the accessibility of the metagenome by random expression cloning techniques**. *Environ Microbiol* 2004, **6**(9):879-886.
4. Rosano GL, Ceccarelli EA: **Recombinant protein expression in *Escherichia coli*: advances and challenges**. *Front Microbiol* 2014, **5**:172.
5. Berlec A, Strukelj B: **Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells**. *J Ind Microbiol Biotechnol* 2013, **40**(3-4):257-274.
6. Vrancken K, Van Mellaert L, Anné J: **Cloning and expression vectors for a Gram-positive host, *Streptomyces lividans***. *Methods Mol Biol* 2010, **668**:97-107.
7. Binda E, Marcone GL, Berini F, Pollegioni L, Marinelli F: ***Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance**. *BMC Biotechnol* 2013, **13**:24.
8. Anné J, Maldonado B, Van Impe J, Van Mellaert L, Bernaerts K: **Recombinant protein production and streptomycetes**. *J Biotechnol* 2012, **158**(4):159-167.
9. Meilleur C, Hupé JF, Juteau P, Shareck F: **Isolation and characterization of a new alkali-thermostable lipase cloned from a metagenomic library**. *J Ind Microbiol Biotechnol* 2009, **36**(6):853-861.
10. Côté A, Shareck F: **Expression and characterization of a novel heterologous moderately thermostable lipase derived from metagenomics in *Streptomyces lividans***. *J Ind Microbiol Biotechnol* 2010, **37**(9):883-891.
11. Sianidis G, Pozidis C, Becker F, Vrancken K, Sjoeholm C, Karamanou S, Takamiya-Wik M, van Mellaert L, Schaefer T, Anné J *et al*: **Functional large-scale production of a novel *Jonesia* sp. xyloglucanase by heterologous secretion from *Streptomyces lividans***. *J Biotechnol* 2006, **121**(4):498-507.
12. Saito A, Fujii T, Yoneyama T, Redenbach M, Ohno T, Watanabe T, Miyashita K: **High-multiplicity of chitinase genes in *Streptomyces coelicolor* A3(2)**. *Biosci Biotechnol Biochem* 1999, **63**(4):710-718.
13. Saito A, Shinya T, Miyamoto K, Yokoyama T, Kaku H, Minami E, Shibuya N, Tsujibo H, Nagata Y, Ando A *et al*: **The dasABC gene cluster, adjacent to dasR, encodes a novel ABC transporter for the uptake of N,N'-diacetylchitobiose in *Streptomyces coelicolor* A3(2)**. *Appl Environ Microbiol* 2007, **73**(9):3000-3008.
14. Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D *et al*: **Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)**. *Nature* 2002, **417**(6885):141-147.

15. Nazari B, Saito A, Kobayashi M, Miyashita K, Wang Y, Fujii T: **High expression levels of chitinase genes in *Streptomyces coelicolor* A3(2) grown in soil**. *FEMS Microbiol Ecol* 2011, **77**(3):623-635.
16. Saito A, Fujii T, Miyashita K: **Chitinase System in *Streptomyces***. *Actinomycetol* 1990, **13**:1-10.
17. Hjort K, Presti I, Elväng A, Marinelli F, Sjöling S: **Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics**. *Appl Microbiol Biotechnol* 2014, **98**(6):2819-2828.
18. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA: **Practical Streptomyces Genetics**: The John Innes Foundation, Norwich UK; 2000.
19. Hsu SC, Lockwood JL: **Powdered chitin agar as a selective medium for enumeration of actinomycetes in water and soil**. *Appl Microbiol* 1975, **29**(3):422-426.
20. Marcone GL, Beltrametti F, Binda E, Carrano L, Foulston L, Hesketh A, Bibb M, Marinelli F: **Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727**. *Antimicrob Agents Chemother* 2010, **54**(6):2465-2472.
21. Gornall AG, Bardawill CJ, David MM: **Determination of serum proteins by means of the biuret reaction**. *J Biol Chem* 1949, **177**(2):751-766.
22. McCreath KJ, Gooday GW: **A rapid and sensitive microassay for determination of chitinolytic activity** *J Microbiol Meth* 1992, **14**:229-227.
23. Schägger H, von Jagow G: **Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa**. *Anal Biochem* 1987, **166**(2):368-379.
24. Job V, Marcone GL, Pilone MS, Pollegioni L: **Glycine oxidase from *Bacillus subtilis*. Characterization of a new flavoprotein**. *J Biol Chem* 2002, **277**(9):6985-6993.
25. Miyashita K, Fujii T, Saito A: **Induction and repression of a *Streptomyces lividans* chitinase gene promoter in response to various carbon sources**. *Biosci Biotechnol Biochem* 2000, **64**(1):39-43.
26. Rigali S, Nothaft H, Noens EE, Schlicht M, Colson S, Müller M, Joris B, Koerten HK, Hopwood DA, Titgemeyer F *et al*: **The sugar phosphotransferase system of *Streptomyces coelicolor* is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development**. *Mol Microbiol* 2006, **61**(5):1237-1251.
27. Nazari B, Kobayashi M, Saito A, Hassaninasab A, Miyashita K, Fujii T: **Chitin-induced gene expression in secondary metabolic pathways of *Streptomyces coelicolor* A3(2) grown in soil**. *Appl Environ Microbiol* 2013, **79**(2):707-713.
28. Rigali S, Titgemeyer F, Barends S, Mulder S, Thomae AW, Hopwood DA, van Wezel GP: **Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by Streptomyces**. *EMBO Rep* 2008, **9**(7):670-675.
29. Leblond P, Redenbach M, Cullum J: **Physical map of the *Streptomyces lividans* 66 genome and comparison with that of the related strain *Streptomyces coelicolor* A3(2)**. *J Bacteriol* 1993, **175**(11):3422-3429.
30. Kontro M, Lignell U, Hirvonen MR, Nevalainen A: **pH effects on 10 *Streptomyces* spp. growth and sporulation depend on nutrients**. *Lett Appl Microbiol* 2005, **41**(1):32-38.
31. Bibb MJ, Janssen GR, Ward JM: **Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of *Streptomyces erythraeus***. *Gene* 1985, **38**(1-3):215-226.
32. Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: discriminating signal peptides from transmembrane regions**. *Nat Methods* 2011, **8**(10):785-786.
33. Colson S, van Wezel GP, Craig M, Noens EE, Nothaft H, Mommaas AM, Titgemeyer F, Joris B, Rigali S: **The chitobiose-binding protein, DasA, acts as a link between chitin utilization and morphogenesis in *Streptomyces coelicolor***. *Microbiology* 2008, **154**(Pt 2):373-382.
34. Rigali S, Schlicht M, Hoskisson P, Nothaft H, Merzbacher M, Joris B, Titgemeyer F: **Extending the classification of bacterial transcription factors beyond the helix-turn-helix motif as an**

alternative approach to discover new cis/trans relationships. *Nucleic Acids Res* 2004, **32**(11):3418-3426.

35. Seo JW, Ohnishi Y, Hirata A, Horinouchi S: **ATP-binding cassette transport system involved in regulation of morphological differentiation in response to glucose in *Streptomyces griseus***. *J Bacteriol* 2002, **184**(1):91-103.
36. Homerová D, Knirschová R, Kormanec J: **Response regulator ChiR regulates expression of chitinase gene, chiC, in *Streptomyces coelicolor***. *Folia Microbiol (Praha)* 2002, **47**(5):499-505.
37. Fujii T, Miyashita K, Ohtomo R, Saito A: **DNA-binding protein involved in the regulation of chitinase production in *Streptomyces lividans***. *Biosci Biotechnol Biochem* 2005, **69**(4):790-799.
38. Nguyen J, Francou F, Virolle MJ, Guérineau M: **Amylase and chitinase genes in *Streptomyces lividans* are regulated by reg1, a pleiotropic regulatory gene**. *J Bacteriol* 1997, **179**(20):6383-6390.

Figure legends

Figure 1. Growth curves and chitinase activity profiles. Panels A and B: *S. lividans* TK24 in YEME (A) or BTSB (B) media. Panels C and D: *S. lividans ΔdasR* in YEME (C) or BTSB (D). Panels E and F: *S. lividans* TK24 in YEME plus glucose (E) or BTSB plus glucose (F). Panels G and H: *S. lividans* TK24/pIJ86::N-His_6-*chi18H8* in YEME plus glucose (G) or BTSB plus glucose (H). Panels I and J: *S. lividans* TK24/pIJ86::C-His_6-*chi18H8* in YEME plus glucose (I) or BTSB plus glucose (J). Growth parameters are: wet weight (■, dashed line), glucose consumption (●, solid line) and pH (▲, dotted line). Chitinase activity, measured on 4-MU-$(GlcNAc)_2$ as substrate, is represented as grey bars and expressed as units per g of cells in wet weight.

Figure 2. Panel A: phenotype of *S. lividans* TK24 (left) and *S. lividans ΔdasR* (right), grown on SFM agar. Panel B: growth of *S. lividans ΔdasR* on CHA and CHA supplemented with 1.0% (w/v) glucose or 1.0% (w/v) *N*-acetylglucosamine (GlcNAc).

Figure 3. Chi18H8 purification from *S. lividans* TK26/pIJ86::C-His_6-*chi18H8* culture broth. The recombinant strain was grown in YEME plus glucose medium for 120 hours. Starting material: 150 mL culture broth, corresponding to 1.6 g_{cells} (wet weight). Panel A: elution profile of Chi18H8 by nickel-affinity chromatography on HiTrap Chelating column. Panel B: chitinolytic activity recorded by fluorimetric assay on 4-MU-$(GlcNAc)_2$ as substrate. Panel C: zymogram analysis of: Ctrl-, crude extract from *S. lividans* TK26/pIJ86, used as negative control; C.E., crude extract from *S. lividans* TK26/pIJ86::C-His_6-*chi18H8*, loaded onto the column; 1 and 2, flow-through chromatographic fractions (see panel A); 8, chromatographic fraction eluted at 50 mM imidazole (see panel A); Std, standard commercial chitinase from *T. viride* (10 µg); Chi18H8, recombinant protein purified from *E. coli* BL21 Star™(DE3)/pET24b(+)::*chi18H8* (5 µg) (Berini *et al.*, manuscript in preparation).

