

**Engineering and optimization of
Bacillus megaterium for the biosynthesis and
biotransformation of terpenoid derived compounds**

Kumulative Dissertation

zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät
der Universität des Saarlandes

von

Philip Hartz

Saarbrücken

2019

Tag des Kolloquiums: 15.11.2019

Dekan: Herr Prof. Dr. Guido Kickelbick

Berichterstatter: Frau Prof. Dr. Rita Bernhardt

Herr Prof. Dr. Gert-Wieland Kohring

Frau Prof. Dr. Vlada Urlacher

Vorsitz: Herr Prof. Dr. Uli Müller

Akad. Mitarbeiter: Herr Dr. Konstantin Lepikhov

Danksagung

Mein größter Dank gilt Frau Prof. Dr. Rita Bernhardt, die mir die Anfertigung dieser Promotionsarbeit in ihrer Gruppe ermöglicht hat. Vielen Dank für das Vertrauen, die Geduld, die Motivation und die intensive wissenschaftliche Betreuung.

Ebenso möchte ich Herrn Prof. Dr. Gert-Wieland Kohring für die Übernahme der Zweitkorrektur danken.

Herrn Dr. Frank Hannemann danke ich nicht nur für die hervorragende Koordination des Projektes, sondern auch für die zahlreichen nützlichen Ratschläge, die wesentlich zum Gelingen dieser Arbeit beigetragen haben.

Im Speziellen bedanke ich mich bei Frau Dr. Martina Schad von OakLabs GmbH und bei Frau Dr. Sandra Trenkamp von Metabolomic Discoveries GmbH für die kooperative Zusammenarbeit und den interessanten Blick hinter die Kulissen während unseres Projekttreffens in Berlin.

Ein herzliches Dankeschön ebenfalls an Antje Eiden-Plach und Birgit Heider-Lips für die nützlichen Tipps und die Unterstützung im Umgang mit dem Bioreaktor. Gabi Schon danke ich für ihre stete Hilfsbereitschaft bei der Bewältigung aller bürokratischen Angelegenheiten und Hürden.

Als Arbeitskollege danke ich Dr. Mohammed Milhim, Dr. Ammar Abdulmughni, Ömer Kurt und Marc Finkler für die konstruktiven Diskussionen, die Inspiration und die Motivation bei der wissenschaftlichen Ausarbeitung dieser Dissertation. Als Freund danke ich euch für euren Zuspruch und die zahlreichen außergewöhnlichen und meist humorvollen Momente.

Ganz besonders danke ich auch meinen Masterstudenten Roxana, Carsten und Manuel, deren Zuarbeit wesentlich zum erfolgreichen Abschluss dieser Promotionsarbeit beigetragen hat.

Weiterhin möchte ich allen namentlich nicht genannten aktuellen sowie ehemaligen Mitarbeiterinnen und Mitarbeitern der Arbeitsgruppe für die angenehme und wissenschaftlich stimulierende Zusammenarbeit danken.

An dieser Stelle ein herzliches Dankeschön an meine Amigos Anna, Carolin, Julia, Paula, Jörn, Pascal, Simon und Stephan für die gemeinsam erlebten, schönen und gelegentlich bizarren, aber unvergesslichen Momente außerhalb des Laboralltages.

Zuletzt möchte ich meiner Familie danken, die mich im Verlauf meiner bisherigen universitären Karriere in allen Belangen bedingungslos unterstützt hat.

Scientific contributions

The thesis is based on six original research papers reproduced in chapter 2. The original manuscripts are printed with the kind of permissions from the Journal of Biotechnology (2.1 Hartz et al (2019), 2.4 Milhim et al (2019) and 2.6 Milhim et al (2016)) and Metabolic Engineering (2.2 Gerber et al (2016), 2.3 König et al (2019) and 2.5 Hartz et al (2018)).

2.1 (Hartz et al., 2019)

Hartz, P., Mattes, C., Schad, M., Bernhardt, R., and Hannemann, F. (2019). Expanding the promoter toolbox of *Bacillus megaterium*. J. Biotechnol. 294, 38-48

The author prepared RNA samples for the transcriptome analyses of *Bacillus megaterium* and identified novel promoter sequences based on the resulting differential gene expression profiles. Furthermore, the author generated a β -galactosidase (LacZ) deficient *B. megaterium* strain and characterized the activity of the novel promoters via LacZ assay. In addition the author applied the novel promoters for steroid bioconversions, interpreted the data and wrote the manuscript.

2.2 (Gerber et al., 2016)

Gerber, A., Milhim, M., **Hartz, P.**, Zapp, J., and Bernhardt, R. (2016). Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one. Metab. Eng. 36, 19–27.

The author carried out the biotransformation and purification of 17 α ,20 α -dihydroxy-4-pregnen-3-one.

2.3 (König et al., 2019)

König, L., **Hartz, P.**, Bernhardt, R., and Hannemann F. (2019). High-yield C11-oxidation of hydrocortisone by establishment of an efficient whole-cell system in *Bacillus megaterium*. Metab. Eng. 55, 59–67.

The author provided promoter sequences for the improved recombinant protein expression, assisted in the conduction of the experiments and interpretation of the results.

2.4 (Milhim et al., 2019)

Milhim, M., **Hartz, P.**, Gerber A., and Bernhardt, R. (2019). A novel short chain dehydrogenase from *Bacillus megaterium* for the conversion of the sesquiterpene nootkatol to (+)-nootkatone. *J. Biotechnol.* *301*, 52-55

The author carried out part of the cloning experiments and participated in the analysis and discussion of the experimental results.

2.5 (Hartz et al., 2018)

Hartz, P., Milhim, M., Trenkamp, S., Bernhardt, R., and Hannemann, F. (2018). Characterization and engineering of a carotenoid biosynthesis operon from *Bacillus megaterium*. *Metab. Eng.* *49*, 47-58.

The author prepared samples for the metabolome analyses of *Bacillus megaterium* and identified precursor metabolites for the biosynthesis of carotenoids. In addition, the author cloned, expressed and characterized the ORFs of the putative carotenoid biosynthesis operon. Furthermore the author contributed to the interpretation and discussion of the experimental results as well as the writing of the manuscript.

2.6 (Milhim et al., 2016)

Milhim, M., Putkaradze, N., Abdulmughni, A., Kern, F., **Hartz, P.**, and Bernhardt, R. (2016). Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. *J. Biotechnol.* *240*, 68–75.

The author participated in the interpretation and discussion of the experimental results and writing of the manuscript.

Contents

Danksagung	I
Scientific contributions	II
Contents	III
Abbreviations	IV
Summary	V
Zusammenfassung	VI
1. Introduction	1
1.1 Microbial strain development	1
1.2 Promoter systems and their contribution to microbial strain development	2
1.3 <i>Bacillus megaterium</i>	5
1.4 Biosynthesis of terpenoids.....	6
1.5 Biosynthesis and biotransformation of steroids.....	8
1.6 Structure, function and biosynthesis of carotenoids	10
1.7 Aim of this work.....	14
2. Scientific articles	15
2.1 Hartz et al (2019).....	15
Expanding the promoter toolbox of <i>Bacillus megaterium</i>	16
Supplementary information.....	27
2.2 Gerber et al (2016).....	34
Genetic engineering of <i>Bacillus megaterium</i> for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one.....	35
Supplementary information.....	44
2.3 König et al (2019).....	51
High-yield C11-oxidation of hydrocortisone by establishment of an efficient whole-cell system in <i>Bacillus megaterium</i>	52
Supplementary information.....	61
2.4 Milhim et al (2019).....	65
A novel short chain dehydrogenase from <i>Bacillus megaterium</i> for the conversion of the sesquiterpene nootkatol to (+)-nootkatone.....	66
2.5 Hartz et al (2018).....	70
Characterization and engineering of a carotenoid biosynthesis operon from <i>Bacillus megaterium</i>	71

2.6 Milhim et al (2016).....	83
Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from <i>Bacillus megaterium</i> with a catalytic activity towards mevastatin.....	84
Supplementary information.....	92
3. General discussion.....	96
3.1 Identification and characterization of novel promoter systems in <i>B. megaterium</i>	96
3.2 Biotransformation of steroids in <i>B. megaterium</i> using new promoters.....	101
3.3 Identification of a novel short chain dehydrogenase for the production of the sesquiterpene (+)-nootkatone in <i>B. megaterium</i>	104
3.4 Biosynthesis of C30 carotenoids in <i>B. megaterium</i>	105
3.5 Biosynthesis of C40 carotenoids in <i>B. megaterium</i>	111
4. Future perspectives.....	114
4.1 Application of the novel promoters and rational strain design of <i>B. megaterium</i> for the efficient biotransformation of steroids.....	114
4.2 Metabolic engineering of <i>B. megaterium</i> towards an improved carotenoid production.....	115
Appendix.....	VII
References.....	VIII

Abbreviations

ABA	Abscisic acid
Acat	Acetyl-CoA-acetyltransferase
ADH	Alcohol dehydrogenase
AdR	Adrenodoxin reductase
Adx	Adrenodoxin
ATP	Adenosine triphosphate
BCO2	<i>Brevibacterium sterolicum</i> cholesterol oxidase 2
Bm	<i>Bacillus megaterium</i>
CCR	Carbon catabolite repression
CDP	Cytidine diphosphate
CD-MEP	4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate
CDP-ME	4-diphosphocytidyl-2-C-methyl-D-erythritol
CoA	Coenzyme A
CO ₂	Carbon dioxide
CrtB	Phytoene synthase
CrtE	GGDP synthase
CrtI	Phytoene desaturase
CrtM	4,4'-diapophytoene synthase
CrtN	4,4'-diapocarotenoid desaturase
CrtP	4,4'-diapocarotenoid oxygenase
CrtY	Lycopene cyclase
CYP/P450	Cytochrome P450
DLPN	4,4'-diapolycopene
DMADP	Dimethylallyl diphosphate
DNA	Deoxyribonucleic acid
DNSA	4,4'-diaponeurosporenic acid
DNSP	4,4'-diaponeurosporene
DOC	11-deoxycorticosterone/21-hydroxyprogesterone
DPFL	4,4'-diapophytofluene
DPHY	4,4'-diapophytoene
DXP	1-deoxy-D-xylulose 5-phosphate
Dxs	DXP synthase
Dxr	DXP reductase
EDTA	Ethylenediaminetetraacetate
EMS	Ethyl methanesulfonate

FabG	3-oxoacyl-(acyl-carrier-protein) reductase
FAD	Flavine adenine dinucleotide
FDP	Farnesyl diphosphate
FMN	Flavine mononucleotide
GC	Gas chromatography
GDP	Geranyl diphosphate
GGDP	Geranylgeranyl diphosphate
GRAS	Generally recognized as safe
h	Hour
HMB-DP	(<i>E</i>)-4-Hydroxy-3-methyl-but-2-enyl-diphosphate
HMG- CoA	3-hydroxy-3-methyl-glutaryl-CoA
Hmgs	HMG-CoA synthase
Hmgr	HMG-CoA reductase
HPLC	High performance liquid chromatography
HSD	Hydroxysteroid dehydrogenase
Idi	IDP isomerase
IDP	Isopentenyl diphosphate
IspD	MEP cytidyltransferase
IspE	CDP-ME kinase
IspF	MEcDP synthase
IspG	HMB-DP synthase
IspH	IDP and DMADP synthase
L	Liter
LacZ	β -galactosidase
Lb	<i>Lactobacillus brevis</i>
LC	Liquid chromatography
M	Molar
MCS	Multiple cloning site
Mdc	Diphosphomevalonate decarboxylase
Mek	Mevalonate kinase
MEcDP	2-C-methyl-D-erythritol-2,4-cyclodiphosphate
MEP	Methylerythritol phosphate
MEV	Mevalonate
mg	Milligram
μ m	Micrometer
min	Minute
MS	Mass spectrometry

NAD(P)H	Nicotinamide adenine dinucleotide (phosphate)
NPQ	Non-photochemical quenching
ORF	Open reading frame
ori	Origin of replication
PAGE	Polyacrylamide gel electrophoresis
Pa	<i>Pantoea ananatis</i>
PDS	Prenyl diphosphate synthase
PHA	Polyhydroxyalkanoate
Pmk	phosphomevalonate kinase
RBS	Ribosomal binding site
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RSS	Reichstein's substance S (11-deoxycortisol)
Sa	<i>Staphylococcus aureus</i>
SDR	Short-chain dehydrogenase/reductase
SDS	Sodium dodecyl sulfate
SL	Strigolactone
wcw	wet cell weight

Summary

Over the past decades, *Bacillus megaterium* has been employed for the industrial production of recombinant proteins and active pharmaceutical compounds. However, the limited repertoire of promoter sequences has largely impaired the potential of *B. megaterium* for the heterologous protein expression and the rational strain design towards a more efficient biosynthesis of these compounds. In the present work, 19 innovative promoter elements of diverse promoter classes were identified based on differential gene expression analyses and characterized via β -galactosidase screening. Their activities ranged from 15% to 145% compared to the reference promoter. Selected promoters were successfully applied to establish the currently most efficient *B. megaterium* based whole-cell systems for the cholesterol oxidase mediated conversion of pregnenolone to progesterone and the 11 β -hydroxysteroid dehydrogenase mediated conversion of cortisol to cortisone. Multigram scaled steroid yields were achieved. Moreover, the novel promoters were used to establish the complex biosynthesis of C30 carotenoids. The involved enzymes in *B. megaterium* had previously been identified within this work along with the underlying biosynthetic route towards the production of the pharmaceutically relevant C30 carotenoid diaponeurosporene. In summary, the versatile range of applications demonstrated the promising potential of the novel promoters for a future use in diverse biotechnological processes.

Zusammenfassung

Bacillus megaterium wird seit Jahren für die industrielle Produktion von rekombinanten Proteinen und pharmazeutisch aktiven Substanzen eingesetzt. Aufgrund der beschränkten Auswahl an Promotorsequenzen wurde das Potenzial von *B. megaterium* für die heterologe Proteinexpression und das rationale Stammdesign allerdings weitestgehend vernachlässigt. Im Rahmen der vorliegenden Arbeit wurden 19 innovative Promotoren unterschiedlicher Promotorklassen auf Grundlage von differentiellen Genexpressionsanalysen im *B. megaterium* Stamm MS941 identifiziert und mittels β -Galaktosidase-Screening charakterisiert. Das Spektrum der Promotoraktivitäten lag im Vergleich zum Referenzpromotor zwischen 15% und 145%. Durch die Anwendung ausgewählter Promotoren konnten die zurzeit effizientesten *B. megaterium* basierten Ganzzellsysteme für den Cholesteroxidase vermittelten Umsatz von Pregnenolon zu Progesteron sowie für die 11 β -Hydroxysteroid Dehydrogenase katalysierte Produktion von Cortison etabliert werden. Die Steroidausbeuten lagen hierbei im Multigramm Bereich. Des Weiteren konnten die neuen Promotoren erfolgreich für die komplexere Biosynthese von C30-Carotinoiden angewendet werden. Die beteiligten Enzyme wurden zuvor in *B. megaterium* identifiziert und der Syntheseweg wurde bis zum pharmazeutisch interessanten C30-Carotinoid Diaponeurosporen aufgeklärt. Das breite Anwendungsspektrum der neuen Promotoren zeigt deren enormes Potenzial für eine zukünftige biotechnologische Nutzung.

1. Introduction

1.1 Microbial strain development

The conventional chemical production of fuels, industrially important bulk chemicals and high-value compounds for the pharmaceutical industry is strongly dependent on petroleum based resources. The reckless exploitation of these raw materials as well as the harsh process and reaction conditions during chemical synthesis significantly contribute to a wide range of severe problems, including environmental pollution, climate change and health risks (Yang et al., 2002). Given the progressive depletion of nonrenewable natural resources, the sustainable and environmentally friendly production of valuable chemicals by microorganisms is considered to be one of the major challenges in modern biotechnology (Chubukov et al., 2016; Gosset, 2008; Lee et al., 2011; Weusthuis et al., 2011). Natural strains of microorganisms were originally used for basic biotechnological applications such as food fermentation, but they rarely meet the requirements of the nowadays highly specialized industrial biotechnology (Ray and Joshi, 2014; Sengun and Karabiyikli, 2011; Sicard and Legras, 2011). Consequently, the demand for more efficient industrial production strains with enhanced properties regarding product formation rates and costs as well as product yields and specificities gave rise to the fascinating interdisciplinary research field of microbial strain development (Parekh et al., 2000).

In general, the strategies for microbial strain development can be conceptualized into an adaptive approach, also known as directed evolution and into an approach which is based on rational design (Zhou and Alper, 2019). Despite fundamental methodological differences, both strategies aim to overcome the same innate cellular limitations of gene regulation, signaling networks, enzyme activities and metabolic fluxes (Bailey, 1991). The adaptive approach of microbial strain development is mainly directed towards the imitation and acceleration of natural evolution by random mutagenesis, thereby generating the largest libraries of diverse phenotypes (Cobb et al., 2013). However, the screening effort for strains with improved properties is considerably high, due to the immense genetic diversity (Reetz et al., 2008). Since random mutagenesis is typically induced with highly mutagenic radiation or highly mutagenic chemicals such as ethyl methanesulfonate (EMS), information about the genetic background is not necessarily required. Furthermore, the resulting microbial strains are not classified as genetically modified and are consequently considered to be particularly compliant with the strict regulations of the food and beverage industry (Jankowicz-Cieslak et al., 2017; Sikora et al., 2011). More recently, novel strategies for continuous evolution processes were established in customized mutator strains with impaired DNA repair systems, thus leading to spontaneous genome-wide insertions during DNA replication (Greener et al., 1997). The mutation frequencies of these strains were significantly improved using DNA polymerases with impaired proofreading activity (Abe et al., 2009; Shimoda et al., 2006). A major limitation of the directed evolution approach, however, is

that not only the location of randomly distributed genome-wide mutations, but also their contribution to the improved phenotypes remain uncharacterized. This loss of information is considered to be extremely detrimental for the rapid and efficient evolution of microbial strains. Contrarily, rational strain design is based on profound knowledge about the genotype-phenotype relationships in the complex network of transcriptional, translational and metabolic control of living cells (Machado and Herrgård, 2015; McCloskey et al., 2013). The minimal prerequisite for this understanding is to have access to the genomic sequence of the desired strain. Since information about the genetic background is not available for any strain or microorganism, the lack of genomic data is considered to be a major drawback of rational strain design. However, as the number of sequenced genomes and molecular tools for genome editing continuously increases, rational strain design will become available for more microbial species (Land et al., 2015; Smith, 2017). Traditional strain design and engineering was predominantly focused on the elimination or optimization of individual enzyme activities to achieve maximal titers of a desired product, thereby disregarding potential negative effects on the overall system performance during industrial fermentation processes. A more systemic and modern approach of rational microbial strain design involves the emerging research area of systems biology. Systems biology relies primarily on comparative “omics”-driven technologies to gain profound insights into any level of the complex regulatory networks in living cells (St. John and Bomble, 2019). The knowledge collectively derived from genomic, transcriptomic, proteomic and metabolomic as well as fluxomic data is utilized for the establishment of complex computational frameworks to iteratively predict, simulate, assess and ultimately determine the best combinatorial genetic manipulations along with optimal fermentation conditions for the efficient production of a desired compound in a specific microorganism (Nielsen and Keasling, 2016). The implementation of these so called “design-build-test-learn” cycles was successfully applied to improve microbial strains with regard to product selectivity and productivity (Cao et al., 2016; Carbonell et al., 2018).