Figures

Figure 1

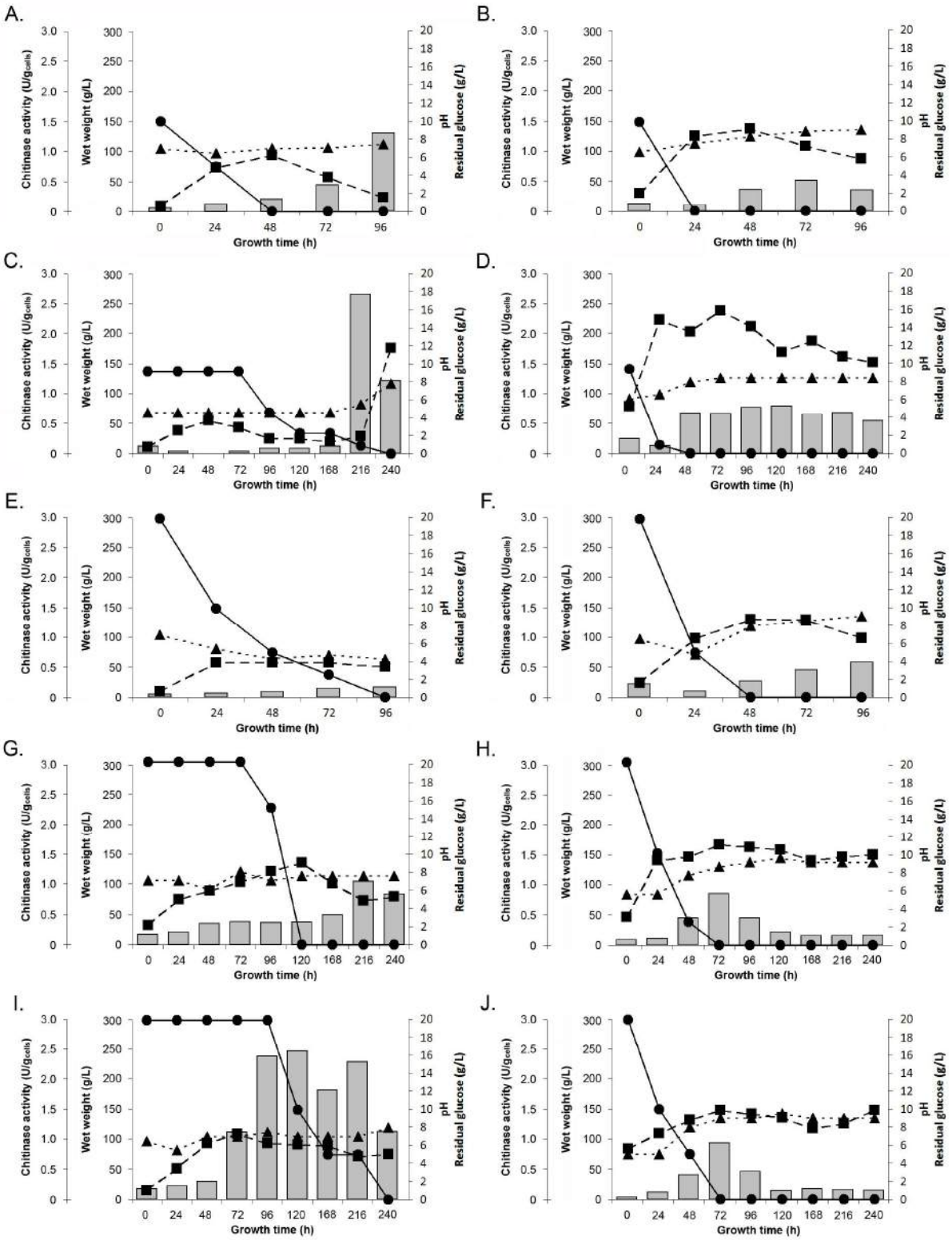


Figure 2

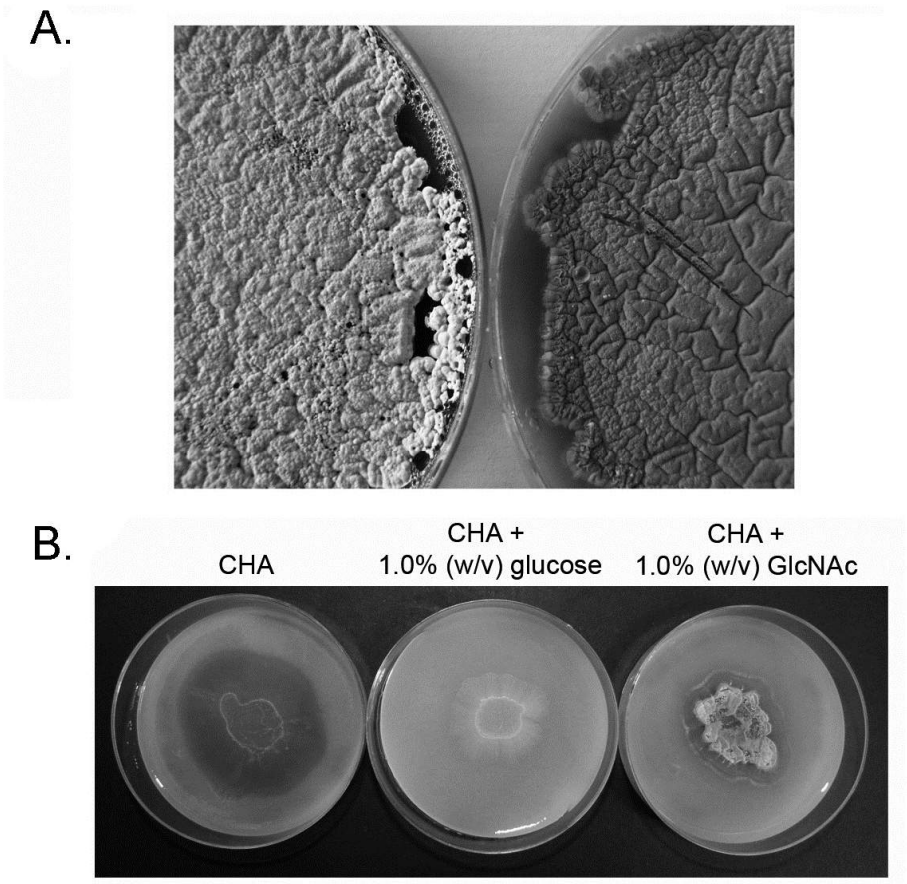
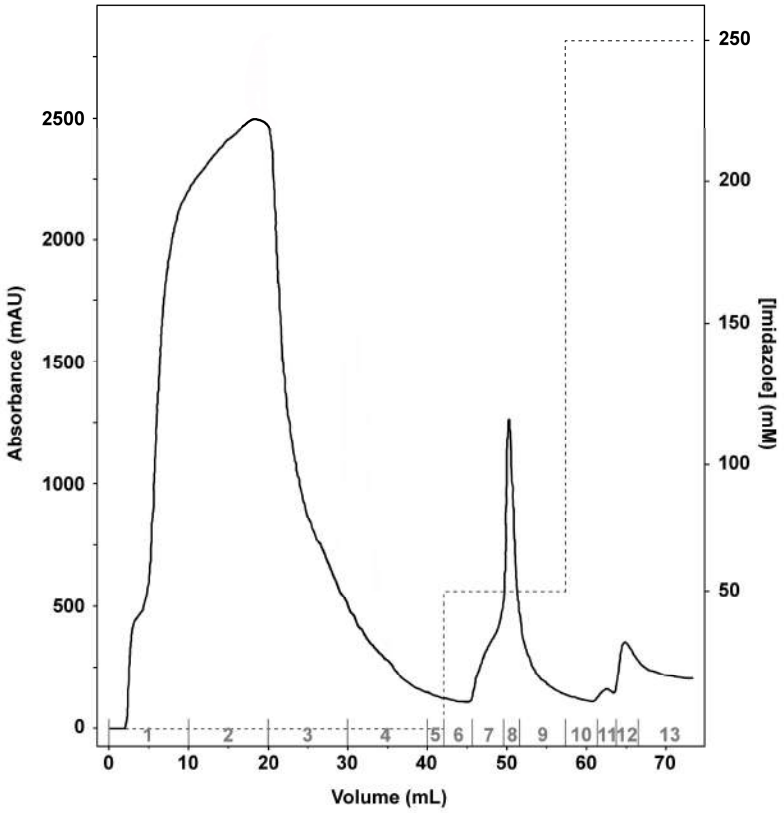


Figure 3

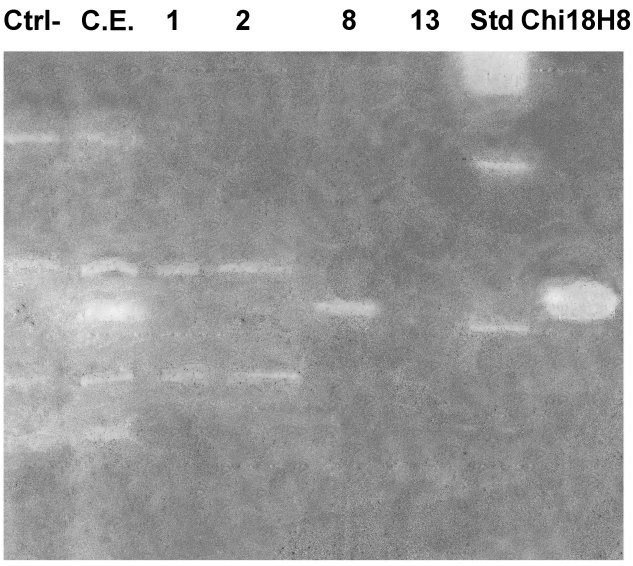
A.



B.

Fraction	1	2	3	4	5	6	7	8	9	10	11	12	13
Activity (total U)	0.85	3.01	0.67	1.01	0.00	0.01	0.01	1.32	0.72	0.45	0.18	0.25	0.00

C.



SECTION 3

“Chitinolytic enzymes as biocontrol agents”

Chapter 5

Effects of *Trichoderma viride* chitinases on the peritrophic matrix of the silkworm, *Bombyx mori*

Francesca Berini[1], Silvia Caccia[2,3], Morena Casartelli[2], Terenzio Congiu[4], Eleonora Franzetti[1], Flavia Marinelli[1], Gianluca Tettamanti[1]

[1]Department of Biotechnology and Life Sciences, University of Insubria, and "The Protein Factory" Research Center, Politecnico di Milano, ICRM CNR Milano and University of Insubria, 21100 Varese, Italy

[2]Department of Biosciences, University of Milan, 20133 Milan, Italy

[3]Department of Agriculture, University of Naples "Federico II", 80055 Portici (Naples), Italy

[4]Department of Surgical and Morphological Sciences, University of Insubria, 21100 Varese, Italy

Running title: Chitinases and silkworm peritrophic membrane

Corresponding author: Gianluca Tettamanti, gianluca.tettamanti@uninsubria.it

Abstract

The peritrophic matrix is a chitin and glycoprotein layer that lines the insect midgut. It is a physical barrier that mainly protects the midgut epithelium from food abrasions and pathogen infections, but it is also involved in the compartmentalisation of digestive enzymes and in the selective transport of nutrients from the lumen to the epithelial cells. Given its fundamental role in insect physiology, it has been considered an excellent target for the development of innovative strategies for pest control. In particular, bacterial, insect and viral chitinases, enzymes able to alter the structural and functional properties of this acellular sheath, have been explored as a tool for the development of integrated pest management strategies. On the contrary, fungal chitinases have never been tested in such approaches. For

this reason, in the present paper we performed a biochemical characterisation of a commercial cocktail of chitinolytic enzymes from *Trichoderma viride* and analysed its effects on the peritrophic matrix of the silkworm *Bombyx mori*, a representative model system among Lepidoptera. The encouraging results obtained on this lepidopteron make us confident on the use of a similar chitinase-based approach on other pest species that represent a serious damage for crops, forestry and pasture.

Keywords: fungal chitinases, insect midgut, Lepidoptera, peritrophic matrix

Introduction

The peritrophic matrix (PM) is a thin acellular sheath that lines the midgut epithelium of most insects and envelops the midgut lumen content (Lehane, 1997; Terra, 2001; Hegedus *et al.*, 2009). The PM consists of a network of chitin fibrils associated with different classes of proteins, glycoproteins and proteoglycans, which confer strength and elasticity to this structure and influence its permeability properties (Lehane, 1997; Tellam *et al*., 1999; Terra, 2001; Wang and Granados, 2001; Hegedus *et al*., 2009). The sieving attributes of the PM are related to the size and the charge of the effective aqueous channels that cross this gel-like structure and discriminate the passage of molecules. For instance, the PM of *Bombyx mori* larvae is largely permeable to methylene blue (320 Da), and almost impermeable to PEG 4000, while the trypsin modulating oostatic factor from *Aedes aegypti* (*Aea*-TMOF) has an intermediate permeability coefficient, in line with its molecular mass (1005 Da) (Fiandra *et al*., 2009). Thanks to its function as permeable barrier, the PM plays several roles in insect digestion, such as the compartmentalisation of digestive enzymes in the endo- and ecto-peritrophic spaces, and the selective transport of nutrients from the lumen to the epithelial cells (Terra, 2001). Moreover, it prevents mechanical lesions of the apical cell membrane caused by food abrasion and non-specific binding to cell surface; it acts as a defence physical barrier against ingested pathogens and toxins, furthermore protecting the midgut from allelochemicals and reactive oxygen species by sequestering and detoxifying ingested toxic materials (Tellam *et al.*, 1999; Barbehenn and Stannard, 2004).

Chitin plays an important role in PM structure and function, being the scaffold for additional components (*i.e.* proteins and glycans), and acting as a structural support for digestive and detoxifying enzymes. Hence, its disruption alters the structural and functional properties of this peculiar viscous layer (Wang and Granados, 2001; Fiandra *et al*., 2010). Chitin, also occurring in the exoskeleton of insects, as well as in the digestive systems of nematodes and in the cell wall of fungi, is an insoluble linear homopolymer of *N*-acetylglucosamine (GlcNAc) linked by *β*-(1-4)-glycosidic bonds, which is hydrolysed by chitinases (Duo-Chuan, 2006). Considerable interest in the chitinolytic enzymes has been stimulated by their possible involvement in enzyme-based integrated pest control strategies against insects, nematodes and fungi (Fiandra *et al*., 2010). Resistance to these pest agents can be imparted by degradation of their vital structures such as PM and cuticle in insects or the cell wall in fungal phytopathogens, or by liberation of substances that subsequently elicit other types of defence responses by the host (Boller, 1987).

In insects, chitinases are usually associated with postembryonic development and turnover of the cuticle (Merzendorfer and Zimoch, 2003; Kramer and Koga, 1986). Insect growth, morphogenesis and metamorphosis strictly depend on structure changes in tissues and organs containing chitin, such as the epidermis, tracheae and PM (Zhuo *et al*., 2014), where chitin levels are maintained by a dynamic and strictly regulated balance of synthesis and decomposition, due to the action of chitin synthases and chitinases, respectively. Interestingly, chitinases and chitin deacetylases associated to the insect midgut are believed to create temporary localised pores in the PM to increase nutrient or enzyme passage (Hegedus *et al*., 2009).

Besides arthropods, chitinases have been characterised in a wide range of organisms, including fungi, viruses, bacteria, higher plants and other animals (Adrangi and Faramarzi, 2013). Some entomopathogenic fungi such as *Metarhizium anisopliae*, *Beauveria bassiana*, *Nomuraea rileyi,* produce a cocktail of chitinolytic and proteolytic enzymes that attacks the insect cuticles, hence facilitating pathogen penetration and infection (El-Sayed *et al*., 1989; St. Leger *et al*., 1991). Soil fungi belonging to common genera such as *Aspergillus*, *Penicillium* and *Trichoderma,* represent a rich source of chitinolytic enzymes degrading a wide range of different chitinous substrates (Hartl *et al*., 2012). Recent studies showed that chitinases from *Trichoderma* spp., specifically of *T. harzianum*, are active and effective against a wide range of phytopathogen fungi, thus becoming particularly attractive for biocontrol strategies

(Lorito *et al*., 2010; Monteiro *et al*., 2010; Mukherjee, *et al*., 2013). In these microbes, enzymatic hydrolysis of chitin is accomplished by the synergistic and consecutive action of two major categories of chitinases (Dahiya *et al*., 2006). Endochitinases cleave chitin randomly at internal sites, generating soluble, low molecular mass multimers of GlcNAc, such as chitotetraose, chitotriose and diacetylchitobiose. Exochitinases include chitobiosidases, which catalyse the progressive release of diacetylchitobiose starting at the non-reducing end of chitin myofibril, and β-(1-4) *N*-acetyl glucosaminidases, which cleave the oligomeric products of endochitinases and chitobiosidases, thus generating monomers of GlcNAc in an exo-type fashion (Duo-Chuan, 2006).

Feeding the insect larvae with viral (Corrado *et al*., 2008; Fiandra *et al*., 2010), bacterial (Regev *et al*., 1996) and plant (Ding *et al*., 2008) chitinases has been tested as an integrated pest management (IPM) strategy. Chitin network rupture may cause a significant damage to the PM, leading to an increased vulnerability of midgut epithelium to pathogens and pathogen-released toxins, and remarkably influencing the physiology of the digestive tract and insect survival (Hegedus *et al*., 2009; Pardo-Loperz *et al*., 2009; Fang *et al*., 2009; Sun *et al*., 2012). To our knowledge, fungal chitinases, although being used as fungal phytopathogen biocontrol agents, have never been tested in such approaches. In the present paper we propose a model study in which (i) we describe the different enzymatic activities present in a commercially available preparation of chitinolytic enzymes from the fungus *Trichoderma viride;* and (ii) we test its action *in vitro* on *Bombyx mori* larvae PM integrity and function. *Bombyx mori* is in fact considered a good model system among Lepidoptera and the information achieved on this insect can be then transferred to other butterfly and moth species. The aim of this study is gaining information on the potential use of fungal chitinases for developing novel IPM strategies.

Materials and methods

Experimental animals

B. mori (polyhybrid strain (126x57)(70x90)) larvae were provided by CRA - Honey Bee and Silkworm Research Unit (Padua, Italy). The larvae were fed an artificial diet (Cappellozza *et al*., 2005) and reared at 25 ± 0.5 °C under a 12L:12D photoperiod and 70% relative humidity.

After animals had ecdysed to the last larval stage (5th instar), they were staged and synchronised according to Franzetti *et al*. (Franzetti *et al*., 2012).

SDS-PAGE and zymogram analysis

The composition of the mixture of chitinolytic enzymes from *Trichoderma viride* (Sigma-Aldrich, St Louis, USA) was analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) with 12% (w/v) polyacrylamide gels, using a Tris-glicine system and Comassie brilliant blue R-250 staining (Schägger and von Jagow, 1987). The molecular weight markers were from GE-Healthcare Sciences, Little Chalfont, UK. Zymogram was employed to detect chitinolytic activities, using 10% (w/v) separating polyacrylamide gels containing 0.7 mg/mL carboxymethyl-chitin-remazol brilliant violet (CM-chitin-RBV) (Loewe Biochemica, Sauerlach, Germany) as previously described (Hjort *et al*., 2014).

Chitinase activity assay

Chitinase activity was quantified with the fluorimetric chitooligosaccharide analogues 4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide (4-MU-GlcNAc), 4-methylumbelliferyl *N,N'*-diacetyl-β-D-chitobioside (4-MU-$(GlcNAc)_2$) and 4-methylumbelliferyl *N,N',N''*-triacetyl-β-D-chitotrioside (4-MU-$(GlcNAc)_3$) (Sigma-Aldrich, St Louis, USA) as substrates, as described in (Hjort *et al*., 2014). One unit (U) of chitinase activity was defined as the amount of enzyme required for the release of 1 µmole of 4-MU per min at 37 °C (Hjort *et al*., 2014; McCreath and Gooday, 1992).

Chitinolytic activity was determined also by the colorimetric method described by Anton and Barrett (Anton and Barrett, 2002) adapted to enzymatic hydrolysis, with colloidal chitin as substrate. Colloidal chitin was prepared starting from chitin flakes from shrimp shells (Sigma-Aldrich, St Louis, USA) according to Hsu and Lockwood (Hsu and Lockwood, 1975) and its pH corrected to 5.0 and 7.0 with 0.1 M NaOH. 250 µL of protein sample were added to an equal volume of 10 g/L colloidal chitin, and the mixture was incubated at 37 °C for 1 h. The reaction was quenched by boiling for 5 min and then centrifuged (20000 x *g*, 25 °C, 15 min); 200 µL of the supernatant were mixed with equal volumes of 0.5 M NaOH and of 3-methyl-2-benzothiazolinonehydrazone (MBTH) reagent. After a 15-min incubation at 80 °C, 400 µL of a solution containing 0.5% (w/v) $FeNH_4(SO_4)_2x12\ H_2O$, 0.5% (w/v) sulfamic acid and 0.25 M HCl were added and allowed to cool to room temperature. After final addition of 1 mL of H_2O,

absorbance at 620 nm was determined. Released reducing sugars were estimated by comparison to a standard curve prepared varying GlcNAc concentrations (0 – 600 µM). One U of chitinase activity was defined as the amount of enzyme that released 1 µmol/mL x h GlcNAc at 37 °C.

pH and temperature curves

The optimum pH for the chitinase activity was determined by the enzymatic assay on the three chitooligosaccharide analogues 4-MU-GlcNAc, 4-MU-$(GlcNAc)_2$ and 4-MU-$(GlcNAc)_3$. The chitinase lyophilised powder and the substrates were diluted in the following buffers (100 mM) at corresponding pH: glycine-HCl (pH 3.0), sodium acetate (pH 4.0 and 5.0), sodium phosphate (pH 6.0 and 7.0), TrisHCl (pH 8.0), and sodium pyrophosphate (NaPPi, pH 9.0). The optimum temperature was assayed by incubating the reaction mixture (in 100 mM sodium acetate buffer pH 5.0, according to the standard protocol) at various temperatures (5 – 70 °C). The relative activity was expressed as a percentage of the highest activity recorded. Long-term stability of the enzyme was tested by the standard fluorimetric assay after pre-incubating the chitinase at pH 7.0 and at different temperatures (4 and 25 °C) from 0 to 192 h.

Structural and ultrastructural analysis of the peritrophic membrane

Isolation of B. mori *peritrophic matrix and incubation with* T. viride *chitinases*

Larvae at the second day of the 5^{th} instar were quickly anesthetised with CO_2 before dissection. They were cut dorsally and the midgut immediately isolated and severed lengthwise to expose the PM with the enclosed intestinal content. The PM was carefully extracted and cut longitudinally. The lumen content was removed by repeatedly rinsing the PM with PBS (phosphate-buffered saline, 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.0). Each PM was cut into four pieces, two of which were used as controls (for SEM and TEM analysis) and the other two exposed to *T. viride* chitinases (for SEM and TEM). Samples were transferred to a 24-multiwell plate and incubated for 90 min at room temperature in the absence or in presence of increasing concentrations of the *T. viride* chitinase mixture: 0.5 mg/mL, corresponding to 25 and 5 total U for the enzyme diluted at pH 5.0 and 7.0, respectively; and 1 mg/mL, corresponding to 50 and 10 total U for the chitinase rinsed at pH 5.0 and 7.0, respectively (total U were calculated as sum of the

activities recorded on the three fluorigenic substrates 4-MU-GlcNAc, 4-MU-$(GlcNAc)_2$ and 4-MU-$(GlcNAc)_3$). The samples were then fixed *in situ* and further processed for SEM and TEM analyses.