1.2 Promoter systems and their contribution to microbial strain development

From the early beginnings of recombinant protein production, multiple concepts have been developed to overcome the cellular limitations for the efficient expression of complex proteins, membrane and even toxic proteins. These concepts not only include various modifications at the DNA/RNA level such as the adaptation of rare codons, the engineering of ribosomal binding sites (RBS) or the rational design of the mRNA structure and stability but also modifications at the protein level like the coexpression of chaperons or fusions with solubilizing protein tags (Francis and Page, 2010). Altogether, these strategies were successfully applied to maximize the yields of individual proteins in many microorganisms (Kaur et al., 2018). However, a more universal approach for the optimization of heterologous protein expression is based on promoters. As important regulatory elements in the complex network of transcriptional control, promoters play a fundamental role during RNA

polymerase recruitment, the initiation of the transcription process as well the determination of the transcription efficiency (Feklistov, 2013; Ghosh et al., 2010; Saecker et al., 2011; Seshasayee et al., 2006). According to the mode of activation, promoter systems are broadly categorized into inducible, constitutive as well as growth phase dependent promoter classes, each of them possessing their own benefits but also limitations. For several decades, inducible promoter systems were preferentially used for the production of recombinant proteins, since they provide the highest flexibility with regard to temporal and quantitative control of protein expression (Briand et al., 2016; Giacalone et al., 2006). The production time and the expression level of a desired protein is simply manipulated by the amount of chemically inducing agents (e.g. trace elements, nutrients, alcohols, carbon sources or antibiotics) (Hemmati and Basu, 2015; Rodríguez-García et al., 2005; Rouch and Brown, 1997; Weinhandl et al., 2014) or the extent of environmental influences (like osmotic stress, temperature differences or light exposure) (Dattananda et al., 1991; Tabor et al., 2011; Taylor et al., 1984). The most commonly used inducible promoter systems are of bacterial origin and comprise promoters of different sugar metabolizing operons such as the lactose or arabinose inducible promoter system from *E. coli* or the xylose inducible promoter system from *B. megaterium* (Greenfield et al., 1978; Hopkins, 1974; Rygus et al., 1991). These promoters were combined with expression systems of bacteriophages for the establishment of the popular and widely used T7 or T5 promoter systems (Brunner and Bujard, 1987; Gamer et al., 2009). The application of these inducible promoters was shown to be especially beneficial for the high yield expression of many challenging proteins including native membrane proteins and toxic proteins (Baneyx, 1999; Joseph et al., 2015; Montigny et al., 2004). However, inducible promoter systems reach their limit when high costs of conventional inducing agents or toxic effects, particularly of the commonly used allolactose analog IPTG, impair the application in industrially scaled processes (Nocadello and Swennen, 2012). Contrarily, constitutive promoters are not subject to the restrictions of inducible promoter systems since the addition of expensive or toxic inducing agents is not necessary (Redden et al., 2015). Though recombinant protein production occurs permanently and can certainly lead to the accumulation of large quantities of protein at early stages of bacterial growth, the static protein formation rates of constitutive promoters do not provide much flexibility in terms of tuning protein expression. As a result, the application of constitutive promoters for the heterologous expression of complex proteins is often associated with considerable formation of inclusion bodies, with growth inhibition or even cell death caused by severe protein toxicity (Donnelly et al., 2001). Growth phase dependent promoters, however, serve as important link between inducible and constitutive expression systems since they combine the useful properties of both systems. They are not dependent on the addition of inducing agents but nonetheless allow sufficient flexibility concerning the time point of protein induction as a result of growth phase specific transcription as well as protein expression patterns (Di Gennaro et al., 2008). In recent years, many growth phase dependent promoters were identified and characterized in a wide variety of different microorganisms and thus became increasingly important to the biotechnological industry, particularly for the efficient

expression of recombinant proteins during late microbial growth phases (Guan et al., 2015; Solera et al., 2004; Studier, 2005).

Apart from their significance for the optimization of recombinant protein yields, different promoter systems have also attracted considerable interest for the application of metabolic pathway engineering in the context of rational strain design (Guiziou et al., 2016; Markley et al., 2015; Vogl et al., 2018). Instead of maximizing the expression levels of single proteins, the rational approach of microbial strain design is rather focused on harmonizing protein expression levels in complex biosynthetic pathways that often involve the coordinated action of multi-enzyme cascades (Farnberger et al., 2017; Hold et al., 2016). The main challenge with increasing reaction network complexity, however, is to gain operational control over simultaneously expressed proteins because the resulting metabolic burden of imbalanced protein expression was shown to impair the overall catalytic performance of microbial biocatalysts remarkably (Wachtmeister and Rother, 2016). In order to meet these requirements, a broad set of promoters with diverse activities is considered fundamental for the dynamic control of protein expression levels in tailored microbial cell factories (Chen et al., 2018; Jayaraman et al., 2018; Müller and Stelling, 2009; Zhang and Zhou, 2014). Promoters with different induction time points as well as transcription efficiencies were deliberately used to exploit native as well as engineered biosynthetic pathways in many microorganisms (Brockman and Prather, 2015a; Gupta et al., 2017; Nevoigt, 2008). The adaption and fine tuning of protein levels in linear multi-enzyme cascades was demonstrated to have a beneficial effect on cell viability and product formation rates (Gardner et al., 2000; Hwang et al., 2018). Moreover, side product formation was significantly reduced (Taniguchi et al., 2017). Increasing numbers of scientific studies furthermore report the redirection of metabolic fluxes in competing biosynthetic pathways, thereby maximizing the yield of desired and valorized natural products. In the course of these studies, particularly weak promoters have gained a prominent role as important tools for the development of knockdown phenotypes of genes that are predominantly located at crucial branch points within the central carbon metabolism (Brockman and Prather, 2015b; Scalcinati et al., 2012). Ultimately, the selection of suitable promoter systems has to be compliant with the experimental approach. Consequently, the availability of a broad promoter set with a wide range of activities is considered indispensable not only for the efficient expression of single proteins but also for the efficient manipulation and engineering of whole metabolic pathways in the emerging field of rational microbial strain design.

1.3 *Bacillus megaterium*

Bacillus megaterium was first described in 1884 by Anton De Bary and belongs to the group of non-pathogenic Gram-positive bacteria (De Bary, 1884). The rod-shaped, aerobic, endospore-forming bacterium is predominantly found in the soil, but is also able to colonize various alternative and unexpected habitats such as seawater, sediment and honey (Vary et al., 2007). Consequently, it is not surprising that *B. megaterium* was observed to grow in minimal media on over 62 different carbon sources, which allows a simple and inexpensive cultivation (Vary, 1994). Due to its large vegetative cell size of up to 4 μm in length and 1.5 μm in diameter, *B. megaterium* has originally been used as a model organism for intensive studies on cytoplasmic membrane and bacterial cell wall biosynthesis, on protein localization, as well as on spore formation and spore structure (Vary, 1992).

However, over the last decades *B. megaterium* has become increasingly important to the biotechnological industry, due to its simple cultivation, high capacity for protein production and plasmid stability (Vary et al., 2007). Unlike other bacteria for recombinant protein expression, *B. megaterium* has been granted with the GRAS (generally recognized as safe) status since it does not produce any endotoxins (Wang et al., 2006). Moreover, the absence of alkaline proteases is considered to be particularly attractive for the downstream processing and recovery of recombinantly expressed and secreted proteins (Priest, 1977). The successful production of bulk enzymes with useful applications in the biotechnological industry was reported for several amylases (Hebeda et al., 1988; Metz et al., 1988), penicillin G acylase (Martín et al., 1995; Suga et al., 1990), glucose dehydrogenase (Pauly and Pfeleiderer, 1975), keratinase (Radha and Gunasekaran, 2007) and xylanase (Sindhu et al., 2006). Apart from being relevant for the industrial production of these recombinant proteins, *B. megaterium* has recently gained considerable interest as microbial cell factory for the biosynthesis of secondary metabolites and biotransformation of pharmaceutical compounds. The elucidation of enzymes involved in these complex biosynthetic pathways and biotransformation systems was significantly accelerated by the publication of the whole genome sequences of the industrially important *B. megaterium* strains DSM319, QMB1551 and WSH-002 in 2011 (Eppinger et al., 2011; Liu et al., 2011). Since then, various novel enzymes for potential biotechnological and pharmaceutical applications have been identified and characterized in *B. megaterium*. The most prominent enzymes certainly belong to the group of cytochromes P450 and auxiliary redox proteins, which were used for the conversion of natural as well as synthetic steroid hormones (Kiss et al., 2015; Putkaradze et al., 2019, 2017a; Schmitz et al., 2018), valorization of vitamins (Abdulmughni et al., 2017; Biedendieck et al., 2010; Ehrhardt et al., 2016), terpenoids (Bleif et al., 2012, 2011; Putkaradze et al., 2017b) as well as other pharmaceutical compounds (Milhim et al., 2016). Likewise, the production of oxetanocin, a potent inhibitor of viral infections was described in *B. megaterium* (Morita et al., 1999; Tseng et al., 1991). Even the biosynthesis of polyhydroxyalkanoates (PHAs), a promising starting material for the sustainable manufacturing of biodegradable plastics, was reported to be very efficient (McCool, 1996;

Rodríguez-Contreras et al., 2013). Particularly, these PHAs were attributed for the intracellular accumulation of hydrophobic organic compounds such as cholesterol, which makes *B. megaterium* an ideal microbial platform for the biotransformation of steroids and steroidal drugs (Gerber et al., 2015).

1.4 Biosynthesis of terpenoids

The biosynthesis of terpenoids is a complex process which involves the coordinated action of multiple enzymes. In general, the backbone structure of terpenoids is derived from the common isoprenoid precursor molecules isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP) (Vattekkatte et al., 2018). Depending on the organism, the intracellular pool of IDP and DMADP can be supplied by two different biosynthetic pathways, the mevalonate (MEV) pathway or the non-mevalonate pathway, often referred as methylerythritol phosphate (MEP) pathway (Boucher and Doolittle, 2000). The MEV-pathway, and variations thereof, is predominantly found in archaea, fungi and higher eukaryotes (Zhao et al., 2013), while the MEP-pathway is mainly distributed among most bacteria and parasitic protozoa (Banerjee and Sharkey, 2014; Guggisberg et al., 2014). Higher plants are considered a special case, since both pathways are present (Kuzuyama and Seto, 2012). The initial biosynthetic steps of the MEV-pathway comprise the successive condensation of three molecules of acetyl-CoA (**9**) to form 3-hydroxy-3-methyl-glutaryl-CoA (HMG-CoA). This condensation cascade first involves the activity of the acetyl-CoA-acetyltransferase (Acat) to yield acetoacetyl-CoA (**10**), which is then further converted to HMG-CoA (**11**) by the HMG-CoA synthase (Hmgs). The subsequent reduction of HMG-CoA to (*R*)-mevalonate (**12**) was demonstrated to be the rate-limiting step within the MEV-pathway and is catalyzed by the HMG-CoA-reductase (Hmgr). This reduction is followed by a phosphorylation cascade, starting with the formation of (*R*)-mevalonate-5-phosphate (**13**), which is then further converted to (*R*)-mevalonate-5-diphosphate (**14**). The involved mevalonate kinase (Mek) and phosphomevalonate kinase (Pmk) are both dependent on adenosine triphosphate (ATP) consumption. The final decarboxylation reaction of (*R*)-mevalonate-5-diphosphate is catalyzed by the diphosphomevalonate decarboxylase (Mdc) and results in the formation of isopentenyl diphosphate (**16**). Intramolecular rearrangements, performed by the IDP-isomerase (Idi), allow the conversion of IDP to dimethylallyl diphosphate (**15**) (Jain, 2014; McGarvey, 1995). The initial reaction step of the MEP-pathway comprises the formation of 1-deoxy-D-xylulose 5-phosphate (**3**) and is catalyzed by the DXP synthase (Dxs) which covalently bonds one molecule of glyceraldehydes-3-phosphate (**1**) to one molecule pyruvate (**2**). Carbon dioxide (CO₂) is released as a byproduct. In the next step, the DXP reductase (Dxr) catalyzes the intramolecular rearrangement and concomitant reduction of DXP to 2C-methyl-D-erythritol-4-phosphate (**4**). This reduction step is followed by the addition of cytidine diphosphate (CDP) to form 4-diphosphocytidyl-2C-methyl-D-erythritol (**5**) and involves the activity of the MEP cytidyltransferase (IspD). The subsequent phosphorylation of CDP-ME to 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate (**6**) is performed by the

corresponding kinase (IspE). The cleavage of cytidine monophosphate (CMP) then leads to the cyclization of CDP-MEP yielding 2-C-methyl-D-erythritol-2,4-cyclodiphosphate (7). This cyclization reaction is catalyzed by the corresponding MEcDP synthase (IspF). The penultimate step of the MEP-pathway is the reduction of MEcDP to (*E*)-4-Hydroxy-3-methyl-but-2-enyl-diphosphate (8) and involves the activity of the HMB-DP synthase (IspG). Finally, HMB-DP is converted by the isopentenyl diphosphate and dimethylallyl diphosphate synthase (IspH) to IDP (16) as well as DMADP (15) (Lichtenthaler, 1999; Rohmer, 1999).

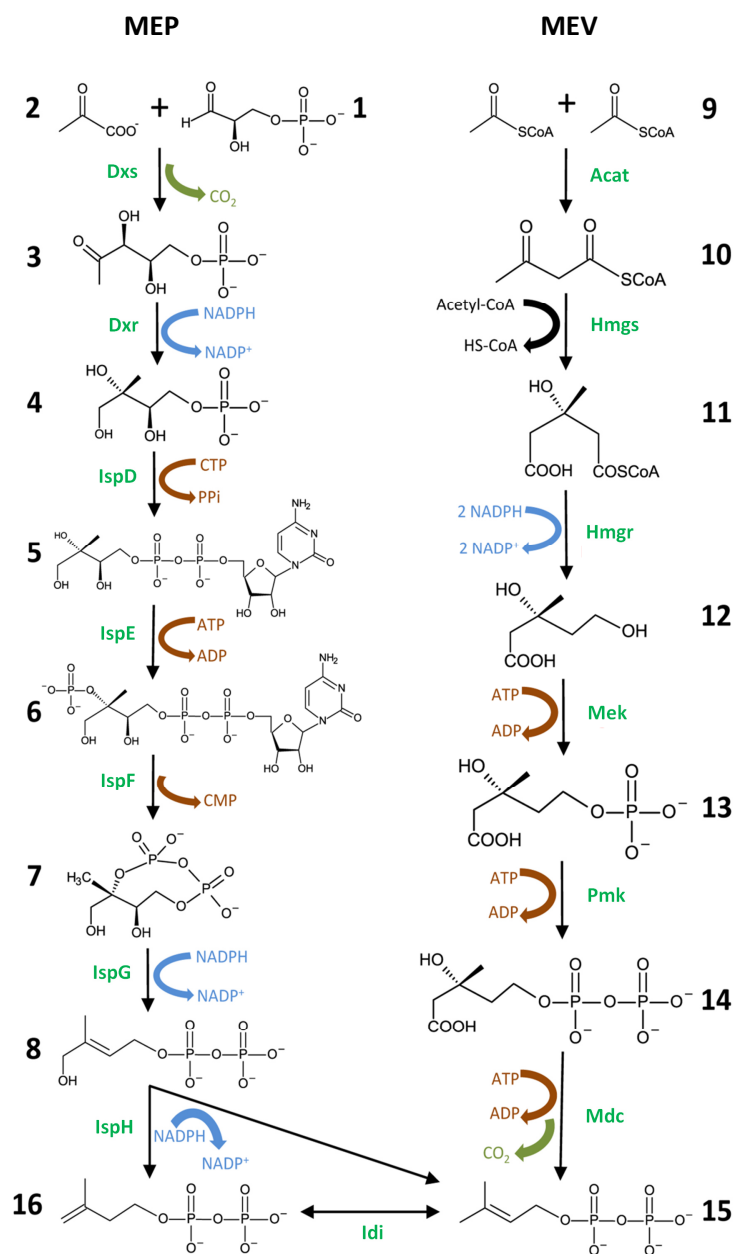


Figure 1.1. Comparison of the biosynthetic steps of the methylerythritol phosphate (MEP) pathway (left) and the mevalonate (MEV) pathway (right). Adopted and modified from (Partow et al., 2012).

Consecutive condensation reactions of IDP and DMADP lead to the formation of terpenoid structures with varying chain lengths, such as the geranyl diphosphate (GDP) derived monoterpenoids, farnesyl diphosphate (FDP) derived sesqui- and triterpenoids or the geranylgeranyl diphosphate (GGDP) derived di- and tetraterpenoids (see Figure 1.2).

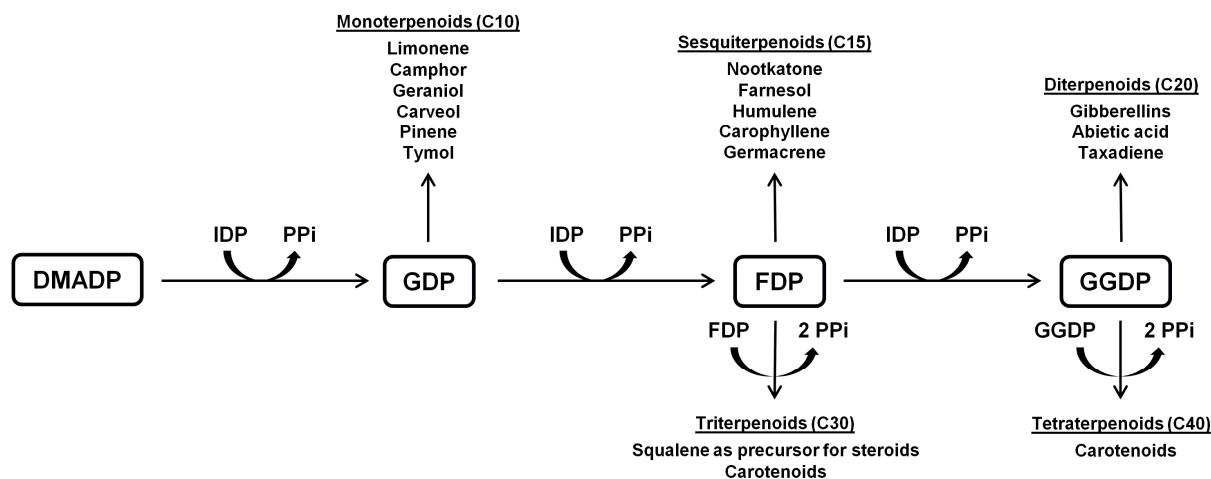


Figure 1.2. Schematic presentation of the terpenoid biosynthesis.

1.5 Biosynthesis and biotransformation of steroids

Natural steroids and steroid-like compounds occur in almost every kingdom of life where they have key roles in the control of many different physiological processes. The most prominent functions of steroids in mammals, including humans, comprise the stabilization of membranes (sterols) (Raffy and Teissié, 1999), the regulation of the salt-water balance (mineralocorticoids) (Funder et al., 1997) and immune response (glucocorticoids) (Munck et al., 1984) as well as the development of sexual characteristics and behavior (progestogens, androgens and estrogens) (Wierman, 2007). The biological functions of steroids are mainly dependent on the oxidation state of the steroidal core structure as well as on the presence, position and nature of attached functional groups (Lednicer, 2011). For this reason, the gonane core of steroids is considered an important lead structure for the development of novel steroid derived drugs with versatile therapeutic applications (Tong and Dong, 2009). The global market value of steroidal active pharmaceuticals is second just behind antibiotics and projected to exceed 17 billion US\$ by the end of 2025 (<https://industrialjournalism.com>). Consequently, the commercial interest for the production of natural steroid hormones and steroid derived drugs is accordingly high. However, the biosynthesis of steroids in general and steroid hormones in the human body is a complex process, which is enabled by the chronologically and spatially coordinated action of multiple enzymes (see Figure 1.3), such as hydroxysteroid dehydrogenases (HSDs) and cytochromes P450 (CYPs) (Ghayee and Auchus, 2007).

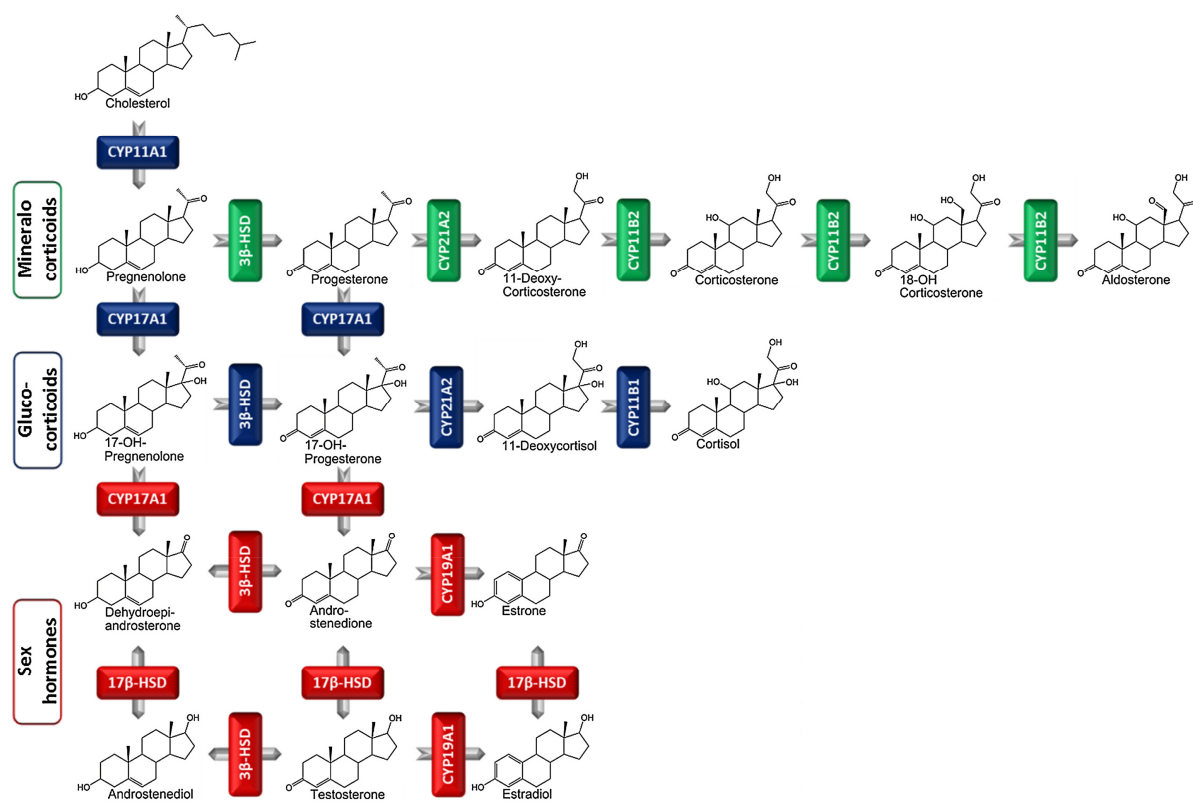


Figure 1.3. Schematic presentation of the steroid hormone biosynthesis in the human body. Figure was taken and modified from (Neunzig et al., 2017).