Scanning electron microscopy (SEM)

To obtain three-dimensional imaging by SEM, PMs were fixed with 4% (w/v) glutaraldehyde in 0.1 M Na-cacodylate buffer (pH 7.4) for 90 min at room temperature. After washes in Na-cacodylate buffer, specimens were post-fixed in a solution of 1% (w/v) osmium tetroxide, 1.25% (w/v) potassium ferrocyanide for 1 h. After dehydration in an increasing series of ethanol and a step in hexamethyldisilazane (2 × 5 min), specimens were mounted on carbonated stubs, gold coated with a Sputter K250 coater, and then observed with a SEM-FEG XL-30 microscope (Philips, Eindhoven, The Netherlands).

Light microscopy (LM) and transmission electron microscopy (TEM)

For LM and TEM analysis, PMs were fixed with 4% (w/v) glutaraldehyde in 0.1 M Na-cacodylate buffer (pH 7.4) overnight at 4 °C. Specimens were postfixed in 1% (w/v) osmium tetroxide for 1 h, dehydrated in an ethanol series, and embedded in an Epon/Araldite 812 mixture. Semithin sections were stained with crystal violet and basic fuchsin and observed by using an Olympus BH2 microscope (Olympus, Tokyo, Japan). Images were acquired with a DS-5M-L1 digital camera system (Nikon, Tokyo, Japan). Thin sections were stained using uranyl acetate in methanol (Milloning, 1976) and lead citrate and observed by using a Jeol JEM-1010 electron microscope (Jeol, Tokyo, Japan). Images were acquired by an Olympus Morada digital camera (Olympus, Münster, Germany).

Permeability measurements

PMs from *B. mori* larvae at the second day of 5^{th} instar were isolated as above described. The PM was cut into two halves, one used as control and the other exposed to *T. viride* chitinases. For these experiments, PM was laid on a thin cotton gauze. The gauze was necessary to maintain the PM extended, thus avoiding its fluttering when mounted in the experimental apparatus; additionally, it had a large mesh in order to avoid any restriction to the permeation of molecules. The gauze with the dissected PM was then mounted as a flat sheet between the two cylindrical plexiglass emichambers of the perfusion apparatus (Ussing chamber, World Precision Instruments, Berlin, Germany). Each opposing emichamber had a round matching hole with a surface area of 12.6 mm^2. Before placing the

PM, the margins of the holes were carefully spread with silicone paste (Baysilone Paste, Sigma–Aldrich, St Louis, USA) to avoid lesions of the PM and fluid leakage. The ectoperitrophic side layered onto the gauze, which, when interposed between the two emichambers, separated the ectoperitrophic compartment from the endoperitrophic one. Both compartments were filled with 500 µL of PBS and the flux of the methylene blue dye (0.5 mg/mL) from the endoperitrophic compartment to the ectoperitrophic one was measured in the absence (control) or in the presence of different concentrations of chitinase: 0.5 mg/mL or 0.8 mg/mL, corresponding to 5 and 8.4 total U per emichamber, respectively (total U were calculated as sum of the activities recorded on the three fluorigenic substrates 4-MU-GlcNAc, 4-MU-$(GlcNAc)_2$ and 4-MU-$(GlcNAc)_3$). After 90 min of incubation at room temperature the solution in the ectoperitrophic compartment was recovered and the amount of methylene blue was determined spectrophotometrically (Ultrospec 3000 Pharmacia Biotech, Cambridge, UK) at the wavelength of 664 nm. A calibration curve was carried out with known amounts of the dye dissolved in PBS. Flux values were expressed as µg/cm^2/h, and mean values ± s.e. were compared by Student's *t*-test.

Results

Enzymatic characterisation of the chitinase cocktail from *Trichoderma viride*

As indicated by the manufacturer, the commercial chitinase preparation from *T. viride* is a mixture of chitinolytic enzymes purified from the culture broth of the fungus (Rogalski *et al*., 1997). The SDS-PAGE analysis revealed the presence of at least four different proteins, with molecular mass ranging from 30 to 80 kDa (Figure 1a). The presence of multiple chitin-degrading enzymes was confirmed also by the zymogram analysis conducted on CM-chitin-RBV as substrate: four different degradation bands were visible in the gel (Figure 1b).

The substrate specificity of the diverse enzyme components in the mixture was assayed on three different length analogues of natural chitooligosaccharides and on colloidal chitin (Table 1). The assays were initially performed at pH 5.0 and 7.0. The chitinase mixture was active at both pHs on all the tested substrates, including the complex colloidal chitin. The prevalent activity was the β-*N*-acetyl-glucosaminidase one. The pH effect on the enzyme

activities was then investigated in details by the fluorimetric assay in the pH range from 3.0 to 9.0 (Figure 2a). The optimum pH was 4.0 for the chitobiosidase activity and 5.0 for β-*N*-acetyl-glucosaminidase and endochitinase ones. Considerable β-*N*-acetyl-glucosaminidase activity was conserved either at pH 3.0 or at pH 7.0. The temperature influence on the chitinase was fluorimetrically assayed in the temperature range from 5.0 to 70.0 °C. The optimum was at mesophilic temperatures (between 35 and 40 °C). However, the enzyme mixture proved to be active in a wide range of temperatures, retaining between 30 and 50% of activity below 25 °C and at 70 °C (Figure 2b). When stored at 4 °C, the enzyme cocktail retained more than 93% and 80% activity after 24 and 72 h respectively, and almost 45% of the initial overall activity was recorded even after 192 h. Also at 25 °C the enzyme cocktail showed appreciable long term stability: 70% activity was maintained for 70 h, and 33% of the initial activity was recovered after 192 h incubation (Figure 2c).

Effect of *T. viride* chitinases on PM structure and morphology

At light microscopy the PM appeared as a membranous sheath whose thickness did not significantly change along its length (Figure 3a). The different layers that form the PM, secreted by the epithelial cells, were well recognisable at SEM (Figure 3b). PM surface was continuous, with a smooth, felt-like texture, even though at higher magnification surface wrinkles were observable (Figure 3c). A limited amount of small pores was visible on the surface (Figure 3d). The morphology of both ectoperitrophic and endoperitrophic sides did not differ significantly, except for the presence of food debris/residues on the latter.

TEM analysis showed that the PM had a well-organised and compact structure (Figure 3e). Indeed, chitin fibrils were properly aligned and aggregated into compact bundles/layers; additionally, electron dense granules with a periodical distribution could be observed (Figures 3e and 3f). Occasionally it was possible to observe bacteria entrapped among chitin fibrils. (Figure 3f).

Treatment with *T. viride* chitinases (at two different concentrations and pHs) significantly altered PM structure, as observed at SEM: ruptures and scrapes of the superficial layers were frequently visible (Figures 4a and 4b). Moreover, chitinase-exposed PM was characterised by a highly porous surface (Figure 4c). Disruptions of the integrity and coalescence of the chitin fibril network were distinctive features of these chitinase-treated PMs (Figures 4d and 4e). General weakening of the PM structure also led to an increased

degree of ruptures due to manipulation of the samples (Figure 4f). The effects on the structure and ultrastructure of PM induced by increasing concentrations of chitinases were similar, although the PM appeared progressively less robust during manipulation. None difference was observed using chitinase cocktail at pH 5.0 or 7.0.

TEM observations after chitinase treatment evidenced a substantial decrease in PM organisation, with separate layers and broken chitin fibrils that formed bundles of fibrils (Figures 4g and 4h). Bacteria previously embedded in the PM were now freely disposed in the spaces among the PM layers (Figure 4i).

Effect of *T. viride* chitinases on PM permeability

To verify if the alterations induced by the *T. viride* chitinolytic enzymes observed by electron microscopy caused an increase of the PM permeability, we determined *in vitro* the flux of methylene blue through PMs isolated from larvae at the second day of the 5th instar. In fact, as reported in Fiandra *et al*. (Fiandra *et al*., 2009) and confirmed by our experiments, the permeability to methylene blue of *B. mori* PM does not change during the 5th instar, except for the few hours following ecdysis, in which a significantly higher permeation is observed. Experiments with Ussing chamber were performed with two different chitinase concentrations but only at pH 7.0, since (i) previous analysis at SEM and TEM did not show significant differences between treatments at pH 5.0 or 7.0, and (ii) the midgut lumen of lepidopteran larvae is characterised by neutral pH values. As reported in figure 5, the chitinases from *T. viride* caused an increase of the PM permeability in a dose-dependent manner. These results indicate that the alterations in chitin organisation induced by these enzymes impair the barrier function of the PM.

Discussion

Chitinases have recently attracted interest for their potential use in a wide range of biotechnological applications (Hjort *et al*., 2014; Dahiya *et al*., 2006; Duo-Chuan, 2006). Chitin-derived compounds are currently employed for medical, pharmaceutical and industrial purposes (Howard *et al*., 2003; Adrangi and Faramarzi, 2013). More recently, chitinases have become attractive also as biocontrol agents for plant protection, against

both insect pests and fungal pathogens (Howard *et al*., 2003; De Boer *et al*., 2001; Fung *et al*., 2002; Liu *et al*., 2002; Karasuda *et al*., 2003; Dahiya *et al*., 2005). For this reason, they are favourite candidates for developing IPM strategies based on enzymes that produce synergistic effects in different organisms. In this context, fungal chitinases, already employed as suppressive agents of fungal phytopathogens and known to attack the insect cuticles, thus permitting pathogen penetration and infection, need to be better explored for their action to insect gut systems. With the aim to overcome this lack of knowledge, in the present work we performed a biochemical characterisation of a commercial mixture of chitinolytic enzymes derived from *Trichoderma viride* and we analysed its effects on the PM of the lepidopteron *Bombyx mori*. *T. viride* is a filamentous mycoparasitic fungus, already reported as a potential biocontrol agent against soil borne plant pathogens and thus worthy to be assayed as a control agent against insect pests (Schuster and Schmoll, 2010). In addition, mycoparasitic fungi as *T. viride* are known producers of multiple chitinolytic enzymes with different substrate specificity and with a synergistic and complementary effect between them (Li, 2006; Rogalski *et al*., 1997; Omumasaba *et al*., 2001; Giridhar *et al*., 2012). The commercial product chosen for this study offered the possibility to test an easily-available cocktail of chitinolytic enzymes, which exhibit exo- and endochitinase activities, and therefore potentially able to strongly affect the PM organisation. The biochemical characterisation of the chitinase confirmed that these diverse chitinolytic activities are on the whole effective in degrading colloidal chitin and stable over a wide range of temperatures and pHs, including those experimental conditions that are compatible with the treatment of insect larvae. To impair a complex and insoluble polymer as chitin, the use of an enzymatic cocktail is definitively preferable to the pure enzymes, and additionally it better suits to environmental changing conditions as those occurring *in vivo.* The potential disadvantage of using an enzyme cocktail consists in the risk of its low reproducibility, which can be overcome by implementing a quality control of its composition as done by the different assays used in this work.