The potential of microorganisms as biocatalysts for the biotransformation of steroids became obvious right after initial attempts for a total chemosynthesis of the important glucocorticoid cortisone (Peterson and Murray, 1952; Sarett, 1946). While total chemosynthesis was typically associated with numerous consecutive reactions under harsh conditions, low selectivity and yields as well as high economic costs, the application of natural microbial strains lead to a significant reduction of chemical steps and production costs, thereby contributing to establish more economically viable processes for a wide variety of steroid hormones (Egorova et al., 2002; El-Kadi and Eman Mostafa, 2004; Fernández-Cabezón et al., 2017; Hakki et al., 2008; Hannemann et al., 2007; Kiss et al., 2015). Given the increasing number of sequenced genomes and molecular tools for genome editing, nowadays, the main challenges for the design and development of microbial strains with improved properties for the production of pharmaceutically important steroidal precursors (synthons) mainly involve the restriction of endogenous steroid degradation pathways and the *de novo* biosynthesis of steroids from cheap and sustainable feedstocks such as glucose, glycerol or second generation biomass (Fernández-Cabezón et al., 2018). Steroid degradation in microorganisms often derives from undesired modifications like hydroxylation and/or dehydrogenation of the gonane ring structure. The successful elimination of these interfering enzyme activities by rational gene deletion has significantly contributed to a better understanding of the microbial steroid catabolism (Gerber et al., 2015; Yao et al., 2014). However, extensive studies on steroid anabolism played an equally important role for the

development of improved microbial steroid transformers. Research on the *de novo* biosynthesis of natural steroid hormones in yeasts was predominantly focused on the optimization of the supply with isoprenoid building blocks for the efficient assembly of squalene, the triterpenoid precursor for the formation of the gonane structure of sterols (Duport et al., 1998; Szczebara et al., 2003). Significant improvements were associated with the engineering of the mevalonate pathway for the biosynthesis of the universal terpenoid precursors isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP) (Ma et al., 2018). The knowledge collectively derived from these prime examples of microbial steroid transformation has paved the way for the engineering of superior strains with enhanced properties for the industrial production of steroid derived drugs.

1.6 Structure, function and biosynthesis of carotenoids

In 1831, Heinrich W. F. Wackenroder first described the isolation of a yellow pigment from carrots and consequently named it "carotin" (Wackenroder, 1831). This pigment was later attributed to be eponymous for one of the most diverse and widespread class of naturally occurring pigments, the carotenoids (Mortensen, 2006). Since then, the number of identified carotenoids has continuously increased to a total of over 1100 structures which are ubiquitously distributed among all domains of life (Yabuzaki, 2017). The colorful world of carotenoids ranges from yellow through orange and red to purple shades (Britton, 1995). The main cause for this wide range of colors is the differing number and location of conjugated double bonds within the carotenoid backbone. The individual composition of alternating double bonds confers unique characteristics to any carotenoid with regard to the absorption of visible light and susceptibility towards electrophilic reagents as well as aggressive radicals (Britton et al., 2008).

Carotenoids are primarily differentiated by the number of carbon atoms in the isoprenoid backbone, which is primarily dependent on the number of the isoprenoid building blocks IDP and DMADP. Both are substrates for the biosynthesis of longer-chained prenyl diphosphates, such as the C10 isoprenoid geranyl diphosphate (GDP), the C15 isoprenoid farnesyl diphosphate (FDP), the C20 isoprenoid geranylgeranyl diphosphate (GGDP) and even higher isoprenoid structures (Tarshis et al., 1996). The class of enzymes which is involved in this process is generally assigned as prenyl diphosphate transferases or prenyl diphosphate synthases, even though most representatives of this enzyme class are more precisely named according to the end product of the prenyl transferase reaction as GDP, FDP or GGDP synthases, etc. (Burke et al., 1999). As illustrated in Figure 1.4, the initial step of the carotenoid biosynthesis (exemplified for β -carotene derivatives) is actually catalyzed by carotenoid synthases (e.g. PSY, phytoene synthase) and comprises the head to head condensation of two molecules of prenyl diphosphates (here GGDP, geranylgeranyl diphosphate) resulting in the formation of the basic, linear carotenoid backbone structure (Moise et al., 2014). The carotenoid backbone is

considered symmetric if both prenyl diphosphates have equal chain lengths, while the condensation of prenyl diphosphates with varying chain lengths leads to the formation of asymmetric carotenoid structures (Perez-Fons et al., 2011). The following reactions of carotenoid biosynthesis involve the activity of carotenoid desaturases (e.g. PDS, phytoene desaturase), which are responsible for the sequential introduction of additional double bonds in the carotenoid backbone, thereby contributing to the diversification of carotenoid structures as well as properties. Carotenoid structures may be further modified through the action of carotenoid cyclases (e.g. LYC, lycopene cyclase) (Mialoundama et al., 2010), oxygenases (Lobo et al., 2012), glycosyl transferases (Pelz et al., 2005) or carotenoid cleavage enzymes (Schwartz et al., 2001).

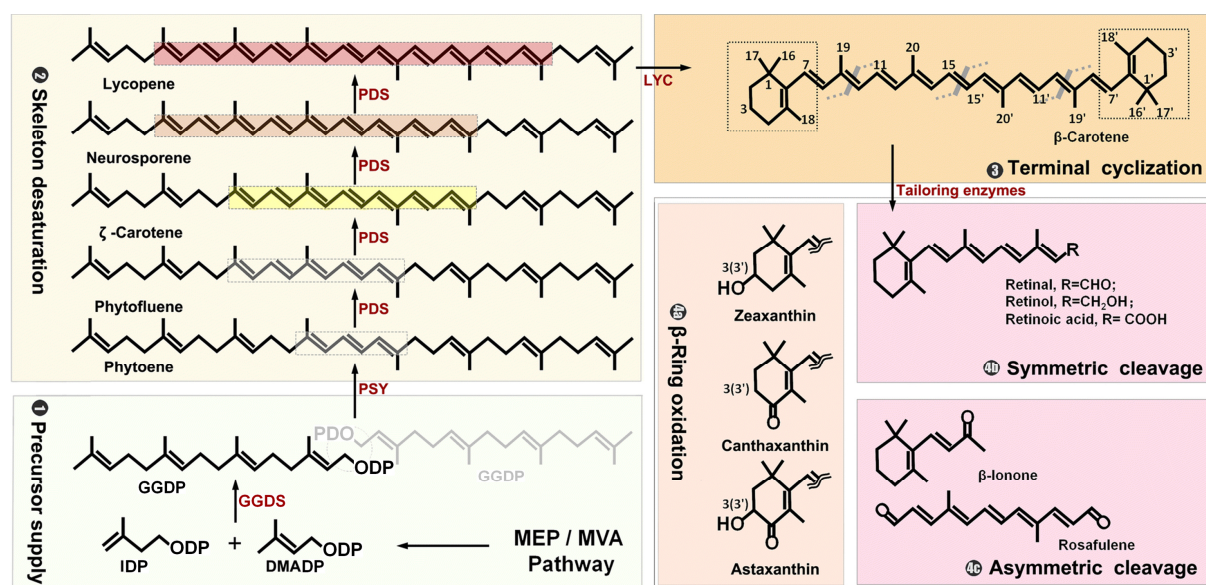


Figure 1.4. Biosynthetic pathway of the common C40 carotenoid β -carotene and its derivatives. Figure was taken and modified from (Wang et al., 2019) .

Among the plethora of all carotenoids, C40 carotenoids which are predominantly found in plants, are considered to be the most abundant class of carotenoid structures in nature. There are also carotenoid structures with noticeably shorter or extended isoprenoid backbones. While C30 carotenoids were described as main natural colorants in the group of pigmented bacilli (Köcher et al., 2009; Perez-Fons et al., 2011; Steiger et al., 2012; Takaichi et al., 1997; Tao et al., 2005; Wieland et al., 1994), the biosynthesis of previously unknown carotenoid structures with C35, C45, C50 and even C60 isoprenoid backbones and novel properties was recently enabled by directed evolution of several carotenoid synthases (Li et al., 2019; Tobias and Arnold, 2006). Given the fact that carotenoids are usually synthesized from C5 isoprene units, the number of carbon atoms in the carotenoid backbone seems to be limited to an interval of 5. However, there are many examples where the cleavage of the carotenoid backbone results in an expanded spectrum of carotenoid structures with irregular numbers of carbon atoms (Auldridge et al., 2006; Fleischmann and Zorn, 2008). These carotenoid structures are mainly classified as carotenoid derived aroma compounds comprising various apocarotenoids and

norisoprenoids (Winterhalter, 1996). Further differentiation criteria for carotenoids are based on the formation of aliphatic or cyclic end groups (Kirti et al., 2014). The most commonly known representative of aliphatic carotenoids is the linear C40 carotenoid lycopene, which is responsible for the red color of several fruits and vegetables, particularly of tomatoes (Khoo et al., 2011). Monocyclic carotenoids which possess both an aliphatic as well as a cyclic end group are less commonly distributed in nature, since they are usually considered as intermediates during the biosynthesis of bicyclic carotenoids, like β -carotene. However, accumulation of monocyclic carotenoids was observed in individual microorganisms (An et al., 1999; Iniesta et al., 2008; Takaichi et al., 1995, 1990). Apart from the classification according to the end group, carotenoids are divided in more detail into carotenes, exclusively composed of carbon as well as hydrogen molecules and xanthophylls, which are functionalized with at least one atom of oxygen originating from hydroxy, carboxy, epoxy, keto or ether groups (Breithaupt and Bamedi, 2001). Even glycosylated carotenoids and carotenoids esterified with fatty acids have been described (Takaichi and Mochimaru, 2007).

The physiological functions of carotenoids are as diverse as their structures but can principally be categorized into photosynthesis, photoprotection and nutrition (Kirti et al., 2014). The majority of carotenoids is found in photosynthetic organisms like plants, algae, euglena and bacteria, where they exert dual functions during photosynthesis. As key components of the chlorophyll containing light-harvesting complexes, carotenoids primarily contribute to expand the absorption spectrum of chlorophyll and support the light induced resonance energy transfer towards the photosynthetic reaction center (Frank and Cogdell, 1996). Moreover, carotenoids protect the photosystems against extreme light exposure via non-photochemical quenching (NPQ), a process in which excessive excitation energy is harmlessly converted to heat, thereby minimizing photooxidative damage (Müller et al., 2001). Apart from their crucial role in photosynthesis and photoprotection, carotenoids also serve as important precursors for the biosynthesis of essential plant hormones and other signaling molecules, such as abscisic acid, strigolactones and rose ketones (Ruiz-Sola and Rodríguez-Concepción, 2012). Abscisic acid (ABA), for example, is generated as a result of carotenoid cleavage by members of the 9-cis-epoxycarotenoid dioxygenase family (Neuman et al., 2014; North et al., 2007). As a key mediator of physiological responses to environmental stress conditions, including high temperatures, aridity and salinity, ABA controls many substantial processes during plant development and plant growth, such as seed dormancy, germination, maturation and stomatal regulation (Kumar et al., 2013). A recently identified class of phytohormones, the strigolactones (SLs), is also proposed to derive from carotenoid metabolites (Chen et al., 2009). Originally identified in parasitic plant seeds, SLs were later characterized as exogenous stimulants of seed germination as well as chemoattractants for symbiotic mycorrhizal fungi (Besserer et al., 2006). SLs, furthermore, participate as endogenous signaling molecules in a variety of fundamental processes, most notably root formation and shoot branching (Waldie et al., 2014). Other degradation products of carotenoids are assigned to the family of rose ketones, which comprises closely related volatile compounds like

ionons, damascones and damascenones, main components of various essential oils. Due to their fragrant properties, rose ketones are not only attractive for pollinating insects or seed dispersing animals, but also highly demanded chemicals for the fragrance industry (Winterhalter and Rouseff, 2001).

Although non-photosynthetic organisms are lacking corresponding photosystems, they can similarly benefit from the photoprotective as well as antioxidant properties of carotenoids. This protective role was demonstrated e.g. for a wide variety of endospore forming bacteria (Duc et al., 2006; Perez-Fons et al., 2011). The biosynthesis of carotenoids during different stages of spore formation was associated with significantly increased survival rates when the bacterial cells were exposed to harsh environmental conditions such as heat and extreme light irradiation (Moeller et al., 2005). As a result of these studies, the accumulation of carotenoids in the cytoplasmic membrane and spore envelope was assumed not only to prevent reactive oxygen species (ROS) and aggressive radicals from damaging integral constituents of biological membranes but also the genomic DNA from suffering harmful or even lethal mutations (Young and Lowe, 2018). Furthermore the presence of carotenoids was proposed to play an essential role in the modulation of membrane fluidity, particularly as stabilizing components in extreme halophilic and thermophilic bacteria (Hara et al., 1999; Hoshino et al., 1994; Yatsunami et al., 2014; Yokoyama et al., 1995). Higher order animals, including humans, are not able to produce carotenoids and are consequently dependent on dietary uptake, particularly of carotenoids with provitamin A function, like α -carotene, β -carotene or β -cryptoxanthin. After metabolic activation, vitamin A (retinol) as well as its cognate forms retinal and retinoic acid are involved in many physiological processes including cell growth, cell differentiation, immune modulation, reproduction and most prominently in the visual process (Chew and Park, 2004; McDevitt et al., 2005; Zile and Cullum, 1983). Furthermore, carotenoids significantly contribute to various health benefits for humans. This has been demonstrated in various scientific studies and is strongly associated with the extraordinary antioxidant potential of carotenoids (Britton et al., 2008). As molecules with hydrophobic character, carotenoids are primarily localized in the lipid bilayer of cytoplasmic membrane and subcellular membrane systems where these antioxidant properties are fully exploited to protect unsaturated fatty acids as well as lipoproteins against ROS and free radicals originating from environmental stress, cellular respiration and other metabolic processes (Sharma et al., 2012). As a result, carotenoids were unsurprisingly shown to exhibit inhibitory effects on the progress of severe diseases such as cancer, cardiovascular and neurodegenerative diseases (Gerster, 1993; Miyake et al., 2011; Obulesu et al., 2011; Rao and Agarwal, 2000).

1.7 Aim of this work

Over the past decades, *Bacillus megaterium* has been widely used for the recombinant production of individual proteins with industrial importance. However, recent developments in microbial strain design increasingly favored the establishment of *B. megaterium* as microbial cell factory for the efficient biotransformation of pharmaceutically active compounds. These compounds include vitamins, terpenoids, cholesterol lowering drugs and steroid-derived drugs. Except for the production of vitamin B12, the potential of *B. megaterium* for the de novo biosynthesis of bioactive compounds from sustainable feedstocks has been largely ignored (Biedendieck et al., 2010; Brey et al., 1986; Moore et al., 2014). The underlying complex reaction networks are usually dependent on the coordinated action of multi-enzyme cascades that necessarily require harmonized protein expression levels. Although promoters are considered as key regulators for harmonized gene expression, the repertoire of useful promoter elements in *B. megaterium* has almost exclusively been limited to the xylose inducible promoter system (Bäumchen et al., 2007; Burger et al., 2003; Korneli et al., 2013). The lack of alternative promoter systems combined with the lack of comprehensive knowledge on biosynthetic pathways with industrial importance has so far evidently impaired the development of *B. megaterium* as efficient microbial cell factory.

In order to overcome these limitations, the present work aims on the characterization of novel promoter systems along with the elucidation of novel biosynthetic pathways in the industrially important *B. megaterium* strain MS941 (DSM319 $\Delta nprM$) (Wittchen and Meinhardt, 1995). Based on the published genome sequence of the *B. megaterium* strain DSM319 (Eppinger et al., 2011), the study will address the identification of inducible, constitutive as well as growth phase dependent promoter classes, thus providing an expanded set of promoters with a broad range of promoter activities. The collaboration with OakLabs GmbH in Berlin is used for the design and validation of microarray chips that allow the genome-wide analyses of differential gene expression profiles for the identification of novel promoters. The activity of selected promoter candidates is subsequently evaluated in a β -galactosidase based reporter gene assay. The collaboration with Metabolomic Discoveries in Potsdam is used to perform a focused metabolite profiling for the identification of novel bioactive compounds and intermediates thereof. The data collectively derived from transcriptomic and metabolomic analyses of *B. megaterium* MS941 are exploited to develop *B. megaterium* into a more efficient microbial cell factory for the transformation and biosynthesis of valuable bioactive compounds.

2. Scientific articles

2.1 (Hartz et al., 2019)

Expanding the promoter toolbox of *Bacillus megaterium*.

Hartz, P., Mattes, C., Schad, M., Bernhardt, R and Hannemann, F.

Journal of Biotechnology. 2019 Jan; 294:38-44

DOI: 10.1016/j.jbiotec.2019.01.018

Reprinted with permission of Journal of Biotechnology. All rights reserved.



Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Expanding the promoter toolbox of *Bacillus megaterium*

Philip Hartz^a, Carsten Mattes^a, Martina Schäd^b, Rita Bernhardt^a, Frank Hannemann^{a,*}

^a Department of Biochemistry, Campus B2.2, 66123, Saarland University, Saarbrücken, Germany

^b OakLabs GmbH, Neuendorfstraße 16b, 16761 Berlin/Hennigsdorf, Germany



ARTICLE INFO

Keywords:

Bacillus megaterium
Genome-wide microarray analysis
Promoter screening
High-level protein expression
Steroid conversion

ABSTRACT

Over the past decades, *Bacillus megaterium* has gained significant interest in the biotechnological industry due to its high capacity for protein production. Although many proteins have been expressed efficiently using the optimized xylose inducible system so far, there is a considerable demand for novel promoters with varying activities, particularly for the adjustment of protein levels in multi-enzyme cascades.

Genome-wide microarray analyses of the industrially important *B. megaterium* strain MS941 were applied to identify constitutive and growth phase dependent promoters for the expression of heterologous proteins from the early exponential to the early stationary phase of bacterial growth. Fifteen putative promoter elements were selected based on differential gene expression profiles and signal intensities of the generated microarray data. The corresponding promoter activities were evaluated in *B. megaterium* via β -galactosidase screening. β -Galactosidase expression levels ranged from 15% to 130% compared to the optimized xylose inducible promoter. Apart from these constitutive promoters we also identified and characterized novel inducible promoters, which were regulated by the addition of arabinose, galactose and the commonly used allolactose analog IPTG. The potential application of the identified promoters for biotechnologically relevant processes was demonstrated by overexpression of the cholesterol oxidase II from *Brevibacterium sterolicum*, thus obtaining product yields of up to 1.13 g/l/d.

The provided toolbox of novel promoters offers versatile promoter strengths and will significantly contribute to harmonize protein expression in synthetic metabolic pathways, thereby pushing forward the engineering of *B. megaterium* as microbial cell factory for the biosynthesis and conversion of valuable compounds.

1. Introduction

Gene expression levels in bacteria are significantly affected by the dynamic range of promoters which constitute important genomic regulatory elements in the complex framework of transcriptional control (Seshasayee et al., 2006). Common features of bacterial promoters, like the -35 and -10 regions, along with binding sites for associated transcription factors, were shown to be crucial for the process of RNA polymerase recruitment and initiation of transcription (Feklistov, 2013; Ghosh et al., 2010; Saecker et al., 2011). The well conserved DNA sequences of these promoter features were subject of numerous studies for the prediction, identification as well as characterization of versatile promoters in many different microorganisms (de Jong et al., 2012; Fickett and Hatzigeorgiou, 1997; Harley and Reynolds, 1987; Ishii, 2001; Lisser and Margalit, 1993; MacLellan et al., 2006). Based on the mode of activation, promoters can be assigned to three different classes

including inducible, growth phase dependent and constitutive ones. A major prerequisite for the successful application in the field of recombinant protein production is that these promoters have to meet the two most important criteria of industrial biotechnology which are reliability as well as cost-effective production of heterologous proteins and therapeutic agents (Panahi et al., 2014). For several decades, inducible expression systems were preferred for the production of heterologous proteins because they give more flexibility compared with constitutive expression systems in terms of tuning protein expression (Briand et al., 2016; Giacalone et al., 2006). Particularly, recombinant production of toxic proteins and membrane proteins was significantly improved using inducible promoters (Baneyx, 1999; Joseph et al., 2015; Montigny et al., 2004). However, the high costs of conventional inducing agents, along with toxic effects, particularly of IPTG, ultimately limit their application in industrially scaled processes (Nocadello and Swennen, 2012). Thus, growth phase specific and constitutive

Abbreviations: BCO2, cholesterol oxidase 2; DOC, 21-hydroxyprogesterone; HSD, hydroxysteroid dehydrogenase; LacZ, β -galactosidase; ONPG, o-nitrophenyl β -D-galactopyranoside; Upp, uracil phosphoribosyltransferase; X-gal, 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

* Corresponding author.

E-mail address: f.hannemann@mx.uni-saarland.de (F. Hannemann).

<https://doi.org/10.1016/j.jbiotec.2019.01.018>

Received 9 November 2018; Received in revised form 18 January 2019; Accepted 22 January 2019

Available online 13 February 2019

0168-1656/ © 2019 Elsevier B.V. All rights reserved.

expression systems are considered to be a suitable alternative to inducible expression systems (Di Gennaro et al., 2008; Guan et al., 2015; Studier, 2005). While constitutive promoters allow the efficient expression of a majority of proteins, they are evidently less applicable for the production of toxic proteins. Growth phase specific promoters combine the beneficial properties of inducible and constitutive promoters, since they allow the initiation of heterologous protein expression without the need of expensive inducers at any time of bacterial growth, thus representing an economical and efficient expression system for industrially scaled fermentation processes (Guan et al., 2016).

While the library of regulatory elements provides a variety of inducible, growth phase specific as well as constitutive promoters for the heterologous protein expression in diverse bacteria, the selection of useful promoters in *B. megaterium* is almost exclusively limited to the xylose inducible promoter of the xylose utilization operon (Bäumchen et al., 2007; Burger et al., 2003; Korneli et al., 2013). Originally discovered and characterized by Rygus and Hillen in 1991, this xylose inducible promoter was later systematically modified and optimized for a high-yield protein production in *B. megaterium* (Rygus and Hillen, 1991; Stammen et al., 2010). This non-pathogenic, Gram-positive bacterium has become increasingly important to the biotechnological industry, due to its simple cultivation, high capacity for protein production and plasmid stability (Vary et al., 2007). These beneficial properties significantly contributed to the expansion of its range of application from the heterologous expression of single proteins towards the biosynthesis and bioconversion of complex secondary metabolites (Gerber et al., 2016). These metabolites include a variety of natural and synthetic steroid hormones (Ehrhardt et al., 2016b; Gerber et al., 2015; Kiss et al., 2015), vitamin D3 (Abdulmughni et al., 2017; Ehrhardt et al., 2016a) and vitamin B12 (Biedendieck et al., 2010; Moore et al., 2014; Wolf and Brey, 1986) as well as oxetanocin, a potent inhibitor of viral infections (Shimada et al., 1986). Moreover, the publication of the whole genome sequences of the industrially important *B. megaterium* strains DSM319, QMB1551 and WSH-002 (Eppinger et al., 2011; Liu et al., 2011) recently enabled the identification and characterization of novel enzymes for potential biotechnological and pharmaceutical applications (Brill et al., 2014; Hartz et al., 2018; Milhim et al., 2016a, 2016b). So far, the lack of a versatile promoter library has impaired the rational design of *B. megaterium* as an efficient microbial cell factory for the biosynthesis and conversion of these valuable compounds, since the complex biosynthetic pathways involve multi-enzyme cascades and rely heavily on well balanced protein expression levels. For this reason, there is an increasing demand of novel promoters that offer an expanded range of promoter activities and thus enable the possibility of dynamic protein expression in *B. megaterium*.

In this work, the published whole genome sequences of the *B. megaterium* strains DSM319, QMB1551 as well as WSH-002 were used for the development and validation of a multiplex DNA-microarray. The designed microarray was not only used for analysis of differential gene expression profiles, but also for the identification of novel promoters in the industrially important *B. megaterium* strain MS941 (DSM319 Δ nprM) (Wittchen and Meinhardt, 1995). The identified promoters were successfully characterized via β -galactosidase assay and ultimately applied in *B. megaterium* for the efficient conversion of pregnenolone to progesterone, the universal precursor of steroid hormones.

2. Materials and methods

2.1. Bacterial strains, expression vectors, reagents and enzymes

Bacterial strains and vectors used in this study are listed in Supplemental Table 1. *E. coli* strain TOP10 (Invitrogen; Karlsruhe, Germany) was used for the assembly of all vectors. *B. megaterium* strain MS941 was used for the preparation of RNA and following microarray analyses as well as for the preparation of genomic DNA and subsequent

amplification of putative promoter elements. *B. megaterium* strain GHH1 was utilized for the deletion of the innate β -galactosidase activity resulting in *B. megaterium* strain GHH9, which was applied for the promoter screening. Conversion experiments with cholesterol oxidase II (BCO2) were carried out in *B. megaterium* strain GHH8.

The isolation kit for bacterial DNA was purchased from nexttec Biotechnologie GmbH (Hilgertshausen, Germany). The NucleoSpin[®] RNA isolation kit was acquired from Macherey-Nagel (Düren, Germany). All restriction enzymes were obtained from New England Biolabs (Ipswich; USA). The Fast-Link DNA Ligation Kit was acquired from Epicentre (Madison, USA). The In-Fusion[®] HD Cloning Kit was acquired from Takara (Saint-Germain-en-Laye, France). LB broth, tryptone and yeast extract were purchased from Becton Dickinson (Franklin Lakes, USA). Isopropyl- β -D-thiogalactopyranoside (IPTG) was acquired from Carbolution Chemicals GmbH (Saarbrücken, Germany). HPLC grade acetonitrile and ethyl acetate were obtained from VWR International GmbH (Darmstadt, Germany). O-nitrophenyl- β -D-galactopyranoside (ONPG) was purchased from TCI Deutschland GmbH (Eschborn, Germany). The purified *E. coli* β -galactosidase enzyme was acquired from Genlantis (San Diego, USA). All other chemicals were obtained from Sigma-Aldrich (St. Louis, USA).

2.2. Molecular cloning and gene deletion

The vectors were assembled with conventional cloning procedures using restriction enzyme digestion and ligation. The In-Fusion[®] HD Cloning Kit (Takara) was used in case that restriction site independent cloning was necessary. All vectors and primers that were used in this study were listed in Supplemental Table 1 and Supplemental Table 2, respectively.

For the promoter screening, the β -galactosidase (*lacZ*) gene was PCR amplified from pcDNA3.1-V5-His-TOPO-*lacZ* vector (Invitrogen; Karlsruhe, Germany) and inserted into the pSMF2.1 vector using *SpeI* and *KpnI* restriction enzymes. The forward primer was extended with the modified *xylA* ribosomal binding site (Stammen et al., 2010). The resulting vector pXyl^{*}-*lacZ* was used as reference for the evaluation and characterization of the novel promoter elements that were identified via microarray analyses. For the cloning of these putative promoters, the xylose inducible promoter system of the reference vector pXyl^{*}-*lacZ* was removed using restriction enzyme digestion and subsequently replaced by the novel, PCR amplified promoters using restriction enzyme digestion and ligation or restriction enzyme independent In-Fusion cloning.

The DNA sequence of the cholesterol oxidase II (BCO2) from *Brevibacterium sterolicum* was codon adapted for the heterologous expression in *B. megaterium* (www.jcat.de) and synthesized by Eurofins Genomics. The N-terminal amino acids were removed as described elsewhere and replaced by a modified hexahistidin tag (MSNNHHHHHH) (Volontè et al., 2010). The modified BCO2 gene was PCR amplified using primers with a *xylA* RBS and inserted into pSMF2.1 derived vectors with differing promoter elements using *SpeI* and *KpnI* restriction enzymes.

The disruption cassette for the deletion of the innate β -galactosidase gene (*BMD_2126*) was constructed as described previously (Gerber et al., 2015). In brief, the flanking regions of *BMD_2126* were PCR amplified from genomic DNA of *B. megaterium* MS941. The resulting PCR fragments were then fused by overlapping extension PCR and inserted into the knockout vector pUCTV2_Upp which contains a heat sensitive origin of replication (*ori*^{ts}). *B. megaterium* strain GHH1 was transformed with the resulting knockout vector pUCTV2_Upp_Δ*BMD_2126* using PEG-mediated protoplast transformation (Brown and Carlton, 1980). Transformants were grown in LB medium (25 g/L) containing 10 μ g/ml tetracycline at 30 °C and 150 rpm for 16 h. Cultures were diluted appropriately, spread on modified minimal medium agar plates [2 g (NH₄)₂SO₄, 6 g KH₂PO₄, 14 g K₂HPO₄, 1 g Na₃C₆H₅O₇, 200 mg MgSO₄, 0.8 mg MnSO₄ · H₂O, 1 mg thiamine

hydrochloride, 4 g glycerol, 2 g glutamate and 16 g agar per 1000 ml H₂O] without tetracycline and incubated over night at non-permissive temperature of 42 °C. Subsequently, colonies were replica plated on modified minimal medium agar plates supplemented with 1 μM 5-fluoro-uracil to counter-select colonies that still contained the knock-out vector. The resulting vector free colonies were screened for successful deletion of β-galactosidase (*BMD_2126*) by PCR.

2.3. RNA extraction

500 ml of modified M9CA medium were inoculated 1:100 with a *B. megaterium* preculture in a Minifors (Infors HT) bioreactor. M9CA medium (pH = 6.0) consisted of 6 g Na₂HPO₄, 3 g KH₂PO₄, 0.5 g NaCl, 1 g NH₄Cl, 2 g casamino acids, 30 g glycerol, 100 μl 1 M CaCl₂, 2 ml 1 M MgSO₄, 100 μl 10 mM thiamine hydrochloride and 1 ml of 1000 x trace elements solution per liter H₂O. 1000 x trace element solution contained 2.5 g EDTA, 695 mg FeSO₄·7H₂O, 99 mg MnCl₂·4H₂O, 68 mg ZnCl₂, 24 mg CoCl₂·6H₂O, 17 mg CuCl₂·2H₂O, 24 mg NiCl₂·6H₂O, 24 mg Na₂MoO₄·2H₂O, 26 mg Na₂SeO₃·5H₂O and 6 mg H₃BO₃ per 50 ml H₂O. *B. megaterium* cells were cultured for 24 h at 37 °C and constant aeration. Samples for RNA extraction were taken at different stages of bacterial growth and stored at –20 °C in 60% methanol.

Bacteria pellets were washed twice with TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH = 8). RNA was prepared with the NucleoSpin® RNA isolation kit of Macherey-Nagel (Düren, Germany) according to the manufacturer's protocol with slight modifications. Since *B. megaterium* cells become resistant to lysozyme treatment following mid-exponential phases of bacterial growth, lysozyme was replaced by equal amounts of endolysin.

2.4. Microarray design and validation

Based on sequence data of three *B. megaterium* strains (DSM319, QMB1551 and WSH-002) retrieved from NCBI, on average 10 isothermal 45–60 mer probes were generated for every gene and represented on 2 × 400 K Agilent microarrays. RNA and DNA samples of *B. megaterium* strain DSM319 were processed and labelled with fluorescent dyes and hybridized on the microarrays. Based on the hybridization data and by weighting several parameters, the best performing probe was selected and represented on 8 × 15 K microarrays. The probe selection algorithm considers the overall probe signal intensities as well as a probe's capability to report a differential expression. A total of 9931 target IDs are represented to detect 13,725 targets of the *B. megaterium* strains.

2.5. RNA labelling and microarray hybridization for gene expression analysis

Concentration, purity and integrity of RNA samples were determined photometrically using the Nanodrop (Thermo Fisher Scientific) and electrophoretically using the RNA 6000 Pico Kit with the 2100 Bioanalyzer (Agilent Technologies). All samples had RNA integrity (RIN) numbers above 5.3. Fluorescent cRNA was generated using the Low-Input QuickAmp Labelling Kit (Agilent Technologies) using 50 ng of total RNA with WT primers following the manufacturer's protocol. Yields and cy3 incorporation rates were measured photometrically (Nanodrop) and all samples yielded more than 5.9 μg with at least 34 pmol cy3 incorporated per μg. 600 ng of each cy3 labelled cRNA were hybridized using the Agilent Gene Expression Hybridization Kit (Agilent Technologies) for 17 h at 65 °C on 8 × 15 K microarrays described in section "Microarray design and validation" followed by microarray washing and scanning on the SureScan Microarray Scanner (Agilent Technology) at a resolution of 3 μm per pixel producing 20-bit TIF files according to the manufacturer's protocols.

2.6. Microarray data analysis

TIF files were extracted using the Feature Extraction Software V11 (Agilent Technologies) and protocol GE1_1105_Oct12. The resulting raw data were analyzed using the DirectArray Software (OakLabs GmbH). The probes were filtered for the DSM319 targets and all samples were quantile normalized and subjected to statistical group-wise comparisons. Here, a Welch's test was applied and log₂ fold changes were calculated. The false discovery rate was controlled by an adaptive Benjamini Hochberg procedure (Benjamini and Hochberg, 2000). Default cutoffs for p-value was 0.05 and for the fold change 2-fold (log₂ < –1 or > 1).

2.7. *B. megaterium* cultivation conditions for protein expression

The LacZ based promoter screening was carried out with cells of the β-galactosidase deficient *B. megaterium* strain GHH9. The BCO2 mediated conversion of pregnenolone to progesterone was performed in the *B. megaterium* strain GHH8 which is deficient of CYP106A1 as well as 20α-HSD catalyzed side product formation of steroids. The corresponding LacZ and BCO2 expression vectors (listed in Supplemental Table 1) were transformed using PEG mediated protoplast transformation. Precultures were started from single colonies in LB medium (25 g/l) supplemented with tetracycline (10 μg/ml) and incubated overnight at 37 °C and 150 rpm. A total volume of 25 ml TB medium (24 g/l yeast extract, 12 g/l tryptone, 0.4% glycerol, 100 mM potassium phosphate buffer, pH 7.4) was supplemented with tetracycline (10 μg/ml) and inoculated with 1:100 of the corresponding preculture. The main cultures were grown in 300 ml baffled flasks at 37 °C and 150 rpm to an optical density of 0.4 at 578 nm. In case of inducible vectors pXyl*, pAra, pSuc, pGal and pBga, protein expression was induced by addition of xylose, arabinose, sucrose, galactose to a final concentration of 5 mg/ml (w/v) and 1 mM IPTG, respectively. Growth phase dependent as well as constitutive promoters were not induced. Cells were further grown at 37 °C and 150 rpm for 21 h and subsequently screened for β-galactosidase activity or used for pregnenolone conversion.

2.8. β-galactosidase assay

After expression of β-galactosidase, bacteria pellets were washed twice with dilution buffer (50 mM potassium phosphate buffer, pH = 7.4) and subsequently suspended in lysis buffer (50 mM potassium phosphate buffer, pH = 7.4, 1 mg/ml muramidase). Cell lysis was carried out as described elsewhere (Stenger et al., 2018). The lysates were diluted appropriately to allow moderate color development. 100 μl of the diluted lysates were transferred into a microtiter plate. Accordingly, a standard curve was prepared with a purified β-galactosidase enzyme from *E. coli* provided by Genlantis. The assay was started by addition of 50 μl of the ONPG substrate solution (50 mM potassium phosphate buffer, pH = 7.4, 3 mg/ml ONPG) and incubated at room temperature for an adequate amount of time. The reaction was stopped by addition of 150 μl of the stop solution (1 M Na₂CO₃) and the color development was monitored at 420 nm on a VICTOR™ X3 multilabel plate reader from PerkinElmer (Waltham, USA). β-galactosidase expression levels were calculated by reference to the standard curve using linear regression analysis.

2.9. In vivo conversion of pregnenolone

After expression of cholesterol oxidase II (BCO2), cells were harvested by centrifugation at 4000g for 15 min and washed twice with conversion buffer (50 mM potassium phosphate buffer, pH = 7.4). For conversion of pregnenolone to progesterone, bacteria pellets were suspended in conversion buffer and cell suspension concentrations were adjusted to 50 g/l (wet cell weight in conversion buffer). The substrate pregnenolone and the internal standard 21-hydroxy-progesterone

(DOC) were added to a final concentration of 5 mM and the cultures were incubated at 37 °C and 200 rpm for 24 h. Samples were taken at the indicated time and extracted twice with equal volumes of ethyl acetate by vigorous shaking. The organic supernatants were combined and evaporated to dryness using a rotary evaporator. Residues were suspended in 10% acetonitrile and subjected to HPLC analysis. Adequate separation of the product progesterone and the internal standard DOC was achieved by RP-HPLC using an EC 125/4 Nucleodur C18 column (Macherey-Nagel) on a Jasco system. The mobile phase consisted of 10% acetonitrile in water on channel A and pure acetonitrile on channel B. The following gradient was applied for HPLC measurements: 0–3 min: 100% A to 90% A; 3.1–5 min: 90% A; 5.1–25 min: 90% A to 0% A; 25.1–30 min 0% A to 100% A. The flow rate was set to 0.8 ml/min. Column oven temperature was set to 40 °C. The injection volume was 50 µl. Detection of progesterone and DOC was carried out at a wavelength of 240 nm.

3. Results and discussion

3.1. Identification of putative promoter elements in *B. megaterium* MS941 via microarray analysis

The identification of novel promoters in the *B. megaterium* strain MS941 was carried out using a genome-wide microarray analysis. The generated data covered a total of 5098 target *open reading frames* of *B. megaterium* MS941 at different time points of bacterial growth, representing early exponential, mid-exponential, late exponential as well as early stationary growth phases. The differential gene expression profiles throughout bacterial growth of *B. megaterium* MS941 were shown in Fig. 1. The gene expression levels during early exponential growth were selected as reference for differential gene expression analysis. As shown in Fig. 1 only a small subset of transcription levels was significantly changed in mid-exponential phase with 1.76% of the genes showing lower and 2.76% of the genes showing higher transcription levels, respectively. During late exponential growth phase a total of 10.08% of the analyzed genes showed significantly decreased transcription levels whereas 10.57% of the genes showed significantly increased transcription levels. The most fundamental changes in gene expression profiles, however, were present during early stationary growth phase in which a decrease of 30.07% and an increase of 25.28% of all analyzed transcript levels was observed.

Predominantly, genes that showed increased transcription levels

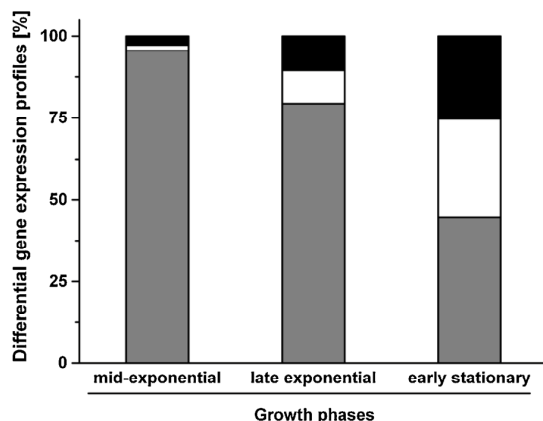


Fig. 1. Differential gene expression profiles during cultivation of *B. megaterium* MS941. Microarray data of all analyzed genes were normalized to their corresponding transcription levels in the early exponential growth phase. Transcription levels were found to be: significantly increased (black bars, log₂ fold change > 1); significantly decreased (white bars, log₂ fold change < -1) or not affected (grey bars); cutoff for p-value was 0.05.

were further used for the identification of putative promoter elements. Throughout analysis of the generated microarray data, these genes were categorized into two groups: genes whose transcript levels showed the highest fold changes but rather low overall signal intensities and genes with transcript levels that showed low fold changes but the highest overall signal intensities. Representative genes of both groups were selected and the corresponding changes of transcript levels during bacterial growth were displayed in Fig. 2. On the one hand, the five genes *BMD_0462*, *BMD_0551*, *BMD_1081*, *BMD_2097* and *BMD_2948* were selected as the open reading frames (ORFs) with the highest fold changes. The functional annotations of these ORFs (listed in Table 1) ranged from thiamine biosynthesis and glycerol transport to proteins with uncharacterized functions.

As shown in Fig. 2A the fold change of transcription levels of *BMD_0551*, *BMD_1081* as well as *BMD_2948* demonstrated a significant increase in early stationary growth phases while the fold change of transcript levels of *BMD_0462* and *BMD_2097* were highest during mid and late exponential phases. Strikingly, the transcript levels of the genes that were located downstream of *BMD_0462* and *BMD_0551* showed similar fold changes, which indicated that these genes were transcribed as operons under the control of the same regulatory element (data not shown).