In vitro exposure of silkworm PM to fungal chitinases demonstrated that the matrix was considerably affected by these enzymes as peeling of the superficial layers and ruptures due to a general weakening of the matrix were observable. These effects were substantiated by the measurements of methylene blue fluxes through the isolated PM in Ussing chamber. The chitinolytic enzymes caused a significant increase of the permeability to methylene blue,

especially at the highest tested dose. These results are similar to those achieved in previous works in which the effects of viral, bacterial and insect chitinases on the PM of Lepidoptera/pests were tested. In particular, the most detailed study of chitinases on PM was performed on ChiA from the *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) (Rao *et al*., 2004; Corrado *et al*.; 2008, Fiandra *et al*.; 2010, Di Maro *et al*.; 2010). The protein has been heterologously expressed and purified in *E. coli* (Rao *et al*., 2004) and in *Nicotiana tabacum* (Corrado *et al*., 2008; Di Maro *et al*., 2010). Its hydrolytic activity was confirmed *in vitro* by treatment of the silkworm *B. mori* (Rao *et al*. 2004; Corrado *et al*., 2008) and of the tobacco budworm *H. virescens* (Di Maro *et al*., 2010) PM. Additionally, its *in vivo* insecticidal activity was proved by feeding *B. mori* and *H. virescens* larvae respectively with an artificial diet supplemented with the purified chitinase (Rao *et al*., 2004; Corrado *et al*., 2008) or directly with the transgenic tobacco leaves (Corrado *et al*., 2008; Fiandra *et al*., 2010). Also bacterial chitinases, from the Gram-positive *Bacillus* sp. (Thamthiankul *et al*., 2004) and *Kurthia zopfii* (Otsu *et al*., 2003), as well as from the Gram-negative *Serratia* spp. (Edwards and Jacobs-Loren, 2000; Regev *et al*., 1996, Huber *et al*., 1991), demonstrated a clear effect on PMs of different insect pests. Additionally, a chitinase purified from *B. mori* caused high mortality in adults of the coleopterum *Monochamus alternatus* after oral ingestion (Kabir *et al*., 2006).

Due to the encouraging results achieved by using a characterised fungal chitinase mixture on an easy-to-handle insect model system such as *Bombyx mori*, our future investigations will be oriented into two directions. Firstly, we would like to extend such type of approach to other insect species of high economic importance as insect pests that reduce crop production or destroy stored food grains. Secondly, we might continue testing and characterising other fungal chitinases. Our work clearly demonstrates that although the *T. viride* chitinase mixture is more active at acid pH (5.0), the activity retained at pH 7.0 is sufficient to induce marked *in vitro* alteration in the PM. Most of the known fungal chitinases, with molecular masses ranging from 20 to 190 kDa, have similar pH and temperature profiles, being usually more active in the pH interval 4.0-7.0 and temperature range 20-40 °C (Li, 2006; Seidl, 2008). Some of them, despite having an optimum at acid or neutral pH, retain their hydrolytic activity also in alkaline environments (Kopparapu *et al*., 2012). Additionally, Li and co-workers (Li *et al*., 2010) reported about a chitinase from the fungus *Thermoascus aurantiacus* var. *levisporus* that has an optimum pH of activity between

8.0 and 10.0. Concluding, it may be worthy to expand our studies on other fungal chitinases from already isolated fungi or searching into metagenomes (Hjort *et al*, 2014), with the final goal to develop an enzyme-based IPM approach: this means that the selected enzymes should be concomitantly employed in field for attacking external and internal layers of insects and degrading fungal cell walls.

Acknowledgements

We are grateful to Dr. Silvia Cappellozza (CRA-API, Padova, Italy) for providing *B. mori* larvae. We acknowledge EUFP7/2007-2011 Metaexplore grant agreement No: 222625 to Flavia Marinelli for funding our research on novel chitinases. This work was also supported by grants from Fondo di Ateneo per la Ricerca to Flavia Marinelli and by MIUR fellowships to Francesca Berini. We appreciate the support from Consorzio Interuniversitario per le Biotecnologie (CIB).

References

- Adrangi S, Faramarzi MA, From bacteria to human: a journey into the world of chitinases. *Biotechnol Adv* **31**: 1786-1795 (2013).
- Anton GE, Barrett DM, Determination of reducing sugars with 3-methyl-2-benzothiazolinonehydrazone. *Anal Biochem* **305**(2):287-289 (2002).
- Barbehenn RV, Stannard J, Antioxidant defense of the midgut epithelium by the peritrophic envelope in caterpillars. *J Insect Physiol* **50**:783–790 (2004).
- Boller T, Hydrolytic enzymes in plant disease resistance. In *Plant-Microbe Interactions: Molecular and Genetic Perspectives*. Vol. 2: ed. Kosuger T and Nester EW; pp. 384-414. New York: Macmillan (1987).
- Cappellozza L, Cappellozza S, Saviane A, Sbrenna G, Artificial diet rearing system for the silkworm *Bombyx mori* (Lepidoptera: Bombycidae): effect of vitamin C deprivation on larval growth and cocoon production. *Appl Entomol Zool* **40**(3):405-412 (2005).
- Corrado G, Arciello S, Fanti P, Fiandra L, Garonna AP, Digilio MC, Lorito M, Giordana B, Pennacchio F, Rao R, The Chitinase A from the baculovirus AcMNPV enhances resistance to both fungi and herbivorous pests in tobacco. *Transgenic Research* **17**:557-571 (2008).
- Dahiya N, Tewari R, Tiwari RP, Hoondal GS, Chitinase production in solid-state fermentation by *Enterobacter* sp. NRG4 using statistical experimental design. *Curr Microbiol* **51**(4):222-228 (2005).
- Dahiya N, Tewari R, Hoondal GS, Biotechnological aspects of chitinolytic enzymes: a review. *Appl Microbiol Biotechnol* **71**(6):773-782 (2006).
- De Boer W, Klein Gunnewiek PJ, Kowalchuk GA, Van Veen JA, Growth of chitinolytic dune soil beta-subclass Proteobacteria in response to invading fungal hypae. *Appl Environ Microbiol* **67**(8):3358-3362 (2001).
- Di Maro A, Terracciano I, Sticco L, Fiandra L, Ruocco M, Corrado G, Parente A, Rao R, Purification and characterization of a viral chitinase active against plant pathogens and herbivores from transgenic tobacco. *J Biotechnol* **147**:1-6 (2010).
- Ding X, Luo Z, Xia L, Gao B, Sun Y, Zhang Y, Improving the insecticidal activity by expression of a recombinant cry1Ac gene with chitinase-encoding gene in acrystalliferous *Bacillus thuringiensis*. *Curr Microbiol* **56**:442-447 (2008).
- Duo-Chuan L, Review of fungal chitinases. *Mycopathologia* **161**(6):345-360 (2006).
- Edwards MJ, Jacobs-Lorena M, Permeability and disruption of the peritrophic matrix and caecal membrane from *Aedes aegypti* and *Anopheles gambiae* mosquito larvae. *J Insect Physiol* **46**:1313-1320 (2000).
- El-Sayed GN, Coudron TA, Ignoffo CM, Riba G, Chitinolytic activity and virulence associated with native and mutant isolates of the entomopathogenic fungus, *Nonomuraea rileyi*. *J Invertebr Pathol* **54**:394-403 (1989).
- Fang S, Wang L, Guo W, Zhang X, Peng D, Luo C, Yu Z, Sun M, *Bacillus thuringiensis* Bel protein enhances the toxicity of Cry1Ac protein to *Helicoverpa armigera* larvae by degrading insect intestinal mucin. *Appl Environ Microbiol* **75**:5237-5243 (2009).
- Fiandra L, Casartelli M, Cermenati G, Burlini N, Giordana B, The intestinal barrier in lepidopteran larvae: Permeability of the peritrophic membrane and of the midgut epithelium to two biologically active peptides. *J Insect Physiol* **55**:10-18 (2009).
- Fiandra L, Terracciano I, Fanti P, Garonna A, Ferracane L, Fogliano V, Casartelli M, Giordana B, Rao R, Pennacchio F, A viral chitinase enhances oral activity of TMOF. *Insect Biochem Mol Biol* **40**:533-540 (2010).
- Fung KL, Zhao KJ, He ZM, Chye ML, Tobacco-expressed *Brassica juncea* chitinase BjCHI1 shows antifungal activity in vitro. *Plant Mol Biol* **50**(2):283-294 (2002).

- Franzetti E, Huang ZJ, Shi YX, Xie K, Deng XJ, Li JP, Li QR, Yang WY, Zeng WN, Casartelli M, Deng HM, Cappellozza S, Grimaldi A, Xia Q, Feng Q, Cao Y, Tettamenti G, Autophagy precedes apoptosis during the remodelling of silkworm larval midgut. *Apoptosis* **17**(3):305-324 (2012).
- Giridhar D, Ravi Sankar N, Kirian Kumar V, Kartheek D, Rajanikanth P, Nagalakshmi Devamma M, Purification, characterization and antifungal activity of chitinase from *Trichoderma viride* N9. *Cell Tissue Res* **12**(2):3187-3192 (2012).
- Hartl L, Zach S, Seidl-Seiboth V, Fungal chitinases: diversity, mechanistic properties and biotechnological potential. A*ppl Microbiol Biotechnol* **93**(2):533-543 (2012).
- Hegedus D, Erlandson M, Gillott C, Toprak U, New insights into peritrophic matrix synthesis, architecture and function. *Annu Rev Entomol* **54**:285-302 (2009).
- Hjort K, Presti I, Elväng A, Marinelli F, Sjöling S, Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics. *Appl Microbiol Biotechnol* **98**(6):2819-2828 (2014).
- Hsu SC, Lockwood JL, Powdered chitin agar as a selective medium for enumeration of actinomycetes in water and soil. *Appl Microbiol* **29**(3):422-426 (1975).
- Howard MB, Ekborg NA, Weiner RM, Hutcheson SW, Detection and characterization of chitinases and other chitin-modifying enzymes. *J Ind Microbiol Biotechnol* **30**(11):627-635 (2003).
- Huber M, Cabib E, Miller LH, Malaria parasite chitinase and penetration of the mosquito peritrophic membrane. *Proc Natl Acad Sci USA* **88**:2807-2810 (1990).
- Kabir KE, Sugimoto H, Tado H, Endo K, Yamanaka A, Tanaka S, Koga D, Effect of *Bombyx mori* chitinase against Japanese pine sawyer (*Monochamus alternatus*) adults as a biopesticide. *Biosci Biotechnol Biochem* **70**(1):219-29 (2005).
- Karasuda S, Tanaka S, Kajihara H, Yamamoto Y, Koga D. Plant chitinase as a possibile biocontrol agent for use instead of chemical fungicides. *Biosci Biotechnol Biochem* **67**(1):221-224 (2003).
- Kopparapu NK, Zhou P, Zhang S, Yan Q, Liu Z, Jiang Z, Purification and characterization of a novel chitinase gene from *Paecilomyces thermophila* expressed in *Escherichia coli*. *Carbohydr Res* **347**(1):155-160 (2012).
- Kramer KJ, Koga D, Insect chitin: physical state, synthesis, degradation and metabolic regulation. *Insect Biochem* **16**: 851–877 (1986).
- Lehane MJ, Peritrophic matrix structure and function. *Annu Rev Entomol* **42**: 525-550 (1997).
- Li AN, Yu K, Liu HQ, Zhang J, Li H, Li DC, Two novel thermostable chitinase genes from thermophilic fungi: cloning, expression and characterization. *Bioresour Technol* **101**:5546-5551 (2010).
- Li DC, Review of fungal chitinases. *Mycopathologia* **161**:345-360 (2006).
- Liu M, Cai QX, Liu HZ, Zhang BH, Yan JP, Yuan ZM, Chitinolytic activities in *Bacillus thuringiensis* and their synergistic effects on larvicidal activity. *J Appl Microbiol* **93**(3):374-379 (2002).
- Lorito M, Woo SL, Harman GE, Monte E, Translational research on *Trichoderma*: from 'omics to the field. *Annu Rev Phytopathol* **48**:395-417 (2010).
- McCreath KJ, Gooday GW, A rapid and sensitive microassay for the determination of chitinolytic activity. *Microbiol Meth* **14**:227-229 (1992).
- Merzendorfer H, Zimoch L, Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J Exp Biol* **206**:4393–4412 (2003).
- Millonig G, Laboratory manual of biological electron microscopy, Saviolo editore, Vercelli, Italy (1976).
- Monteiro VN, do Nascimento Silva R, Steindorff AS, Costa FT, Noronha EF, Ricart CAO, Valle de Sousa M, Vainstein MH, Ulhoa CJ, New insights in *Trichoderma harzianum* antagonism of fungal plant pathogens by secreted protein analysis. *Curr Microbiol* **61**: 298-305 (2010).