On the other hand, the genes *BMD_0123*, *BMD_0450*, *BMD_0706*, *BMD_1214*, *BMD_2386*, *BMD_2618*, *BMD_3115*, *BMD_3537*, *BMD_4632* and *BMD_4756* were selected as ORFs that exhibited some of the highest overall signal intensities (Fig. 2B). The encoded enzymes seem to be involved in a variety of cellular processes, ranging from central carbon metabolism to protein biosynthesis as well as solute transport and stress tolerance (Table 1). Among the selected genes, *BMD_0123*, *BMD_0706* and *BMD_4756* showed the highest signal intensities during early exponential phases, while *BMD_0450* seemed to be specifically transcribed following mid-exponential phases. Genes with highest transcription levels in late exponential phases were represented by the ORFs *BMD_2386*, *BMD_2618*, *BMD_3537* as well as *BMD_4632*. Although the genes *BMD_1214* and *BMD_3115* showed consistent transcription levels during bacterial growth, they were still included as constitutively transcribed genes with high signal intensities.

3.2. Generation of a β -galactosidase-deficient *B. megaterium* strain as suitable host for a β -galactosidase based promoter screening system

B. megaterium is known to possess an innate gene which encodes a β -galactosidase enzyme (LacZ) that was predicted to interfere with β -galactosidase based screening systems (Schmidt et al., 2005). The basal activity of LacZ was eliminated by engineering a β -galactosidase-deficient *B. megaterium* strain GHH9 (Supplemental Table 1).

The disruption cassette for the genomic deletion of *BMD_2126* (*lacZ*) was constructed as described in the material and method section (2.2). Cells of the *Bacillus megaterium* strain GHH1 were transformed with the resulting vector pUCTV2_Upp_Δ*BMD_2126*. The subsequent knockout procedure was carried out as described previously (Gerber et al., 2015). The genomic DNA of the resulting colonies was screened for successful *lacZ* deletion via PCR (Fig. 3A). A colony with the deletion was identified and designated as GHH9. The PCR product with genomic DNA of the *lacZ* knockout strain GHH9 was found to be truncated by approximately 2900 bp compared to the PCR product with genomic DNA of the parental strain GHH1 (Fig. 3B). Moreover, this PCR product also exhibited the same size as the PCR product of the knockout vector pUCTV2_Upp_Δ*BMD_2126*, indicating the successful deletion of *BMD_2126* (*lacZ*). Consistently, phenotypical characterization of the *lacZ* knockout strain GHH9 showed no detectable β -galactosidase activity (Fig. 3C).

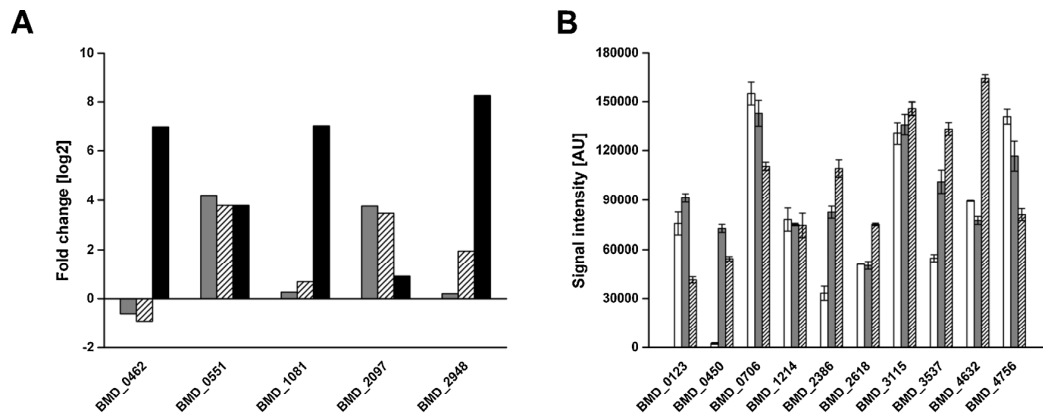


Fig. 2. Microarray analysis of transcription levels for the identification of putative promoter elements in *B. megaterium* MS941. **A:** Transcription levels of ORFs with the highest log2 fold changes. Microarray data were normalized to the transcription levels in the early exponential growth phase. **B:** Transcription levels of ORFs with highest overall signal intensities. White bars indicate early exponential stages, grey bars represent mid-exponential stages, lined bars show late exponential stages and black bars represent early stationary stages of bacterial growth. Experiments were performed in biological triplicates (n = 3).

Table 1
Computational analysis of selected putative promoter sequences and their adjacent open reading frame.

ORF	Annotation	-35 box	-10 box	Transcription factors
BMD_0123	50S ribosomal protein L10	TTGACT	GATTATAAT	LexR, PdhR, CpxR
BMD_0450	thiamine biosynthesis protein ThiC	TTGATT	TGGTAAACT	LexA, ArgR, ArgR2, Ihf
BMD_0462	putative glycerol-3-phosphate ABC transporter	TCGTTA	CCCTATTAT	GlpR
BMD_0551	putative thiamine-phosphate pyrophosphorylase	TTGTAT	TGTTATAAT	RpoD17, ArgR, Fnr
BMD_0689	galactokinase of the galactose utilization operon	TTTTTA	ATCTATTAT	Fnr, PhoB, RpoD19, Crp, SoxS
BMD_0706	oligopeptide-binding protein OppA	TTGCAA	ATTTATAAT	RpoD17, RpoD18, Crp, Lrp, ArgR2
BMD_1081	conserved hypothetical protein	ATGAAT	TAGTAGATT	Lrp, LexA, ArgR, ArgR2, ArcA
BMD_1214	polyhydroxyalkanoic acid synthase, PhaR subunit	TTGTTG	ATTTACATT	RpoD16, RpoD17, ArgR, OmpR, LexA, TyrR
BMD_1565	sucrose permease	TTTACA	TTTTTAAAT	RpoD17; RpoH2, LexA, Fis
BMD_1858	xylose isomerase of the xylose utilization operon	TTTCAA	AGTTATAAT	RpoH2, LexA
BMD_2097	tena/thi-4 family domain protein	TTTACT	GTTTATAAA	MetJ, ArgR, ArgR2, Ihf
BMD_2126	β -galactosidase	TTGATA	CGTTATCAT	RpoS17, RpoD17, RpoD16, ArgR2, SoxS, Fur, FhlA, PhoB, RpoD19
BMD_2386	putative D-3-phosphoglycerate dehydrogenase	TTGAAC	TGCTAGAAT	Fis, RpoD17
BMD_2618	NADH dehydrogenase	TTAACA	GTGTTAGAT	PdhR; RpoD19
BMD_2948	conserved hypothetical protein	TTAAAT	GTTTAAATAT	Fis, RpoD18, TyrR, GlpR
BMD_3115	malate dehydrogenase	TTTAAG	TTTTATGTT	RpoS17
BMD_3532	carbohydrate kinase of the arabinose utilization operon	TTGACA	AAATATACT	RpoD17, RpoD16, Fnr, Fis, OmpR
BMD_3537	putative ferrous iron transport protein	TTAAAA	ACATAACCT	ArgR2, Lrp, RpoH2, LexA, RpoD17
BMD_4632	conserved hypothetical protein, probably heat tolerance	TTGTTA	TTGTAAAAAT	SoxS, ArcA, PhoB, Hns, RpoD15, RpoD17
BMD_4756	citrate synthase II	TTGATT	GTTTATAAT	LexA, RpoD18, SoxS, DnaA, ArcA, Fur, RpoD17, PhoB

3.3. Evaluation and characterization of putative promoter elements in *B. megaterium* with a β -galactosidase based screening system

The BPROM online tool was used to predict the most common features of promoters, namely the -35 and -10 regions along with putative binding sites for well characterized transcription factors (www.softberry.com). The identified consensus sequences were listed in Table 1.

For the evaluation of promoter activities and comparison of the corresponding promoter strengths, the intergenic regions upstream the analyzed ORFs were PCR amplified from genomic DNA of *B. megaterium* MS941 and inserted into the pSMF2.1 vector containing β -galactosidase (*lacZ*) as reporter gene. The optimized xylose inducible promoter was used as reference. This promoter was modified by Stammen et al. and showed nearly 8-fold higher expression levels compared with the wild type xylose inducible promoter (Stammen et al., 2010). The resulting vectors were transformed into the *lacZ* deficient *B. megaterium* strain GHH9.

The β -galactosidase activities of the transformants after 21 h of cultivation were shown in Fig. 4. Although the microarray data suggested high fold changes in transcription levels, the β -galactosidase activities of all tested high fold change promoters were considerably

lower compared with the optimized xylose inducible promoter (Fig. 4A). Among these promoters, p1081 and p2097 showed higher β -galactosidase activities in complex medium than in M9CA medium. Their highest β -galactosidase activities were found to be 700 mU and 900 mU which corresponded to approximately 15% and 20% activity compared with 5200 mU for the optimized xylose inducible promoter pXyl*. The β -galactosidase activities driven by p0462 and p0551 and p2948 were found to be negligible compared with the reference promoter.

In contrast to the fold change promoters, the regulatory elements of genes that showed high overall signal intensities in microarray analysis seemed to drive β -galactosidase expression more efficiently (Fig. 4B). Protein expression levels of the most promising promoters were displayed in Supplemental Fig. 1. Despite the fact that low β -galactosidase activities around 1000 mU were observed for promoters p0123, p0450, p2618 and p3115, moderate β -galactosidase activities of around 2000 mU were achieved with promoters p1214 as well as p4632. Highly active promoter candidates were identified in promoters p0706 as well as p4756. Both of them were specific for early to mid-exponential growth phases and showed β -galactosidase activities between 4000 mU and 4700 mU, which were comparable to the activity of the optimized xylose promoter. Furthermore, heterologous expression using these two

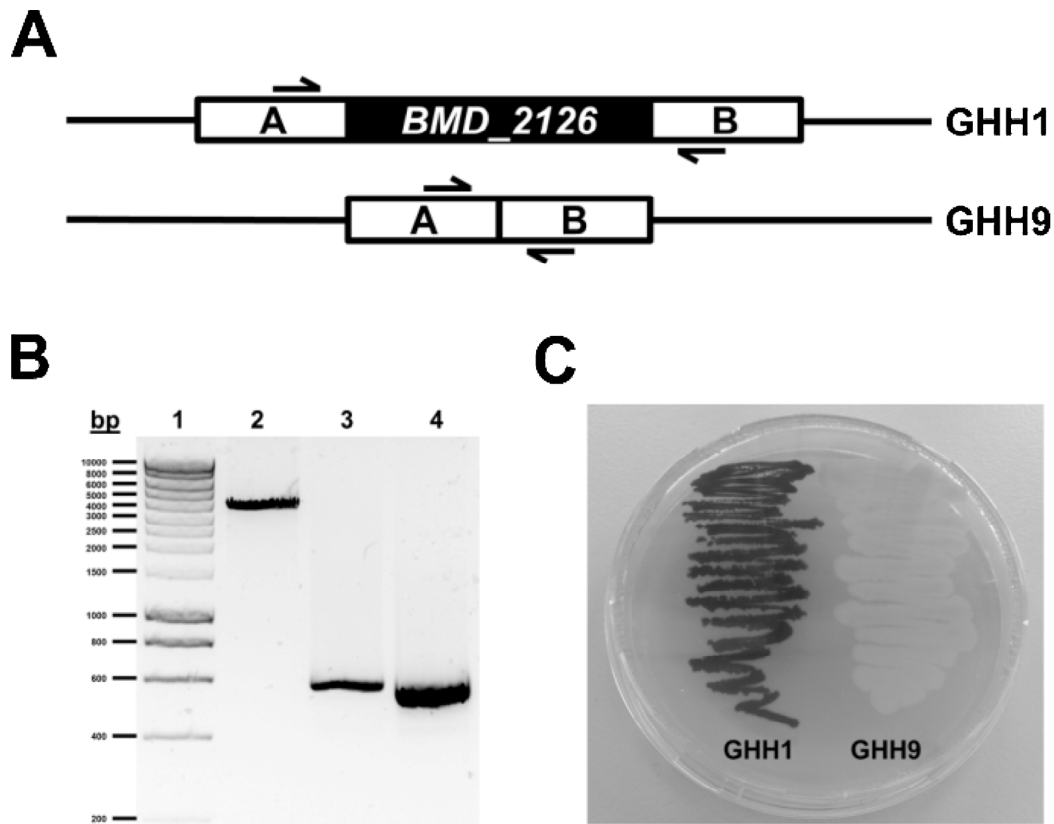


Fig. 3. Genotypic and phenotypic verification of the genomic deletion of *BMD_2126* (*lacZ*) **A:** Genomic organization of the ORF *BMD_2126* and flanking regions (A and B) in the parental strain GHH1 and the *lacZ* knockout strain GHH9. Black arrows indicate the position of the primers used for the knockout screening (primer names and sequences are listed in Supplemental Table 2). **B:** Agarose gel electrophoresis showing PCR amplification products with genomic DNA of the parental strain GHH1 (lane 2), with genomic DNA of the *lacZ* knockout strain GHH9 (lane 3) and the knockout vector pUCTV2_Upp_Δ*BMD_2126* (lane 4) as template. The first lane shows the DNA molecular weight marker SmartLadder (Eurogentec). The PCR product of the *lacZ* knockout strain GHH9 (lane 3) is truncated by approx. 2900 bp compared to the PCR product of the control strain GHH1 (lane 2) and corresponds to the PCR product size of the knockout vector (lane 4). **C:** Parental strain GHH1 as well as *lacZ* knockout strain GHH9 were grown overnight on an agar plate containing X-gal (200 µg/ml). In contrast to the control strain GHH1, the knockout strain GHH9 is not able to convert X-gal, indicating the successful genomic deletion of *BMD_2126* (*lacZ*).

promoter candidates seemed to be very robust since similar high β-galactosidase expression levels were obtained independent of the cultivation medium. The consistency of the β-galactosidase activities along with high expression levels are both very important criteria for promoters in biotechnological applications. The highest β-galactosidase activities with 5600 mU and 6300 mU, however, were obtained with

the promoters p2386 and p3537. The β-galactosidase activities of these promising promoter candidates exceeded the β-galactosidase activity of the optimized xylose inducible promoter by 10% and 30%, respectively. Additionally, they both were specific for late exponential growth phases.

Intriguingly, the promoter p2386 seemed to have a strong

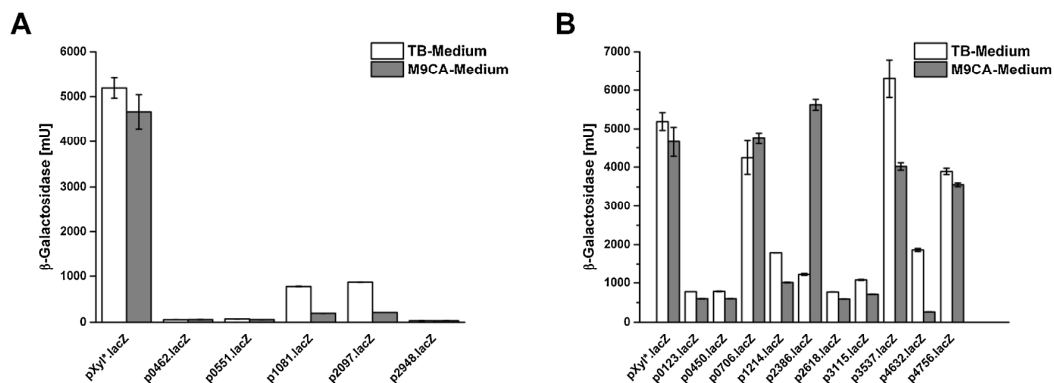


Fig. 4. β-Galactosidase assay with novel growth dependent promoter elements in *B. megaterium*. β-Galactosidase activities were measured in transformed *B. megaterium* strain GHH9 after 21 h of cultivation. **A:** Promoters with highest log₂ fold change. **B:** Promoters with highest signal intensities. Experiments were performed in biological triplicates (n = 3).

preference for the modified M9CA minimal medium, while the promoters p3537 and p4632 showed higher β -galactosidase activities in complex medium. Since the corresponding open reading frame of p2386 is predicted to encode a putative D-3-phosphoglycerate dehydrogenase, which is possibly involved in the metabolism of glycerol, we hypothesized that the elevated concentration of glycerol in the M9CA minimal medium is responsible for the higher β -galactosidase activity compared to the activity in complex medium. However, titration experiments with glycerol did not show any beneficial effect on β -galactosidase activities neither in M9CA medium nor in complex medium (Supplemental Fig. 2A). Thus, an explanation for the different specificities of promoter p2386 for the cultivation medium remains unclear. The open reading frame downstream of the promoter p3537 was predicted to encode a putative iron transporter, which is possibly involved in iron uptake. While titration experiments with increasing iron concentrations did not change the β -galactosidase activity of p3537 in M9CA medium, the β -galactosidase activity of p3537 was significantly increased to 7560 mU in complex medium which corresponds to a 20% higher activity than in case of no iron addition (Supplemental Fig. 2B). Thus, the newly identified promoter p3537 seems to respond to elevated iron levels in complex medium. Initially, promoter p4632 showed moderate β -galactosidase activities of approximately 2000 mU in complex medium. However, since the corresponding open reading frame encodes a putative stress inducible protein, several cultivation parameters, including salt concentration, pH value as well as cultivation temperature were tested for an activatory effect on promoter p4632. Among all tested stress conditions, a temperature shift towards 42 °C significantly improved the β -galactosidase activity of p4632 to 4875 mU which corresponded to a 2.6 fold increase compared with the culture at 37 °C. Intriguingly, the β -galactosidase activity of the xylose inducible promoter dropped from approximately 5200 mU to 2019 mU (Supplemental Fig. 2C). To the best of our knowledge, promoter p4632 represents the first described heat inducible promoter in *B. megaterium*. This promoter will be particularly beneficial for the heterologous expression of proteins that require higher temperatures to achieve optimal catalytic activities.

3.4. Identification and characterization of novel inducible promoters in *B. megaterium*

The class of inducible promoters allows controllable expression of recombinant proteins in *B. megaterium* and has almost exclusively been limited to the xylose inducible promoter (Bleif et al., 2012; Rygus et al., 1991). Originally discovered and characterized by Rygus and Hillen in 1991, this xylose inducible promoter was later systematically modified and optimized for a high-yield protein production in *B. megaterium* (Stammen et al., 2010). The binding of a xylose repressor has been reported to be responsible for the tight control of the xylose inducible promoter. The open reading frame of this xylose repressor was found to be located immediately upstream and in opposite direction of the xylose promoter (Rygus and Hillen, 1991). Hence, bioinformatic analyses for the prediction of novel sugar-inducible promoter elements in *B. megaterium* were carried out on the basis of the architecture of the xylose promoter-repressor system. Putative promoter and repressor elements of arabinose, sucrose, galactose as well as lactose utilization operons were determined in the genome of *B. megaterium* MS941 (Fig. 5A).

The predicted inducible promoters were amplified from genomic DNA of *B. megaterium* MS941 along with their corresponding repressor elements and inserted into the pSMF2.1 vector, containing β -galactosidase (*lacZ*) as reporter gene. The optimized xylose inducible promoter was used as reference. The resulting vectors were transformed into the *lacZ* deficient *B. megaterium* strain GHH9. The β -galactosidase expression was induced by the addition of corresponding sugars (5 mg/ml) or IPTG (1 mM) at OD₅₇₈ 0.4 and β -galactosidase activities were measured after 21 h of cultivation. As shown in Fig. 5B the distinct sugar-inducible promoters exhibited different levels of β -galactosidase activity.

Although the putative sucrose inducible promoter was located in proximity to an operon for sucrose metabolism, β -galactosidase expression was inefficiently induced using sucrose and other potential sugars. Moderate β -galactosidase expression levels of 1245 mU and 1810 mU were achieved using the galactose and lactose inducible promoter system, respectively. Both promoters showed considerably lower basal β -galactosidase activities than the reference promoter, indicating a tighter regulation. However, the arabinose inducible promoter evidently exhibited the highest β -galactosidase activity as well as tightest regulation among all newly identified promoter elements. While the β -galactosidase activity was slightly decreased to 3914 mU using the arabinose inducible promoter, its basal activity was shown to be significantly improved compared to the optimized xylose inducible promoter. Particularly, expression of complex or toxic proteins will benefit from the tight regulation of the novel arabinose inducible promoter, thus providing an ideal system for the controllable recombinant protein expression in *B. megaterium*.