- Mukherjee PK, Horwitz BA, Herrera-Estrella A, Schmoll M, Kenerley CM, *Trichoderma* research in the genome era. *Annu Rev Phytopathol* **51**:105-29 (2013).
- Omumasaba CA, Yoshida N, Ogawa K, Purification and characterization of a chitinase from *Trichoderma viride*. *J Gen Appl Microbiol* **47**:53-61 (2001).
- Otsu Y, Mori H, Komuta K, Shimizu H, Nogawa S, Matsuda Y, Nonomura T, Sakuratani Y, Tosa Y, Mayama S, Toyoda H, Suppression of leaf feeding and oviposition of phytophagous ladybird beetles (Coleoptera: Coccinellidae) bychitinase gene-transformed phylloplane bacteria and their specific bacteriophages entrapped in alginate gel beads. *J Econ Entomol* **96**(3):555-563 (2003).
- Pardo-López L, Muñoz-Garay C, Porta H, Rodríguez-Almazán C, Soberón M, Bravo A, Strategies to improve the insecticidal activity of Cry toxins from *Bacillus thuringiensis*. *Peptides* **30**(3):589-595 (2009).
- Rao R, Fiandra L, Giordana B, de Eguileor M, Congiu T, Burlini N, Arciello S, Corrado G, Pennacchio F, AcMNPV ChiA protein disrupts the peritrophic membrane and alters midgut physiology of *Bombyx mori* larvae. *Insect Biochem Mol Biol* **34**(11):1205-1213 (2004).
- Regev A, Keller M, Strizhov N, Synergistic activity of a *Bacillus thuringiensis* delta-endotoxin and a bacterial endochitinase against *Spodoptera littoralis* larvae. *Appl Environ Microbiol* **62**: 3581-3586 (1996).
- Rogalski J, Krasowska B, Glowiak G, Wojcik W, Targonski Z, Purification and some properties of extracellular chitinase produced by *Trichoderma viride* F-19. *Acta Microbiol Pol* **46**:363-375 (1997).
- Schägger H, von Jagow G, Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**(2):368-379 (1987).
- Schuster A, Schmoll M, Biology and biotechnology of *Trichoderma*. *Appl Microbiol Biote*chnol **87**:787-799 (2010).
- Seidl V, Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. *Fungal Biol Rev* **22**:36-42 (2008).
- Sun S, Cheng Z, Fan J, Chen X, Pang Y, The utility of camptothecin as a synergist of *Bacillus thuringiensis* var. *kurstaki* and nucleopolyhedroviruses against *Trichoplusia ni* and *Spodoptera exigua*. *J Econ Entomol* **105**:1164-1170 (2012).
- St. Leger RJ, Cooper RM, Charnley AJ, Characterization of chitinase and chitobiosidase produced by the entomopathogenic fungus *Metarhizium anisopliae*. *J Invertebr Pathol* **58**:415-426 (1991).
- Thamthiankul S, Moar WJ, Miller ME, Panbangred W. Improving the insecticidal activity of *Bacillus thuringiensis* subsp. *aizawai* against *Spodoptera exigua* by chromosomal expression of a chitinase gene. *Appl Microbiol Biotechnol* **65**:183-192 (2004).
- Tellam RL, Wijffels G, Willadsen P, Peritrophic matrix proteins. *Insect Biochem Mol Biol* **29**(2):87-101 (1999).
- Terra WR, The origin and functions of the insect peritrophic membrane and peritrophic gel. *Arch Insect Biochem Physiol* **47**(2):47-61 (2001).
- Wang P, Granados RR, Molecular structure of the peritrophic membrane (PM): identification of potential PM target sites for insect control. *Arch Insect Biochem Physiol* **47**(2):110-118 (2001).
- Zhuo W, Chu F, Kong L, Tao H, Sima Y, Xu S, Chitin synthase B: a midgut-specific gene induced by insect hormones and involved in food intake in *Bombyx mori* larvae. *Arch Insect Biochem Physiol* **85**(1):36-47 (2014).

Figure legends

Figure 1. Panel a: SDS-PAGE analysis of the chitinase cocktail from *T. viride*. LMW: molecular weight markers (GE-Healthcare Sciences, Little Chalfont, UK). Panel b: zymogram analysis of the chitinase with CM-chitin-RBV as substrate. Proteins are indicated by arrows.

Figure 2. Panels a and b: enzyme properties of the *T. viride* chitinase mixture, using 4-MU-(GlcNAc) (■), 4-MU-$(GlcNAc)_2$ (●) and 4-MU-$(GlcNAc)_3$ (▲) as substrates. Panel a: pH profile. Buffer employed (final concentration 100 mM) were glycine-HCl (pH 3.0), sodium acetate (pH 4.0 and 5.0), sodium phosphate (pH 6.0 and 7.0), TrisHCl (pH 8.0) and sodium pyrophosphate (pH 9.0). Panel b: temperature profile. Assays were performed in 100 mM sodium acetate buffer pH 5.0. Panel c: long-term stability of *T. viride* chitinase preparation; residual activities are expressed as sum of the single chitinolytic activities measured on the three fluorimetric substrates. The enzymes were incubated in PBS pH 7.0 at 4 °C (◆, solid line) or 25 °C (◇, dashed line) for several days. For each graph, the values represent the means of three independent experiments (mean ± standard deviation) and enzymatic activities are expressed as relative to the maximal activity recorded.

Figure 3. Morphology of control peritrophic matrix. Semi-thin cross section (a), SEM (b-d), TEM (e, f). The peritrophic matrix appears as a thin structure that lines the midgut epithelium (a). It is formed by a series of overlaid layers (b) that generally have a smooth appearance although, at higher magnification, wrinkles can be observed (c). In untreated larvae, some small pores are visible on PM surface (d). Ultrastructural analysis evidences the compact organisation of the chitin fibrils in the PM (e, f), which are linked by electron-dense structures (f). Bacteria are embedded among the PM layers (f).
Bars: 30 µm (a), 10 µm (b), 2 µm (c), 1 µm (d, e), 500 nm (f)

Figure 4. Morphology of peritrophic matrix treated with *T. viride* chitinases. SEM (a-f), TEM (g-i). PM treated with chitinases shows ruptures (a) and peeling (b) of the superficial layers. The number of pores increases significantly (c). The integrity of fibril network is compromised (d) and coalescence of fibrils occurs (d, e). The alteration of the PM structure

leads to a general weakening of the matrix, with consequent breakages probably due to manipulation (f). Ultrastructural analysis confirms a general collapse of the PM structure and a massive alteration in the organisation of the chitin fibrils (g-i).

Bars: 200 µm (a), 10 µm (b, f), 2 µm (c, e), 5 µm (d), 1 µm (g), 500 nm (h, i)

Figure 5. Flux of 0.5 mg/mL methylene blue across the PM isolated from *B. mori* larvae at the second or third day of the 5^{th} instar, in the absence (control) or in the presence of different amounts of chitinases. Chitinolytic activity is expressed as the sum of the activities recorded on the three fluorigenic substrates 4-MU-GlcNAc, 4-MU-$(GlcNAc)_2$ and 4-MU-$(GlcNAc)_3$. Bars represent mean values ± standard deviation of at least four replicates. * $P < 0.01$, ** $P < 0.001$ *vs* control, Student's *t* test.

Figures

Figure 1

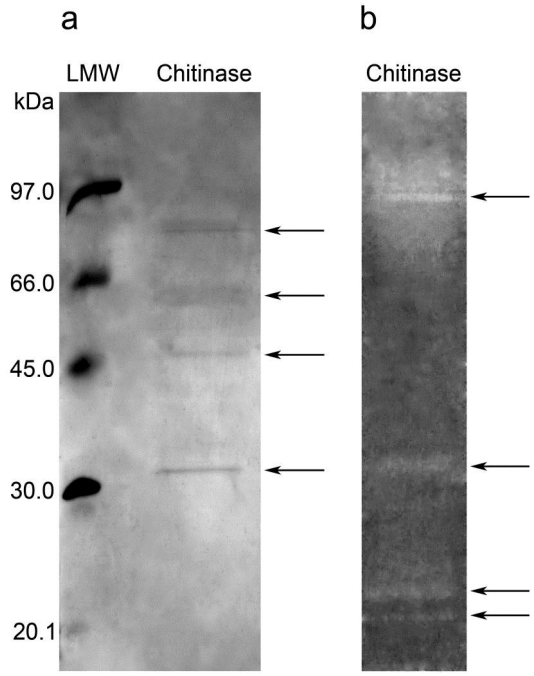


Figure 2

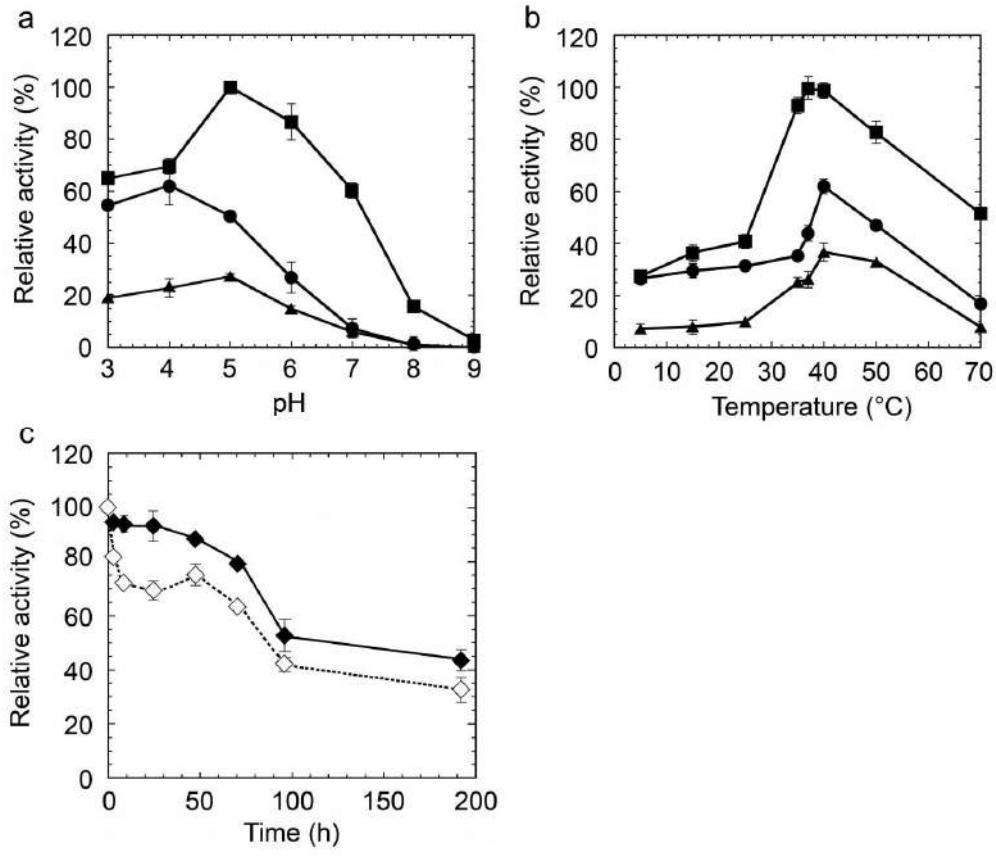


Figure 3

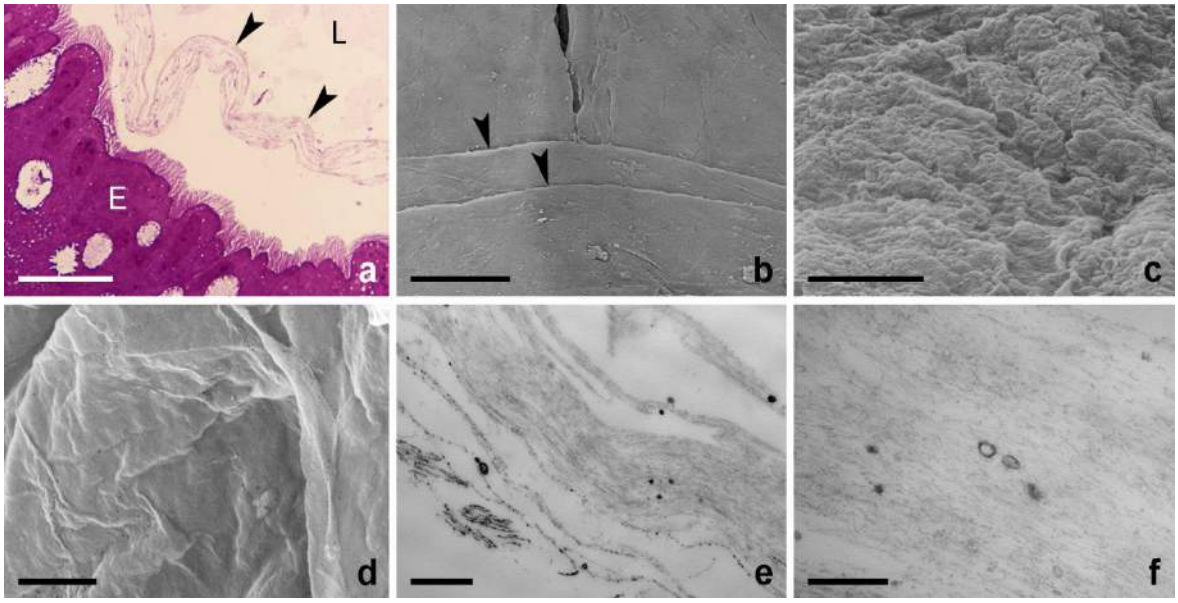


Figure 4

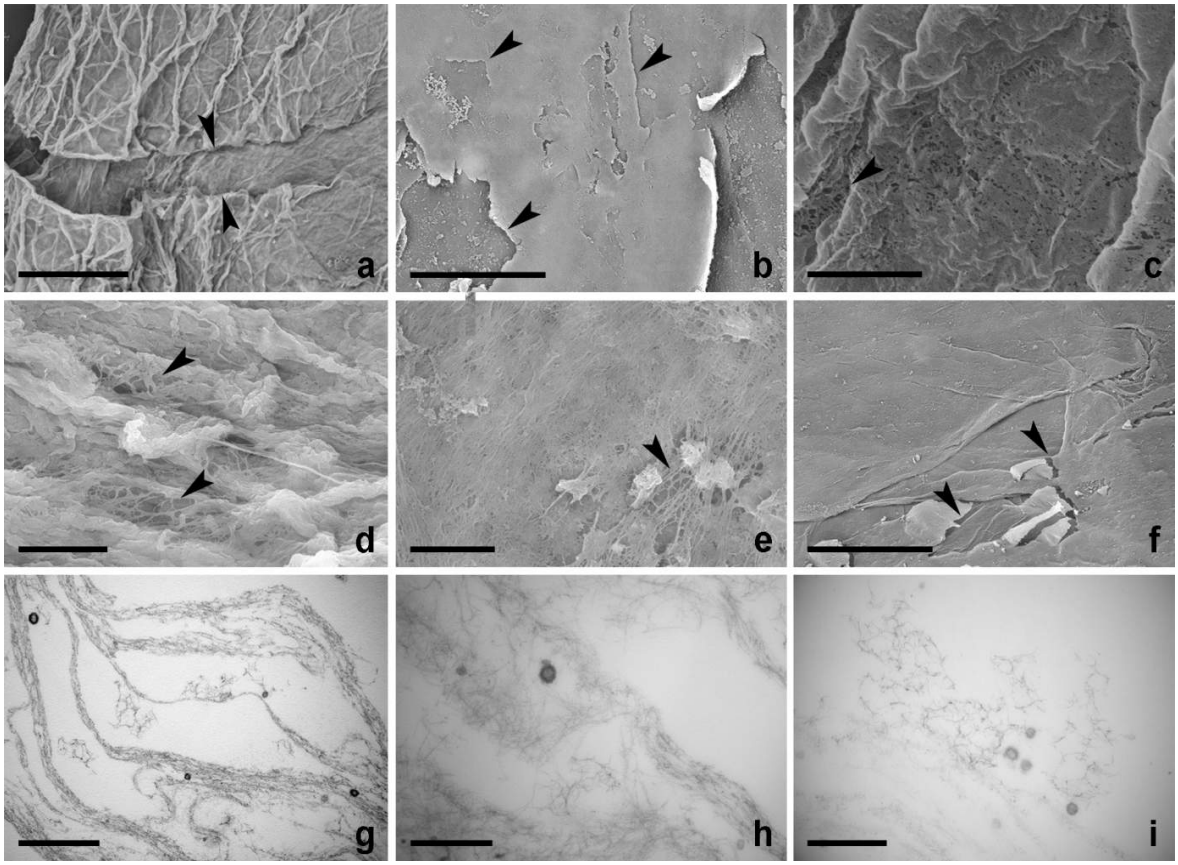
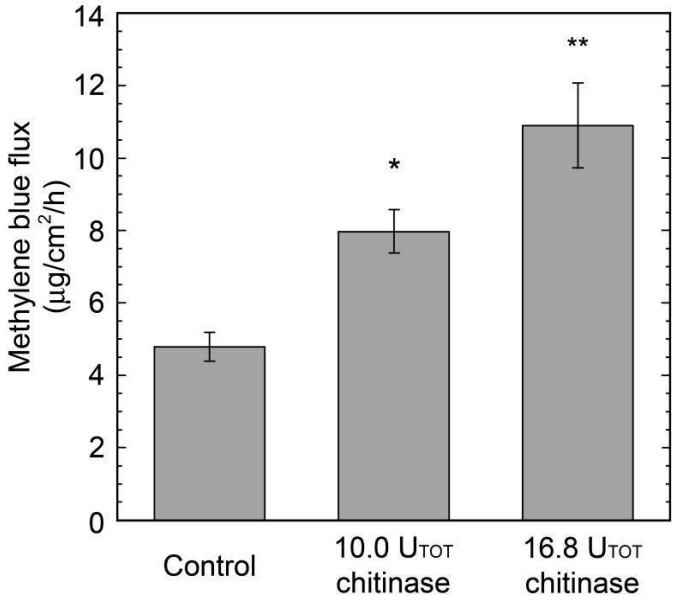


Figure 5



Tables

Table 1. *T. viride* chitinase activity on different substrates (mean ± standard deviation from at least three independent experiments).

Substrate	Type of activity detected	Specific activity (U/mg)	
		pH 5.0	pH 7.0
4-MU-GlcNAc	β-*N*-acetyl-glucosaminidase	35.87 ± 0.01	16.21 ± 0.04
4-MU-$(GlcNAc)_2$	Chitobiosidase	18.08 ± 0.03	2.35 ± 0.17
4-MU-$(GlcNAc)_3$	Endochitinase	12.14 ± 0.09	2.00 ±0.06
Colloidal chitin		5.87 ± 0.16	2.48 ± 0.09

CONCLUSIONS

My PhD has been accomplished in the frame of the FP7 European project MetaExplore, whose goal was the development of metagenomic tools and techniques for the identification of novel enzymes involved in the biodegradation of recalcitrant natural polymers, like chitins and lignins. The driving force of this project was the need from European industries of new biocatalysts endowed with innovative features and enhanced activities, to be employed in a vast range of industrial and agricultural processes. Nowadays, biocatalysis accounts for a large part of the chemical industry, but its contribution is foreseen to exponentially increase in the next years, as a central feature of the sustainable economic future of industrialised societies [1]. Microorganisms, the oldest form of life, able to live not only in nutrient-rich environments but also in the less-hospital habitats, encompass by far the largest resource of metabolic and genetic diversity encountered on Earth [2]. Unfortunately, since a major part of the microbiota in natural ecosystems (up to 99 – 99.9%) is unculturable by traditional microbiological methods, this unparalleled biodiversity risks to remain encrypted and underexploited [2, 3]. This prompted the development of culture-independent techniques, among which metagenomics is currently though to be the most promising one. Since their introduction, in fact, metagenomic approaches allowed the discovery and characterisation of a significant number of novel biocatalysts or molecules with high potential for use in pharmaceutical products or production processes. It is moreover conceivable that metagenomics, together with protein engineering and *in vitro* evolution technologies, might be employed to find suitable natural enzymes that can serve as backbone to produce ideal biocatalysts, *i.e.* improved tailored enzymes that optimally fit specific process requirements [1].