3.5. Validation and application of the novel promoter elements for steroid conversion in *B. megaterium*

Apart from the successful characterization via β -galactosidase assay, the activities of the novel promoters were ultimately validated in *B. megaterium*, applying the cholesterol oxidase II (BCO2) from *Brevibacterium sterolicum* for the conversion of pregnenolone to progesterone, which is a key intermediate for steroid production (Sanderson, 2006). Cholesterol oxidases are bifunctional flavoenzymes that catalyze the oxidative conversion of Δ^5 -3 β -hydroxysteroids to the corresponding steroids with Δ^4 -3-keto configuration (Coulombe et al., 2001; Croteau and Vrielink, 1996). The formation of this 3-keto-4-ene structure allows the detection of the resulting steroids at a characteristic wavelength of 240 nm. Previous studies have demonstrated the importance of cholesterol oxidases for clinical, agricultural and biocatalytic applications (Pollegioni, 2009).

Prior to pregnenolone conversion, several untransformed *B. megaterium* strains were tested for side product formation regarding progesterone as the desired end product of the cholesterol oxidase reaction as well as potential internal standards for the calculation of extraction efficiency. As shown in Supplemental Fig. 3A, the *B. megaterium* strain MS941 was not suitable for the establishment of the cholesterol oxidase reaction as considerable side product formation was observed after 21 h of cultivation with progesterone. Some of these side products were shown to derive from innate activity of CYP106A1, while others are presumably resulting from hydroxysteroid dehydrogenases (HSDs) (Lee et al., 2015). Fortunately, a CYP106A1 as well as 20 α -HSD deficient *B. megaterium* strain was engineered and available in our laboratory (Gerber et al., 2016). As expected, this *B. megaterium* GHH8 did not show any side product formation, thus representing the ideal strain for the cholesterol oxidase (BCO2) mediated conversion of pregnenolone to progesterone. Among all tested compounds, 21-hydroxy-progesterone (DOC) was found to be a suitable internal standard since metabolization to unwanted side products was neither observed in *B. megaterium* strain MS941 nor in strain GHH8 (Supplemental Fig. 3B).

Finally, the validation of the novel promoters via cholesterol oxidase (BCO2) mediated conversion of pregnenolone to progesterone was carried out in *B. megaterium* strain GHH8 (Fig. 6). Representative promoter candidates of novel inducible as well as growth-phase dependent promoters were tested for cholesterol oxidase (BCO2) activity. The optimized xylose inducible promoter was used as reference.

Strikingly, all tested promoter candidates showed BCO2 activities comparable to results of the β -galactosidase assays, thus emphasizing the reliability and efficiency of the established promoter screening system in *B. megaterium*. Compared to the reference promoter, the arabinose inducible promoter showed slightly decelerated progesterone formation kinetics but ultimately yielded the same amount of 3.5 mM progesterone after 24 h of conversion. Likewise, the growth-phase

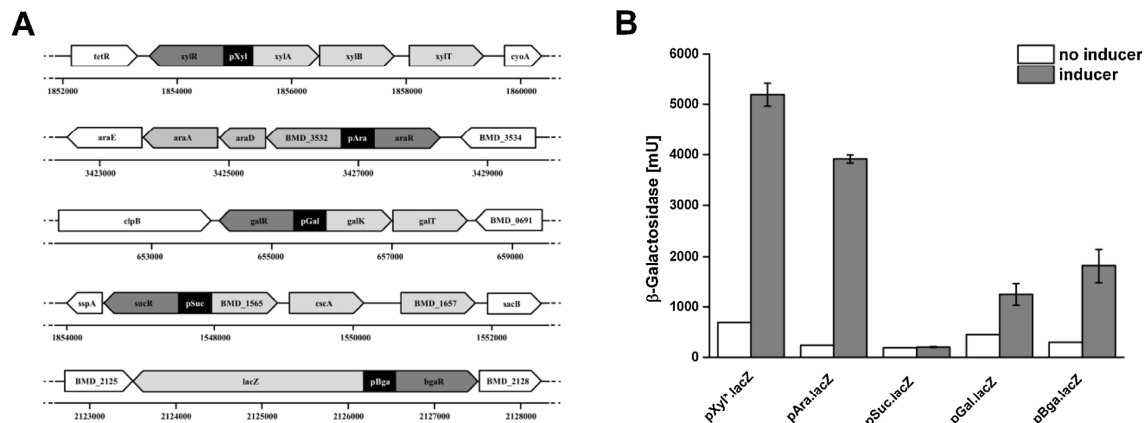


Fig. 5. Identification and characterization of novel sugar inducible promoters in *B. megaterium* MS941. **A:** Genomic context of sugar inducible promoters. Putative promoter elements are shown in black along with their corresponding repressor elements in dark grey. Open reading frames under the control of the respective sugar inducible promoter are represented in light grey (the annotation and location of the genes was adopted from MegaBac v9 database). **B:** β -Galactosidase activities were measured in transformed *B. megaterium* strain GIII19 after 21 h of cultivation. Experiments were performed in biological triplicates ($n = 3$).

dependent promoters p0706 and p4756, both specific for early exponential phases of bacterial growth, showed slower progesterone formation rates than the optimized xylose inducible promoter. While the final product yield with promoter p4756 was limited to 2.9 mM progesterone, a total yield of 3.3 mM progesterone was obtained when BCO2 expression was driven by promoter p0706. Based on β -galactosidase expression levels, the late exponential phase promoter p3537 was anticipated to yield the highest progesterone titers. As expected, final progesterone yields of 3.6 mM were found to be slightly increased compared with the reference promoter.

Considering the fact that all novel promoter elements are wildtype DNA sequences, their progesterone yields were additionally compared with the progesterone yields achieved with the non-optimized, wildtype xylose inducible promoter system in order to further emphasize their potential for recombinant protein expression in *B. megaterium*. BCO2 mediated formation of progesterone using the wildtype xylose inducible promoter system barely resulted in more than 1.4 mM progesterone (Fig. 6F). All novel promoters exceeded this progesterone yield at least by the factor 2, thus offering an innovative and competitive option for heterologous protein production in *B. megaterium*.

4. Conclusion

Previous studies have demonstrated the importance of promoters as regulatory elements for the efficient heterologous expression of proteins in bacterial hosts (Guan et al., 2016; Ming et al., 2010; Wilkinson et al., 2002; Yu et al., 2008, 2015). Finding a suitable promoter for the production of a particular protein is often based on trial and error approaches and therefore considered to be very laborious and time consuming (Bland et al., 2010; Won et al., 2008). This may be the reason why recombinant protein production in *B. megaterium* has exclusively been limited to the xylose inducible promoter system for several decades. Although there have been a few attempts for the establishment of alternative promoter systems, such as the sucrose inducible *sacB* promoter or the T7 system, the optimized xylose inducible promoter is still the most efficient and most frequently used promoter system for the heterologous protein production in *B. megaterium* (Biedendieck et al., 2007; Gamer et al., 2009). While the application of a single promoter is often considered to be sufficient for the high-yield production of a particular protein, rational engineering of complex metabolic pathways involving multi-enzyme cascades necessarily relies on finely tuned transcription levels, thus resulting in well balanced protein expression (Wang et al., 2005; Jones et al., 2015; Pitera et al., 2007; Xu et al., 2014). Given the scarcely described repertoire of promoters for *B.*

megaterium, there is no possibility to develop and employ *B. megaterium* as microbial cell factory for the sophisticated biosynthesis and conversion of natural products and pharmaceutical compounds. For this reason, there is an increasing demand for novel promoters that offer an expanded range of promoter activities, thereby facilitating the modulation as well as the optimization of proteomic and metabolomic networks in *B. megaterium*.

Applying genome-wide microarray analyses, we were able to identify novel promoters in the industrially important *B. megaterium* strain MS941. The provided library of novel promoter elements comprises innovative growth phase dependent and constitutive promoters as well as sugar inducible promoters. Our results demonstrated that most identified promoters were able to produce sufficient levels of active β -galactosidase ranging from approximately 15% to 130% compared with the optimized xylose inducible promoter. A major concern of common promoter screening approaches is the fact that many variables, but mainly RNA structure and stability, are supposed to significantly affect the expression levels of the used reporter gene (Arnold et al., 1998; Baim and Sherman, 1988; Dvir et al., 2013; Kabardin and Bläsi, 2006; Klaff et al., 1996; Niepel et al., 1999). For this reason, selected promoter activities were validated by heterologous expression of the cholesterol oxidase II (BCO2) from *Brevibacterium sterolicum*. Corresponding activities were shown to be consistent with β -galactosidase levels of the promoter screening thereby indicating that the evaluated promoter strengths neither seemed to be influenced by the RNA structure nor by the RNA stability of the reporter gene itself. Moreover, the open reading frames of the most active promoters were also found by Wang and colleagues among the most abundant proteins during batch cultivation of *B. megaterium* MS941 (Wang et al., 2005). All together, these findings emphasize the accuracy and reliability of the generated microarray data for the efficient promoter screening system.

To the best of our knowledge, this is the most comprehensive study for the identification and characterization of useful promoter elements in *B. megaterium*. The different promoter strengths ranging almost over one order of magnitude will significantly contribute to the engineering of *B. megaterium* as microbial cell factory for the biosynthesis and conversion of valuable compounds.

Competing interests

The authors declare no competing interest.

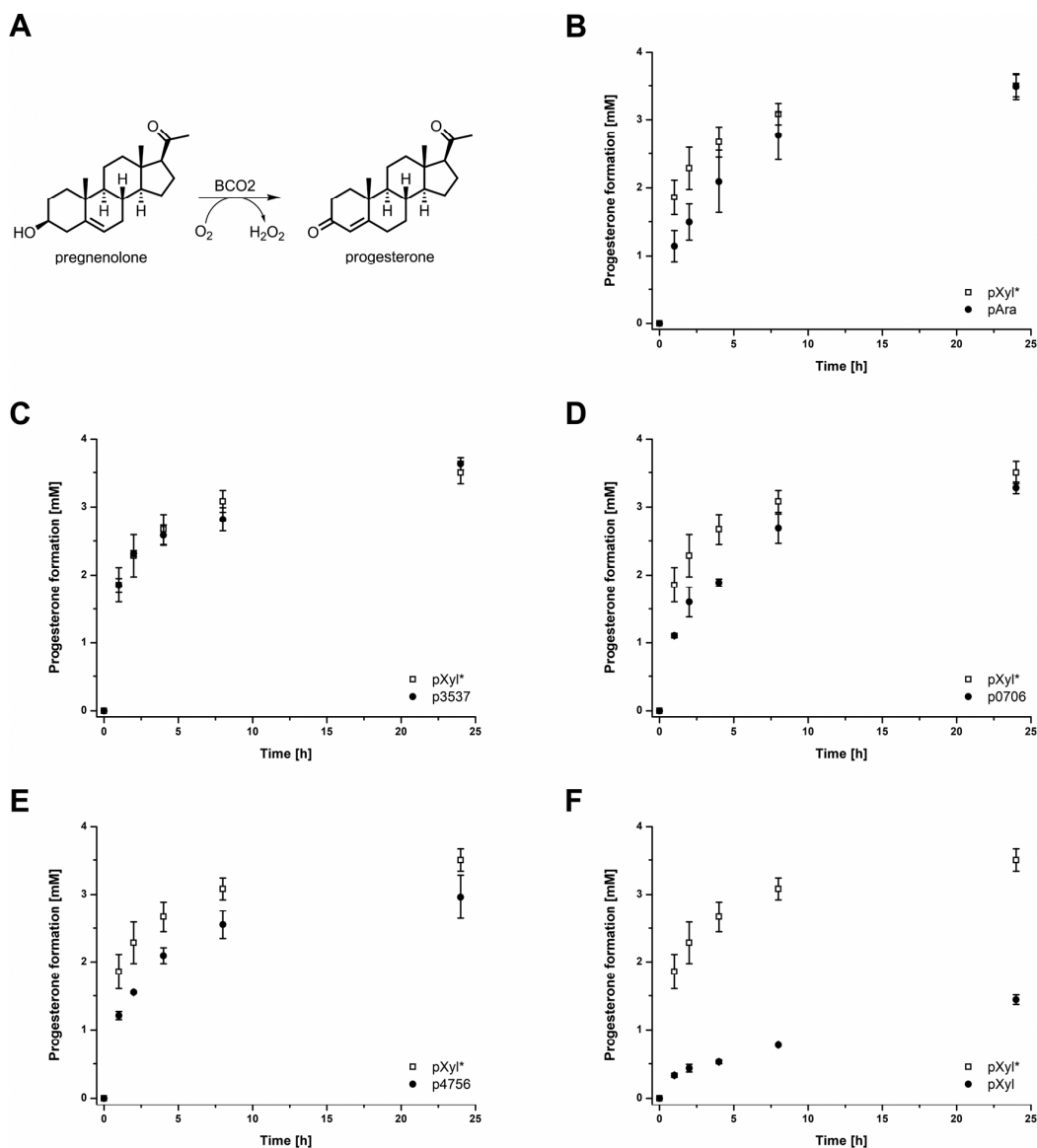


Fig. 6. *In-vivo* conversion of pregnenolone to progesterone with resting cells of the *B. megaterium* strain GHH8 (50 g/l). **A:** Activity of cholesterol oxidase II (BCO2) with pregnenolone as substrate. **B-F:** Heterologous expression of BCO2 under the control of the arabinose inducible promoter pAra, late exponential growth phase promoter p3537, early exponential growth phase promoters p0706 and p4756, as well as the wildtype xylose inducible promoter pXyl. The corresponding oxidase activities were compared to the optimized xylose inducible promoter pXyl*. Experiments were performed in biological triplicates ($n = 3$).

Author's contribution

P.H performed all experiments except for microarray design and transcriptome analyses. P.H analyzed and interpreted the data. C.M participated in the cloning and characterization of the promoters. P.H wrote the manuscript. M.S designed the microarrays and carried out transcriptome analyses. F.H and R.B supervised the experiments, participated in the interpretation and discussion of the data as well as the writing of the manuscript.

Funding

This work was supported by a grant of the Deutsche Bundesstiftung Umwelt (DBU).

Acknowledgements

We would like to thank Antje Eiden-Plach and Birgit Heider-Lips for their assistance with the bioreactor and preparation of RNA.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.jbiotec.2019.01.018>.

References

- Abdulmughni, A., Jóźwik, I.K., Putkaradze, N., Brill, E., Zapp, J., Thunnissen, A.-M.W.H., Hannemann, F., Bernhardt, R., 2017. Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D3hydroxylase. *J. Biotechnol.* 243, 38–47.
- Arnold, T.E., Yu, J., Belasco, J.G., 1998. mRNA stabilization by the ompA 5' untranslated

- region: two protective elements hinder distinct pathways for mRNA degradation. *RNA* 4, 319–330.
- Baim, S.B., Sherman, F., 1988. mRNA structures influencing translation in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 8, 1591–1601.
- Baneyx, F., 1999. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* 10, 411–421.
- Bäumchen, C., Roth, A.H.F.J., Biedendieck, R., Malten, M., Follmann, M., Sahn, H., Bringer-Meyer, S., Jahn, D., 2007. D-Mannitol production by resting state whole cell biotransformation of D-fructose by heterologous mannitol and formate dehydrogenase gene expression in *Bacillus megaterium*. *Biotechnol. J.* 2, 1408–1416.
- Benjamini, Y., Hochberg, Y., 2000. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* 25, 60–83.
- Biedendieck, R., Gamer, M., Jaensch, L., Meyer, S., Rohde, M., Deckwer, W.-D., Jahn, D., 2007. A sucrose-inducible promoter system for the intra- and extracellular protein production in *Bacillus megaterium*. *J. Biotechnol.* 132, 426–430.
- Biedendieck, R., Malten, M., Barg, H., Bunk, B., Martens, J.-H., Deery, E., Leech, H., Warren, M.J., Jahn, D., 2010. Metabolic engineering of cobalamin (vitamin B12) production in *Bacillus megaterium*. *Microb. Biotechnol.* 3, 24–37.
- Bland, C., Newsome, A.S., Markovets, A.A., 2010. Promoter prediction in *E. coli* based on SIDD profiles and Artificial Neural Networks. *BMC Bioinformatics* 11, S17.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- β -boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl. Microbiol. Biotechnol.* 93, 1135–1146.
- Briand, L., Marcion, G., Kriznik, A., Heydel, J.M., Artur, Y., Garrido, C., Seigneuric, R., Neiers, F., 2016. A self-inducible heterologous protein expression system in *Escherichia coli*. *Sci. Rep.* 6, 33037.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2014. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1701–1717.
- Brown, B.J., Carlton, B.C., 1980. Plasmid-mediated transformation in *Bacillus megaterium*. *J. Bacteriol.* 142, 508–512.
- Burger, S., Tatge, H., Hofmann, F., Genth, H., Just, I., Gerhard, R., 2003. Expression of recombinant *Clostridium difficile* toxin A using the *Bacillus megaterium* system. *Biochem. Biophys. Res. Commun.* 307, 584–588.
- Coulombe, R., Yue, K.Q., Ghisla, S., Vrieling, A., 2001. Oxygen access to the active site of cholesterol oxidase through a narrow channel is gated by an Arg-Glu pair. *J. Biol. Chem.* 276, 30435–30441.
- Croteau, N., Vrieling, A., 1996. Crystallization and preliminary X-ray analysis of cholesterol oxidase from *Brevibacterium sterolicum* containing covalently bound FAD. *J. Struct. Biol.* 116, 317–319.
- de Jong, A., Pietersma, H., Cordes, M., Kuipers, O.P., Kok, J., 2012. PePPER: a webserver for prediction of prokaryote promoter elements and regulons. *BMC Genomics* 13, 299.
- Di Gennaro, P., Ferrara, S., Bestetti, G., Sello, G., Solera, D., Galli, E., Renzi, F., Bertoni, G., 2008. Novel auto-inducing expression systems for the development of whole-cell biocatalysts. *Appl. Microbiol. Biotechnol.* 79, 617–625.
- Dvir, S., Velten, L., Sharon, E., Zeevi, D., Carey, L.B., Weinberger, A., Segal, E., 2013. Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 110, E2792–2801.
- Ehrhardt, M., Gerber, A., Hannemann, F., Bernhardt, R., 2016a. Expression of human CYP27A1 in *B. megaterium* for the efficient hydroxylation of cholesterol, vitamin D3 and 7-dehydrocholesterol. *J. Biotechnol.* 218, 34–40.
- Ehrhardt, M., Gerber, A., Zapp, J., Hannemann, F., Bernhardt, R., 2016b. Human CYP27A1 catalyzes hydroxylation of β -sitosterol and ergosterol. *Biol. Chem.* 397, 513–518.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin, M., Elbourne, L.D.H., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukyanchuk, V., Ball, B., Ul-Qamar, R., Seibel, J., Bremer, E., Jahn, D., Ravel, J., Vary, P.S., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J. Bacteriol.* 193, 4199–4213.
- Feklistov, A., 2013. RNA polymerase: in search of promoters. *Ann. N. Y. Acad. Sci.* 1293, 25–32.
- Fickett, J.W., Hatziigeorgiou, A.G., 1997. Eukaryotic promoter recognition. *Genome Res.* 7, 861–878.
- Gamer, M., Fröde, D., Biedendieck, R., Stammen, S., Jahn, D., 2009. A T7 RNA polymerase-dependent gene expression system for *Bacillus megaterium*. *Appl. Microbiol. Biotechnol.* 82, 1195–1203.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Fact.* 14, 107.
- Gerber, A., Milhim, M., Hartz, P., Zapp, J., Bernhardt, R., 2016. Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progesterones 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one. *Metab. Eng.* 36, 19–27.
- Ghosh, T., Bose, D., Zhang, X., 2010. Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol. Rev.* 34, 611–627.
- Giacalone, M.J., Gentile, A.M., Lovitt, B.T., Berkley, N.L., Gunderson, C.W., Surber, M.W., 2006. Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *BioTechniques* 40, 355–364.
- Guan, C., Cui, W., Cheng, J., Zhou, L., Guo, J., Hu, X., Xiao, G., Zhou, Z., 2015. Construction and development of an auto-regulatory gene expression system in *Bacillus subtilis*. *Microb. Cell Fact.* 14, 150.
- Guan, C., Cui, W., Cheng, J., Zhou, L., Liu, Z., Zhou, Z., 2016. Development of an efficient autoinducible expression system by promoter engineering in *Bacillus subtilis*. *Microb. Cell Fact.* 15, 66.
- Harley, C.B., Reynolds, R.P., 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* 15, 2343–2361.
- Hartz, P., Milhim, M., Trenkamp, S., Bernhardt, R., Hannemann, F., 2018. Characterization and engineering of a carotenoid biosynthesis operon from *Bacillus megaterium*. *Metab. Eng.* 49, 47–58.
- Ishii, T., 2001. DBTBS: a database of *Bacillus subtilis* promoters and transcription factors. *Nucleic Acids Res.* 29, 278–280.
- Jones, J.A., Vernacchio, V.R., Lachance, D.M., Lebovich, M., Fu, L., Shirke, A.N., Schultz, V.L., Cress, B., Linhardt, R.J., Koffas, M.A.G., 2015. ePathOptimize: a combinatorial approach for transcriptional balancing of metabolic pathways. *Sci. Rep.* 5, 11301.
- Joseph, B.C., Pichaimuthu, S., Srimeenakshi, S., 2015. An overview of the parameters for recombinant protein expression in *Escherichia coli*. *J. Cell Sci. Ther.* 06.
- Kaberdin, V.R., Bläsi, U., 2006. Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol. Rev.* 30, 967–979.
- Kiss, F.M., Khatri, Y., Zapp, J., Bernhardt, R., 2015. Identification of new substrates for the CYP106A1-mediated 11-oxidation and investigation of the reaction mechanism. *FEB Lett.* 589, 2320–2326.
- Klaff, P., Riessner, D., Steger, G., 1996. RNA structure and the regulation of gene expression. *Plant Mol. Biol.* 32, 89–106.
- Korneli, C., Biedendieck, R., David, F., Jahn, D., Wittmann, C., 2013. High yield production of extracellular recombinant levansucrase by *Bacillus megaterium*. *Appl. Microbiol. Biotechnol.* 97, 3343–3353.
- Lee, G.-Y., Kim, D.-H., Kim, D., Ahn, T., Yun, C.-H., 2015. Functional characterization of steroid hydroxylase CYP106A1 derived from *Bacillus megaterium*. *Arch. Pharm. Res.* 38, 98–107.
- Lisser, S., Margalit, H., 1993. Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res.* 21, 1507–1516.
- Liu, L., Li, Y., Zhang, J., Zou, W., Zhou, Z., Liu, J., Li, X., Wang, L., Chen, J., 2011. Complete genome sequence of the industrial strain *Bacillus megaterium* WSH-002. *J. Bacteriol.* 193, 6389–6390.
- MacLellan, S.R., MacLean, A.M., Finan, T.M., 2006. Promoter prediction in the rhizobia. *Microbiology (Read. Engl.)* 152, 1751–1763.
- Milhim, M., Gerber, A., Neunzig, J., Hannemann, F., Bernhardt, R., 2016a. A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase. *J. Biotechnol.* 231, 83–94.
- Milhim, M., Putkaradze, N., Abdulmughni, A., Kern, F., Hartz, P., Bernhardt, R., 2016b. Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. *J. Biotechnol.* 240, 68–75.
- Ming, Y.M., Wei, Z.W., Lin, C.Y., Sheng, G.Y., 2010. Development of a *Bacillus subtilis* expression system using the improved Pglv promoter. *Microb. Cell Fact.* 9, 55.
- Montigny, C., Penin, F., Lethias, C., Falson, P., 2004. Overcoming the toxicity of membrane peptide expression in bacteria by upstream insertion of Asp-Pro sequence. *Biochim. Biophys. Acta (BBA) – Biomembr.* 1660, 53–65.
- Moore, S.J., Mayer, M.J., Biedendieck, R., Deery, E., Warren, M.J., 2014. Towards a cell factory for vitamin B12 production in *Bacillus megaterium*: bypassing of the cobalamin riboswitch control elements. *N. Biotechnol.* 31, 553–561.
- Niepel, M., Ling, J., Gallie, D.R., 1999. Secondary structure in the 5'-leader or 3'-untranslated region reduces protein yield but does not affect the functional interaction between the 5'-cap and the poly(A) tail. *FEB Lett.* 462, 79–84.
- Nocadello, S., Swennen, E.F., 2012. The new pLAI (lux regulon based auto-inducible) expression system for recombinant protein production in *Escherichia coli*. *Microb. Cell Fact.* 11, 3.
- Panahi, R., Vashghani-Farahani, E., Shojasoadati, S.A., Bamba, B., 2014. Auto-inducible expression system based on the SigB-dependent ohrB promoter in *Bacillus subtilis*. *Mol. Biol. (Mosk.)* 48, 970–976.
- Pitera, D.J., Paddon, C.J., Newman, J.D., Keasling, J.D., 2007. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab. Eng.* 9, 193–207.
- Pollegioni, L., 2009. Cholesterol oxidase: a model flavoprotein oxidase and a biotechnological tool. *FEB J.* 276, 6825.
- Rygu, T., Hillen, W., 1991. Inducible high-level expression of heterologous genes in *Bacillus megaterium* using the regulatory elements of the xylose-utilization operon. *Appl. Microbiol. Biotechnol.* 35, 594–599.
- Rygu, T., Scheler, A., Allmansberger, R., Hillen, W., 1991. Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. *Arch. Microbiol.* 155, 535–542.
- Saecker, R.M., Record, M.T., Dehaseth, P.L., 2011. Mechanism of bacterial transcription initiation: RNA polymerase – promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol.* 412, 754–771.
- Sanderson, J.T., 2006. The steroid hormone biosynthesis pathway as a target for endocrine-disrupting chemicals. *Toxicol. Sci.* 94, 3–21.
- Schmidt, S., Wolf, N., Strey, J., Nahrstedt, H., Meinhardt, F., Waldeck, J., 2005. Test systems to study transcriptional regulation and promoter activity in *Bacillus megaterium*. *Appl. Microbiol. Biotechnol.* 68, 647–655.
- Seshasayee, A.S.N., Bertone, P., Fraser, G.M., Luscombe, N.M., 2006. Transcriptional regulatory networks in bacteria: from input signals to output responses. *Curr. Opin. Microbiol.* 9, 511–519.
- Shimada, N., Hasegawa, S., Harada, T., Tomisawa, T., Fujii, A., Takita, T., 1986. Oxetanocin, a novel nucleoside from bacteria. *J. Antibiot.* 39, 1623–1625.
- Stammen, S., Müller, B.K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E., Jahn, D., 2010. High-yield intra- and extracellular protein production using *Bacillus megaterium*. *Appl. Environ. Microbiol.* 76, 4037–4046.
- Stenger, B., Gerber, A., Bernhardt, R., Hannemann, F., 2018. Functionalized poly(3-hydroxybutyric acid) bodies as new in vitro biocatalysts. *Biochim. Biophys. Acta* 1866, 52–59.