At the beginning of the MetaExplore project (in May 2009) the premise of metagenomics as a source of new technology was not fully realised, primarily because of challenges in screening and producing the desired enzymatic activities. Indeed, at that date bacterial chitinases, which represent the core of this PhD dissertation, had been identified almost exclusively by conventional molecular or functional screening approaches. Only a couple of studies had focused on culture-independent screening of chitinolytic enzymes from bacterial isolates, but none of them had resulted in an in-depth characterisation of these new biocatalysts. In 1999, Cottrell and co-workers [4] screened libraries from coastal and estuarine waters with 4-MU-$(GlcNAc)_2$ as substrate and identified eleven putative clones, whose corresponding enzymes were classified by zymogram analysis and activity assays. The

first example of metagenomic library screened through a sequence-based approach, on the contrary, dated back to 2004 [5]: PCR screening of clone libraries from ten aquatic environments, with a degenerate primer set for family 18 chitinases, led to the identification of several chitinase genes, all new if compared to the previously identified sequences. Finally, in 2006 the development of IAN-PCR (inverse affinity nested PCR) for metagenome walking allowed the successful fishing of complete family 18 genes from groundwater metagenome [6].

During the five years of the MetaExplore project, a few more papers have been published, concerning the application of metagenome and metagenome-like approaches for the discovery of new chitinolytic enzymes in different environments. Most of them, however, report exclusively on the identification of putative chitinase sequences, without conducting further experiments to characterise them. A partial characterisation of library-sourced chitinolytic enzymes is present only in [7] and [8]. In the first paper, a metagenome-like library was constructed using DNA extracted from a cell mixture of pure-cultured chitinolytic bacteria, followed by functional-based screening and heterologous expression of the most promising gene, *chi22718_II*. Activity assays, performed directly on the crude enzyme solution, revealed that Chi22718_III is a thermolabile chitinase, a typical characteristic of cold-active enzymes [7]. On the other hand, the recent paper of Stöveken and co-workers describes the metagenomic analysis of chitin-enriched soil samples, which led to the identification of several putative genes for chitin and chitosan modifying enzymes, including the full-length chitinase gene *chiA01*. The sequence was codon-optimised, the protein heterologously expressed in *E. coli* and its activity confirmed by glycol chitin dot assay [8].

Nevertheless, the two chitinases described in the first section of the present thesis, Chi18H8 and 53D1, represent the real first two examples of chitinolytic enzymes identified by metagenome library screening, expressed and fully characterised. The two chitinases show 47% amino acid sequence identity between each other and less than 50% similarity with any other known chitinase. Chi18H8 and 53D1 display some common features, for instance the prevalent chitobiosidase activity and the higher stability at acid pH and mesophilic temperatures. This could be easily explained taking into account that the two enzymes belong to the same group (family 18 of glycosyl hydrolases) and both have been identified in metagenomic libraries of sub-acidic temperate soils. Nonetheless, the two chitinases possess

also peculiar characteristics, which differentiate one from the other and, more importantly, from other already characterised chitinolytic enzymes.

The first chitinase, Chi18H8, revealed a remarkable antagonistic activity against common plant phytopathogens, thus representing an environmental-friendly alternative to synthetic fungicides [9]. Nowadays diseases of plants caused by fungal pathogens globally contribute to extensive loss of crops important for food and energy production, an effect that is increased by the norm of monoculture practise [10]. Microbiological control of fungal diseases, by employing bacteria with antifungal action or enzymatic formulations possessing antiphytopathogenic activities, represents the only sustainable solution for limiting the use of toxic chemicals. Biotechnologically relevant is also Chi18H8 high solvent-tolerance. If one considers that time-consuming immobilisation, mutagenesis and protein engineering procedures often need to be applied to increase enzymatic activity and stability in organic solvents, the intrinsic solvent-tolerance of Chi18H8 can represent a key factor and a great advantage for its exploitation in non-aqueous enzymology [11].

On the contrary, the most interesting feature of 53D1 is its remarkable stability and activity in the presence of high salt concentrations. This, combined with its activity also on the complex substrate colloidal chitin, suggested a possible application of this biocatalyst in the downstream processing of exoskeletal waste (the carapace) in the industrialisation of foodstuff. Nowadays, chitin extraction and derivatisation are usually exerted by harsh treatments with acids and high temperatures. The employment of chitinolytic enzymes can therefore represent a sustainable alternative to these traditional chemical treatments; it can help in solving the environmental and economic problem connected to the high amount of recalcitrant marine wastes generated annually by food industries and, meanwhile, in producing from these waste materials value-added compounds (chitosan, glucosamines and chitooligosaccharides) with high pharmaceutical and nutritional potential. It has been estimated, for example, that by 2015 the market for chitin and chitooligosaccharides will reach up to 63 billion US$, while that for chitosan up to 21 billion US$ [12]. It is therefore easy to understand why industries are interested in developing cheap, efficient and ecological processes for the treatment of such recalcitrant polymers.

Before Chi18H8 and 53D1 could be actually used for in-field applications, additional experiments need to be done; it is moreover relevant to develop reproducible and robust process to produce these proteins and sustain their development. Anyhow, the work

accomplished in these three years clearly demonstrates the great potential of metagenomics for the identification of novel valuable biocatalysts. Metagenomic approaches are still in development but have the high potential to substantially impact industrial production.

The potential application of chitinolytic enzymes as biocontrol agents has been investigated and demonstrated not only against phytopathogen fungi, but also in respect to insect pests. In the third section of the present dissertation, the commercial mixture of chitinases from *Trichoderma viride* proved to be effective in altering *in vitro* the structure and the permeability of the peritrophic membrane of the model lepidopteron *Bombyx mori*. Control of insect pests by the application of enzymes holds great promise as an alternative to the use of chemical pesticides. It is, in fact, generally recognised that biological control agents are safer and more environmentally sound than is reliance on the use of high volumes of chemical pesticides, which, moreover, may be lethal to beneficial insects and microorganisms populating the soil, and may enter the food chain. Additionally, enzyme formulations can represent an advantage even to the use of entire microorganisms, since living organisms are less verifiable and can show short shelf lives or inconsistent performances in field [10].
The fungal chitinases herein tested represent promising candidates for integrated pest management. Interesting could be verifying their effect *in vivo*, even in combination with other bioactive peptides and lytic enzymes, as found in natural systems, as well as confirming their activity also on pest Lepidoptera. Anyway, the preliminary work done confirms that biotechnology, alone or in conjunction with conventional breeding programs, can make significant contributions to sustainable agriculture.

Finally, the work performed during my PhD had shed light on the use of *E. coli* and streptomycetes as expression platforms for the production of recombinant enzymes. It is impossible to unconditionally claim which microorganism is the best expression system, both having peculiar advantages and disadvantages. *E. coli* continues to dominate the bacterial expression systems and remains the first choice for laboratory investigations and initial development in commercial activities. The unparalleled fast growth on cheap substrates and the availability of extensive genetic toolkits are the major factors supporting *E. coli* employment as heterologous host [13]. However, as demonstrated for both Chi18H8 and

53D1 chitinases, recombinant protein expression in this host often results in enzyme accumulation into inclusion bodies, from which protein recovery is not always effective. Indeed, the process for the identification of a suitable protocol for Chi18H8 solubilisation was strenuous and time-consuming. Nevertheless, now that it has been defined, high amount of recombinant protein could be potentially produced in a fast, reproducible and economically feasible way. On the other hand, extracellular production of proteins is highly desirable as it could reduce the complexity of downstream processes and possibly improve product quality. In this respect, the use of streptomycetes constitutes a great advantage on *E. coli*, where protein secretion into the culture broth or even to the periplasmic space is rare. Expression of $VanY_n$ in *S. venezuelae* ATCC 10595 greatly improved the production yield if compared with the previous expression of this D,D-peptidase/D,D-carboxypeptidase in *E. coli*. For Chi18H8, *S. lividans* TK24 represented a promising system thanks to the chitinase secretion into culture broth that makes its purification procedure much easier; actually it cannot be considered competitive in production yield if compared with *E. coli* and further trials should be done for optimising the codon usage of the gene sequence and the signal peptide. Concluding, both *E. coli* and *Streptomyces* spp. proved to be valuable and powerful expression platforms. Even if a general rule for heterologous protein production does not exist and the procedures must be tailored to each single protein, the achieved results and the methods developed in these three years for both microorganisms, can be eventually applied and used for the heterologous expression of other enzymes (metagenome-sourced and not).

1. Lorenz P, Eck J: **Metagenomics and industrial applications**. *Nat Rev Microbiol* 2005, **3**(6):510-516.
2. Simon C, Daniel R: **Metagenomic analyses: past and future trends**. *Appl Environ Microbiol* 2011, **77**(4):1153-1161.
3. Ekkers DM, Cretoiu MS, Kielak AM, Elsas JD: **The great screen anomaly--a new frontier in product discovery through functional metagenomics**. *Appl Microbiol Biotechnol* 2012, **93**(3):1005-1020.
4. Cottrell MT, Moore JA, Kirchman DL: **Chitinases from uncultured marine microorganisms**. *Appl Environ Microbiol* 1999, **65**(6):2553-2557.
5. LeCleir GR, Buchan A, Hollibaugh JT: **Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions**. *Appl Environ Microbiol* 2004, **70**(12):6977-6983.
6. Uchiyama T, Watanabe K: **Improved inverse PCR scheme for metagenome walking**. *Biotechniques* 2006, **41**(2):183-188.
7. Kim D, Park HJ, Kim IC, Yim JH: **A new approach for discovering cold-active enzymes in a cell mixture of pure-cultured bacteria**. *Biotechnol Lett* 2014, **36**(3):567-573.
8. Stöveken J, Singh R, Kolkenbrock S, Zakrzewski M, Wibberg D, Eikmeyer FG, Pühler A, Schlüter A, Moerschbacher BM: **Successful heterologous expression of a novel chitinase identified by sequence analyses of the metagenome from a chitin-enriched soil sample**. *J Biotechnol* 2014.
9. Hjort K, Presti I, Elväng A, Marinelli F, Sjöling S: **Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics**. *Appl Microbiol Biotechnol* 2014, **98**(6):2819-2828.
10. Herrera-Estrella A, Chet I: **Chitinases in biological control**. *EXS* 1999, **87**:171-184.
11. Gupta A, Khare SK: **Enzymes from solvent-tolerant microbes: useful biocatalysts for non-aqueous enzymology**. *Crit Rev Biotechnol* 2009, **29**(1):44-54.
12. Lodhi G, Kim YS, Hwang JW, Kim SK, Jeon YJ, Je JY, Ahn CB, Moon SH, Jeon BT, Park PJ: **Chitooligosaccharide and its derivatives: preparation and biological applications**. *Biomed Res Int* 2014, **2014**:654913.
13. Rosano GL, Ceccarelli EA: **Recombinant protein expression in *Escherichia coli*: advances and challenges**. *Front Microbiol* 2014, **5**:172.