- Studier, F.W., 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* 41, 207–234.
- Vary, P.S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W.-D., Jahn, D., 2007. *Bacillus megaterium*-from simple soil bacterium to industrial protein production host. *Appl. Microbiol. Biotechnol.* 76, 957–967.
- Volontè, F., Pollegioni, L., Molla, G., Frattini, L., Marinelli, F., Piubelli, L., 2010. Production of recombinant cholesterol oxidase containing covalently bound FAD in *Escherichia coli*. *BMC Biotechnol.* 10, 33.
- Wang, W., Hollmann, R., Fürch, T., Nimtz, M., Malten, M., Jahn, D., Deckwer, W.-D., 2005. Proteome analysis of a recombinant *Bacillus megaterium* strain during heterologous production of a glucosyltransferase. *Proteome Sci.* 3, 4.
- Wilkinson, C.J., Hughes-Thomas, Z.A., Martin, C.J., Böhm, I., Mironenko, T., Deacon, M., Wheatcroft, M., Wirtz, G., Staunton, J., Leadlay, P.F., 2002. Increasing the efficiency of heterologous promoters in actinomycetes. *J. Mol. Microbiol. Biotechnol.* 4, 417–426.
- Wittchen, K.D., Meinhardt, F., 1995. Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl. Microbiol. Biotechnol.* 42, 871–877.
- Wolf, J.B., Brey, R.N., 1986. Isolation and genetic characterizations of *Bacillus megaterium* cobalamin biosynthesis-deficient mutants. *J. Bacteriol.* 166, 51–58.
- Won, H.-H., Kim, M.-J., Kim, S., Kim, J.-W., 2008. EnsemPro: an ensemble approach to predicting transcription start sites in human genomic DNA sequences. *Genomics* 91, 259–266.
- Xu, P., Li, L., Zhang, F., Stephanopoulos, G., Koffas, M., 2014. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proc. Natl. Acad. Sci.* 111, 11299–11304.
- Yu, F., Harada, H., Yamasaki, K., Okamoto, S., Hirase, S., Tanaka, Y., Misawa, N., Utsumi, R., 2008. Isolation and functional characterization of a β -eudesmol synthase, a new sesquiterpene synthase from *Zingiber zerumbet* Smith. *FEBS Lett.* 582, 565–572.
- Yu, X., Xu, J., Liu, X., Chu, X., Wang, P., Tian, J., Wu, N., Fan, Y., 2015. Identification of a highly efficient stationary phase promoter in *Bacillus subtilis*. *Sci. Rep.* 5, 18405.

2.1 (Hartz et al., 2019)

Supplemental information

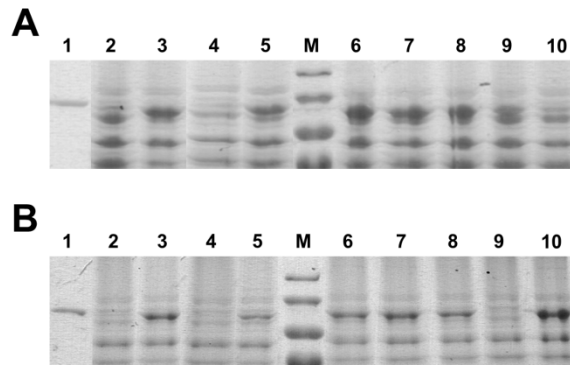
Expanding the promoter toolbox of *Bacillus megaterium*.

Hartz, P., Mattes, C., Schad, M., Bernhardt, R and Hannemann, F.

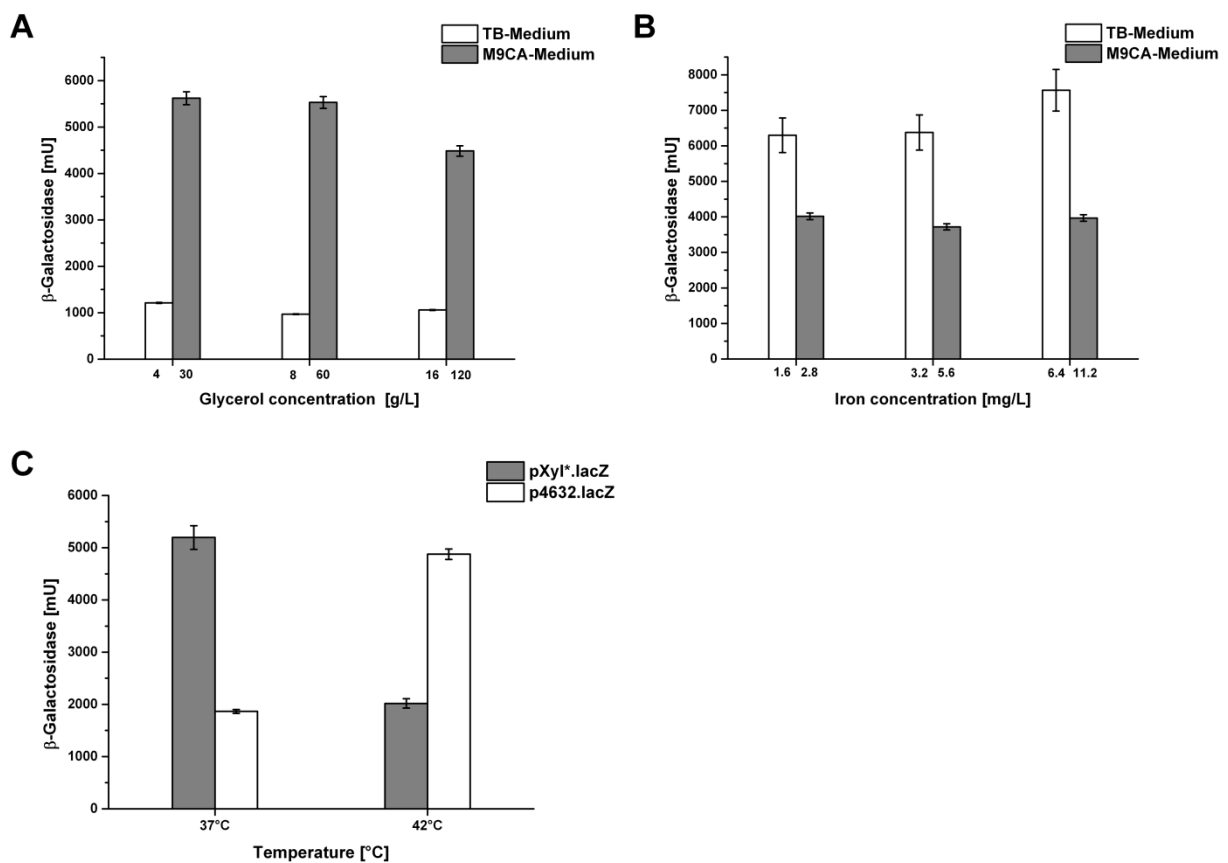
Journal of Biotechnology. 2019 Jan; 294:38-44

DOI: 10.1016/j.jbiotec.2019.01.018

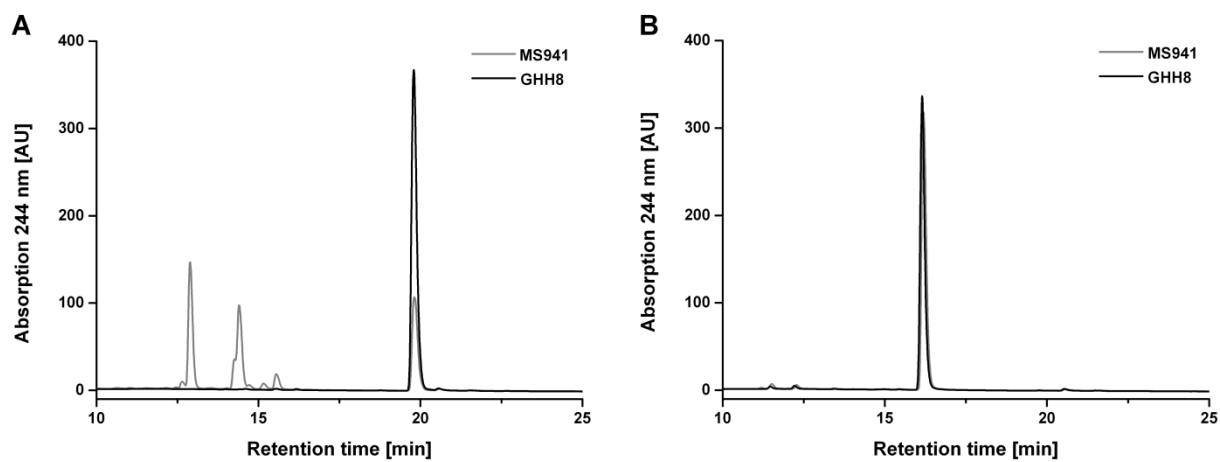
Reprinted with permission of Journal of Biotechnology. All rights reserved.



Supplemental Figure 1: SDS-PAGE analysis of β -galactosidase expression levels with the novel promoters in *B. megaterium* strain GHH9 cultivated in **A:** TB-medium; **B:** modified minimal medium (M9CA). **Lane 1:** Purified β -Galactosidase standard (116 kDa) of β -Gal Assay Kit (Genlantis). **Lane 2 and 4:** pXyl*.lacZ and pAra.lacZ without the corresponding inducer. **Lane 3 and 5:** pXyl*.lacZ and pAra.lacZ with the corresponding inducer (5mg/ml). **Lane 6-10:** p3537.lacZ, p0706.lacZ, p4756.lacZ, p4632.lacZ and p2386.lacZ, respectively.



Supplemental Figure 2: Cultivation condition dependent expression of β -galactosidase. LacZ activities were measured in transformed *B. megaterium* strain GHH9 after 21 h of cultivation. **A:** Influence of glycerol concentrations on the β -galactosidase expression under the control of promoter p2386. **B:** Effect of iron concentrations on the β -galactosidase expression driven by promoter p3537. **C:** Impact of increased temperature on the putative heat inducible promoter p4632. Experiments were performed in biological triplicates (n=3).



Supplemental Figure 3: *In-vivo* conversions with different *B. megaterium* strains. **A:** Conversion of progesterone. **B:** Conversion of DOC as internal standard (IS).

Supplemental Table 1. List of bacterial strains and vectors used in this study.

Bacterial strain	Description	Reference
<i>E. coli</i> Top10	F ⁻ <i>mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\Phi 80lacZ\Delta M15 \Delta lacX74 recA1 araD139$ $\Delta(ara leu) 7697 galU galK rpsL$ (StrR) <i>endA1 nupG</i>	Invitrogen (Karlsruhe, Germany)
<i>B. megaterium</i> MS941	Mutant of DSM319; $\Delta nprM$ (extracellular protease)	Wittchen and Meinhardt (1995)
<i>B. megaterium</i> GHH1	Mutant of MS941; Δupp (selection marker)	Gerber <i>et al.</i> (2015)
<i>B. megaterium</i> GHH8	Mutant of GHH1; $\Delta cyp106A1$; ΔBMD_0912 ; ΔBMD_1595 ; ΔBMD_1068 ; ΔBMD_3715 (20 α HSDs)	Gerber <i>et al.</i> (2016)
<i>B. megaterium</i> GHH9	Mutant of GHH1; ΔBMD_2126 (<i>lacZ</i>)	This study
Plasmid name	Description	Reference
pSMF2.1	Shuttle vector for xylose inducible protein expression with optimized P _{xyIA}	Bleif <i>et al.</i> (2012), based on Stammen <i>et al.</i> (2010)
pXyl*.lacZ	Derivative of pSMF2.1 for xylose-inducible expression of <i>lacZ</i>	This study
pAra.lacZ	Derivative of pSMF2.1 for arabinose-inducible expression of <i>lacZ</i>	This study
pSuc.lacZ	Derivative of pSMF2.1 for sucrose-inducible expression of <i>lacZ</i>	This study
pBga.lacZ	Derivative of pSMF2.1 for lactose/IPTG-inducible expression of <i>lacZ</i>	This study
pGal.lacZ	Derivative of pSMF2.1 for galactose-inducible expression of <i>lacZ</i>	This study
p0462.lacZ	Derivative of pSMF2.1 for protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_0462</i>	This study
p0551.lacZ	Derivative of pSMF2.1 for protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_0551</i>	This study
p1081.lacZ	Derivative of pSMF2.1 for protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_1081</i>	This study
p2097.lacZ	Derivative of pSMF2.1 for protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_2097</i>	This study
p2948.lacZ	Derivative of pSMF2.1 for protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_0462</i>	This study
p0706.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_0706</i>	This study
p4756.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_4756</i>	This study
p0123.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_0123</i>	This study
p3115.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_3115</i>	This study
p3537.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_3537</i>	This study

p0450.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_0450</i>	This study
p1214.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_1214</i>	This study
p2386.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_2386</i>	This study
p2618.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_2618</i>	This study
p4632.lacZ	Derivative of pSMF2.1 for constitutive protein expression of <i>lacZ</i> under the control of the putative promoter of <i>BMD_4632</i>	This study
pUCTV2_Upp	Backbone vector for gene deletion	Gerber <i>et al.</i> (2015), based on Wittchen and Meinhardt (1995)
pUCTV2_Upp_Δ <i>BMD_2126</i>	Genomic deletion of <i>BMD_2126</i> (<i>lacZ</i>)	This study

Supplemental Table 2. List of primers used in this study. Restriction sites are shown in bold letters.

Ribosomal binding sites are underlined.

Name	Sequence (5'-3')	Description
Δ2126-A-for	TGTACGA ATTC GTCTCGAAGCATATTGCTTCGACT TTTTTTGC	Amplification of flanking region A upstream of <i>BMD_2126</i> and insertion in knockout vector pUCTV2_Upp with <i>EcoRI</i>
Δ2126-A-rev	CGGCAAGAA ATTT CATTATACAGTTTGA ACTGG AGGCTTACGATC ATTTT TAC	Overlapping extension PCR with flanking region B
Δ2126-B-for	ATCGTAAGCCTCCAGTTCAAAGCTGTATAATGA A ATTTCT TGCCGG	Overlapping extension PCR with flanking region A
Δ2126-B-rev	TGTACGA ATTC TATCCCTCCTTTTTGTTTCTTCTC TC	Amplification of flanking region A upstream of <i>BMD_2126</i> and insertion in knockout vector pUCTV2_Upp with <i>EcoRI</i>
lacZ-screen-for	AAGCCATTGTTCTTGC GT CAGCAGAACC	Screening for genomic deletion of <i>BMD_2126</i>
lacZ-screen-rev	CCGGAAAGTACATACGGTCCAC ATTACTTCC	Screening for genomic deletion of <i>BMD_2126</i>
Ec-lacZ-for	TGTAC ACTAGT <u>AAATCAAGGAGGTGAATGTACA</u> GTGGTTTTACAACGTCGTGAC	Amplification of <i>E. coli lacZ</i> gene and insertion in pSMF2.1 vector with <i>SpeI</i>
Ec-lacZ-rev	TGTAC GGTACCTT ATTTTTGACACCAGACCAAC TGGTAATGG	Amplification of <i>E. coli lacZ</i> gene and insertion in pSMF2.1 vector with <i>KpnI</i>
BCO2-for	TGTACA ACTAGT <u>AAATCAAGGAGGTGAATGTAC</u> <u>ATG</u>	Cloning of cholesterol oxidase II (BCO2) in pSMF2.1 vector with <i>SpeI</i>
BCO2-rev	TGTAC AGGTACCTT ATGTAGAATCTAATCCTAA TAAAGGATTAGG	Cloning of cholesterol oxidase II (BCO2) in pSMF2.1 vector with <i>KpnI</i>
pAra-for	GTCTGTACGTT CCTT AAGTTAACGAACTTCATCA CGCTTCGC ATTTAACTC	Cloning of putative promoter sequence upstream of <i>BMD_3532</i> incl. <i>BMD_3531</i> as putative repressor

pAra-rev	GTTTATCCATCAGCTAGCCAATACCTATAGCAGT AAGGAAGCGCTTTCC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_3532</i> incl. <i>BMD_3531</i> as putative repressor
pBga-for	TGTACCTTAAGCCTAGATAAAACCCGATATATG TAGC	Cloning of putative promoter sequence upstream of <i>BMD_2126</i> incl. <i>BMD_2127</i> as putative repressor with <i>AflIII</i>
pBga-rev	TGTACACTAGTGTGCGATCCCTCATTCTAAAGTA TTAG	Cloning of putative promoter sequence upstream of <i>BMD_2126</i> incl. <i>BMD_2127</i> as putative repressor with <i>SpeI</i>
pGal-for	GTCTGTACGTTCCCTTAAGTTATTTTGGCAATGCA AAGTGCTCATC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_0689</i> incl. <i>BMD_0688</i> as putative repressor
pGal-rev	ACCTCCTTGATTTACTAGTTTAATTAAGTAGTTT TTTGGAAACGGATTCATATGTAC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_0689</i> incl. <i>BMD_0688</i> as putative repressor
pSuc-for	GTCTGTACGTTCCCTTAAGTTACGTTGTTTGTCCCTT CTAGAAG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_1565</i> incl. <i>BMD_1564</i> as putative repressor
pSuc-rev	ACCTCCTTGATTTACTAGTTTAATTAAGATATAA AATCTTTTATCTAACAAAATGAAAGCAC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_1565</i> incl. <i>BMD_1564</i> as putative repressor
p0123-for	CTGTACGTTCCCTTAAGATCAACGTAAATTTGGAA ATAAATCGGTGTTGAC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_0123</i>
p0123-rev	CTCCTTGATTTACTAGTTTAATTAACCTTACACC TCCTGTAGATTTCC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_0123</i>
p0450-for	CGTCTGTACGTTCCCTTAAGTTGTCTAGCTTTTTTTT ATTTAAGGGTTG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_0450</i>
p0450-rev	TCCTTGATTTACTAGTTTAATTAATACATATTCC TCCGGTTTCG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_0450</i>
p0462-for	TGTACCTTAAGAGCCGTTGAGAAATCAACGG	Cloning of putative promoter sequence upstream of <i>BMD_0462</i> with <i>AflIII</i>
p0462-rev	TGTACGCTAGCTAAAAACGCTCCTTTCTCTATC ATTC	Cloning of putative promoter sequence upstream of <i>BMD_0462</i> with <i>NheI</i>
p0551-for	TGTACCTTAAGTCATATAAAAAGCAGAATGATG CGG	Cloning of putative promoter sequence upstream of <i>BMD_0551</i> with <i>AflIII</i>
p0551-rev	TGTACGCTAGCATGATCGAAACCTTTCTATTTCG G	Cloning of putative promoter sequence upstream of <i>BMD_0551</i> with <i>NheI</i>
p0706-for	TGTACCTTAAGCAAATCTTCCTTTCATCTGCACA TATATGATGC	Cloning of putative promoter sequence upstream of <i>BMD_0706</i> with <i>AflIII</i>
p0706-rev	TGTACACTAGTTGCATGTTACCCCTATTAAATT TTGTTTG	Cloning of putative promoter sequence upstream of <i>BMD_0706</i> with <i>SpeI</i>
p1081-for	TGTACCTTAAGTGTTTAAAAGAGGCTAGAACAT AAG	Cloning of putative promoter sequence upstream of <i>BMD_1081</i> with <i>AflIII</i>

p1081-rev	TGTACGCTAGCGCTATCCCTCCGTCCTTTG	Cloning of putative promoter sequence upstream of <i>BMD_1081</i> with <i>NheI</i>
p1214-for	TAAGCCGTCTGTACGTTCTTAAGACAAAGATTT AGAATTGTTTATTTTG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_1214</i>
p1214-rev	TCCTTGATTTACTAGTTTAATTA AAACTCCATCT CCTTTCTTGTGTACGT	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_1214</i>
p2097-for	TGTACCTTAAGTTTGTCTCTCTTCTTCTATTT ATATGGG	Cloning of putative promoter sequence upstream of <i>BMD_2097</i> with <i>AflIII</i>
p2097-rev	TGTACGCTAGCCAAAATCCTCCTTTAATATGAG GAAGTC	Cloning of putative promoter sequence upstream of <i>BMD_2097</i> with <i>NheI</i>
p2386-for	CGTCTGTACGTTCTTAAGAAGAAGTTCTCCTTT CCAAATAGGGCT	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_2386</i>
p2386-rev	ACCTCCTTGATTTACTAGTTTAATTAATTTGTAC TCTCCTCTGATTATC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_2386</i>
p2618-for	GTCTGTACGTTCTTAAGGTTATATCTCCTTACA TAAGATG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_2618</i>
p2618-rev	ACCTCCTTGATTTACTAGTTTAATTAAGTCGAT CCTCCGTCAC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_2618</i>
p2948-for	TGTACCTTAAGGAAGTACTTCTTAAGAAAAGGC GG	Cloning of putative promoter sequence upstream of <i>BMD_2948</i> with <i>AflIII</i>
p2948-rev	TGTACGCTAGCTTCCACTCTCCTCATTTTTAGT TG	Cloning of putative promoter sequence upstream of <i>BMD_2948</i> with <i>NheI</i>
p3115-for	GTCTGTACGTTCTTAAGATCAACGTGAGCGCTT GTATGCTTTCTAAG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_3115</i>
p3115-rev	CTCCTTGATTTACTAGTTTAATTA AAAATTGTATT CTCTCCTTTTTTCTATAC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_3115</i>
p3537-for	GTCTGTACGTTCTTAAGATCAACGAAAAGCCT CCTAAAGTGATTTTTTGAGG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_3537</i>
p3537-rev	CTCCTTGATTTACTAGTTTAATTAACGTCATCAC CTACTTTTTTCAC	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_3537</i>
p4632-for	CTGTACGTTCTTAAGATCAACGTCACCTTCACC TCACACG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_4632</i>
p4632-rev	CTCCTTGATTTACTAGTTTAATTAATACTTCAGC TCCCTTTTTTATG	Restriction site independent cloning of putative promoter sequence upstream of <i>BMD_4632</i>
p4756-for	TGTACCTTAAGCTTATTTATTTATGAAATTAACC GAACTTTTTTCGTAG	Cloning of putative promoter sequence upstream of <i>BMD_4756</i> with <i>AflIII</i>
p4756-rev	TGTACACTAGTAAATCTCCACTCCTTTACAGAA AAAATTCC	Cloning of putative promoter sequence upstream of <i>BMD_4756</i> with <i>SpeI</i>

2.2 (Gerber et al., 2016)

Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one.

Gerber, A., Milhim, M., **Hartz, P.**, Zapp, J. and Bernhardt, R.

Metabolic Engineering. 2016 Feb; 36:19-27

DOI: 10.1016/j.ymben.2016.02.010

Reprinted with permission of Metabolic Engineering. All rights reserved.



Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens $17\alpha,20\beta$ -di- and $17\alpha,20\beta,21\alpha$ -trihydroxy-4-pregnen-3-one



Adrian Gerber^a, Mohammed Milhim^a, Philip Hartz^a, Josef Zapp^b, Rita Bernhardt^{a,*}

^a Saarland University, Institute of Biochemistry, Campus B2.2, 66123 Saarbrücken, Germany

^b Saarland University, Institute of Pharmaceutical Biology, Campus C2.3, 66123 Saarbrücken, Germany

ARTICLE INFO

Article history:

Received 19 October 2015

Received in revised form

1 February 2016

Accepted 23 February 2016

Available online 11 March 2016

Keywords:

Bacillus megaterium

Hydroxysteroid dehydrogenase

Steroid hormones

$17\alpha,20\beta$ -Dihydroxy-4-pregnen-3-one

$17\alpha,20\beta,21\alpha$ -Trihydroxy-4-pregnen-3-one

MIH

ABSTRACT

$17\alpha,20\beta$ -Dihydroxy-4-pregnen-3-one ($17\alpha,20\beta$ DiOH-P) and $17\alpha,20\beta,21\alpha$ -trihydroxy-4-pregnen-3-one (20β OH-RSS) are the critical hormones required for oocyte maturation in fish. We utilized *B. megaterium*'s endogenous 20β -hydroxysteroid dehydrogenase (20β HSD) for the efficient production of both progestogens after genetically modifying the microorganism to reduce side-product formation. First, the gene encoding the autologous cytochrome P450 CYP106A1 was deleted, resulting in a strain devoid of any steroid hydroxylation activity. Cultivation of this strain in the presence of 17α -hydroxyprogesterone (17α OH-P) led to the formation of $17\alpha,20\alpha$ -dihydroxy-4-pregnen-3-one ($17\alpha,20\alpha$ DiOH-P) as a major and $17\alpha,20\beta$ DiOH-P as a minor product. Four enzymes were identified as 20α HSDs and their genes deleted to yield a strain with no 20α HSD activity. The 3-oxoacyl-(acyl-carrier-protein) reductase FabG was found to exhibit 20β HSD-activity and overexpressed to create a biocatalyst yielding 0.22 g/L $17\alpha,20\beta$ DiOH-P and 0.34 g/L 20β OH-RSS after 8 h using shake-flask cultivation, thus obtaining products that are at least a thousand times more expensive than their substrates.

© 2016 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

1. Introduction

The oocyte maturation inducing hormones (MIHs) $17\alpha,20\beta$ DiOH-P and 20β OH-RSS are the main progestogens in teleosts. In female fishes, MIH production in the ovarian follicle layers is stimulated by gonadotropin synthesized in the pituitary. MIH induces the formation of the maturation promoting factor (MPF) leading to further oocyte maturation (Nagahama and Yamashita, 2008). In addition, MIHs are produced in the testis of male and, among other functions, regulate spermiation and sperm motility (Scott et al., 2010). Both MIHs are produced from

cholesterol in a series of reaction steps involving the cytochromes P450 CYP11A1, CYP17, CYP21 as well as 3β - and 20β HSD.

The administration of MIHs to different species of aquacultured fish has been extensively investigated and shown to greatly increase spawning rates due to, for instance, overcoming ovulation disorder, improving sperm volume and concentration, attracting fish to spawning sites and induction of ovulation (Yamamoto et al., 2015; Martins Pinheiro et al., 2003; Hong et al., 2006; Haider and Rao, 1994; King and Young, 2001; Ohta et al., 1996; Miwa et al., 2001). Induced spawning is of particular interest for economically significant species of fish that do not reproduce spontaneously under captive cultivation conditions (Lee and Yang, 2002). However, the widespread use of both synthetic MIHs in aquacultures has been hindered by their high cost.

Steroid hormones are produced mainly through a combination of chemical and microbial conversion steps, since single-step stereospecific modifications of the unreactive steroid nucleus by chemical means alone are often not possible (Donova and Egorova, 2012). The chemical synthesis of $17\alpha,20\beta$ DiOH-P is tedious, requiring multiple reaction steps to prepare the steroid stereoselectively. One method to produce the progestogen includes the chlorination, dechlorination and reduction of RSS to produce the substance with an overall yield of 64 % (Ouedraogo et al., 2013). Another approach consists of reducing 17α OH-P by NaBH_4 in

Abbreviations: AKR, aldo-keto reductase; 17α OH-P, 17α -hydroxyprogesterone; $17\alpha,20\alpha$ DiOH-P, $17\alpha,20\alpha$ -dihydroxy-4-pregnen-3-one; $17\alpha,20\beta$ DiOH-P, $17\alpha,20\beta$ -dihydroxy-4-pregnen-3-one; 20β OH-RSS, $17\alpha,20\alpha,21\alpha$ -trihydroxy-4-pregnen-3-one; DOC, 11-deoxycorticosterone; FabG, 3-oxoacyl-(acyl-carrier-protein) reductase; HSD, hydroxysteroid dehydrogenase; LB, lysogeny broth; MIH, maturation inducing hormone; MPF, maturation promoting factor; NMR, nuclear magnetic resonance; ORF, open reading frame; PHB, poly(3-hydroxybutyrate); RP-HPLC, reverse phase high performance liquid chromatography; RSS, 11-deoxycortisol; SOE-PCR, splicing by overlapping polymerase chain reaction; TB, terrific broth; Upp, uracil phosphoribosyltransferase

* Corresponding author.

E-mail address: ritabern@mx.uni-saarland.de (R. Bernhardt).

<http://dx.doi.org/10.1016/j.ymben.2016.02.010>

1096-7176/© 2016 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

Table 1

List of *B. megaterium* strains and plasmids used in this study.

Strains or plasmids	Description	Reference
Strains		
MS941	Mutant of DSM319, $\Delta nprM$ (extracellular protease)	Wittchen and Meinhardt (1995)
GHH1	Mutant of MS941, Δupp (selection marker)	Gerber et al. (2015)
GHH2	Mutant of GHH1, $\Delta cyp106A1$	This study
GHH5	Mutant of GHH2, ΔBMD_0912 (20 α HSD)	This study
GHH6	Mutant of GHH5, ΔBMD_1068 (20 α HSD)	This study
GHH7	Mutant of GHH6, ΔBMD_3715 (20 α HSD)	This study
GHH8	Mutant of GHH7, ΔBMD_1595 (20 α HSD)	This study
Plasmids		
pSMF2.1	Backbone vector for protein overexpression	Bleif et al. (2012)
pSMF2.1_0912	Expression of <i>BMD_0912</i> (20 α HSD)	This study
pSMF2.1_1595	Expression of <i>BMD_1595</i> (20 α HSD)	This study
pSMF2.1_1068	Expression of <i>BMD_1068</i> (20 α HSD)	This study
pSMF2.1_3715	Expression of <i>BMD_3715</i> (20 α HSD)	This study
pSMF2.1_FabG	Expression of FabG (20 β HSD)	This study
pUCTV2_Upp	Backbone vector for gene deletion	Gerber et al. (2015), based on Wittchen and Meinhardt (1995)
pUCTV2_Upp_ $\Delta 106A1$	Genomic deletion of <i>cyp106A1</i>	This study
pUCTV2_Upp_ $\Delta 0912$	Genomic deletion of <i>BMD_0912</i>	This study
pUCTV2_Upp_ $\Delta 1595$	Genomic deletion of <i>BMD_1595</i>	This study
pUCTV2_Upp_ $\Delta 1068$	Genomic deletion of <i>BMD_1068</i>	This study
pUCTV2_Upp_ $\Delta 3715$	Genomic deletion of <i>BMD_3715</i>	This study

methanol, resulting in a ~30 % yield of 17 α ,20 β DiOH-P (Kovganko et al., 2001). In microbial fermentations with strains of *B. megaterium*, *Yarrowia lipolytica*, *Saccharomyces cerevisiae* or *Bifidobacterium adolescentis* 17 α ,20 β DiOH-P appeared only as a minor side-product (Shkumatov et al., 2003; Winter et al., 1982; Koshcheenko et al., 1976). Genetically engineered microorganisms are often used to increase the accumulation of desired products by, for instance, deletion of genes to avoid by-product formation (Yao et al., 2014), augmenting the activity of endogenous enzymes by overexpression (Yao et al., 2013) or mutagenesis (Hunter et al., 2011), or increasing the tolerance of the organism towards metabolite and substrate stress (Nicolaou et al., 2010).

In this study, we present a genetically modified *B. megaterium* strain allowing the high yield production of both 17,20 β DiOH-P and 20 β OH-RSS from cheap steroid precursors. This Gram-positive, rod-shaped bacterium has gained considerable interest in recent years as a recombinant expression host due to its high protein production capacity, plasmid stability, ability to take up a variety of hydrophobic substrates as well as its large cell size, which allows detailed microscopic analyses (Vary et al., 2007). Like other *Bacillus* species, *B. megaterium* can be genetically engineered through homologous recombination with exogenous DNA, with or without applying a marker gene (Dong and Zhang, 2014). The classic procedure consists of protoplast transforming cells with the deletion vector, the integration of parts of the vector into the chromosome, screening the target locus for the mutation by PCR and finally curing the vector from the cells (Biedendieck et al., 2011). A more recent approach allows the one-step deletion of genes by using a transconjugation protocol with *Escherichia coli* as plasmid-donor cells (Richhardt et al., 2010). *B. megaterium* has been shown to be particularly suitable for the bioconversion of hydrophobic steroidal compounds, including cholesterol and analogs (Gerber et al., 2015), natural steroid hormones derived from cholesterol such as pregnenolone and dehydroepiandrosterone (Schmitz et al., 2014), synthetic steroid hormones such as prednisolone and dexamethasone (Kiss et al., 2015) and the secosteroid vitamin D3 (Ehrhardt et al., 2016). The genome of *B. megaterium* DSM319, the precursor of the strain used in this study, contains the gene for the cytochrome P450 CYP106A1, resulting in an endogenous steroid hydroxylation activity. Cytochromes P450 form a superfamily of heme-thiolate proteins, which are, in bacteria, mainly involved in the metabolism of xenobiotics and the

production of secondary metabolites (Bernhardt, 2006; Bernhardt and Urlacher, 2014). CYP106A1 is able to convert a wide variety of steroids, including testosterone, progesterone, 17-hydroxyprogesterone, RSS, DOC and cortisol (Kiss et al., 2015).

The aim of this study was to construct a biocatalyst for the production of the fish progestogens 17 α ,20 β DiOH-P and 20 β OH-RSS by utilizing an endogenous 20 β HSD and abolishing side-product formation from the likewise endogenous CYP106A1 and four 20 α HSDs through genetic engineering.

2. Materials and methods

2.1. Materials

20 β OH-RSS (4-pregnen-17 α ,20 β ,21-triol-3-one, catalog ID: Q4080-000), 20 β OH-cortisol (4-pregnen-11 β , 17, 20 β , 21-tetrol-3-one, Q3790-000) and 20 β OH-cortisone (4-pregnen-17, 20 β , 21-triol-3, 11-dione, Q3960-000) were purchased from Steraloids. 17OH-P (4-pregnen-17 α -ol-3,20-dione, H5752) RSS (4-Pregnen-17 α ,21-diol-3,20-dione, R0500), cortisol (4-pregnen-11 β ,17 α ,21-triol-3,20-dione, H4001), cortisone (pregnen-17 α ,21-diol-3,11,20-trione, C2755) and 17 α ,20 α DiOH-P (4-pregnen-17, 20 β -diol-3-one, P6285), xylose (95729), tetracycline (T7660), ethyl acetate (34858) and acetonitrile (34851) were from Sigma-Aldrich. LB broth (244610), yeast extract (212750) and tryptone (211705) were obtained from Becton Dickinson.

2.2. Molecular cloning and gene deletion

All plasmids were constructed by conventional cloning using restriction enzyme digestion and ligation. *E. coli* strain TOP10 (Invitrogen) was used for the propagation of plasmids. *B. megaterium* was transformed according to the PEG-mediated protoplast transformation method (Biedendieck et al., 2011). Table 1 lists all plasmids used in this study. Plasmid maps displaying backbone vectors and the restriction sites used for cloning are depicted in Supplemental Fig. 1. The xylose-inducible vector pSMF2.1 was used as a backbone for the overexpression of genes. ORFs of *BMD_0912*, *BMD_1595*, *BMD_1068*, *BMD_3715* and *FabG* were amplified from genomic DNA of *B. megaterium* MS941 and cloned into pSMF2.1. Genomic DNA was prepared using a genomic DNA

isolation kit (nexttec). For gene deletions, pUCTV2_Upp was used as a backbone vector (Gerber et al., 2015; Wittchen and Meinhardt, 1995). The knockout of *cyp106A1*, *BMD_0912*, *BMD_1595*, *BMD_1068* and *BMD_3715* was carried out as previously described (Gerber et al., 2015). In brief, flanking regions of the target genes were amplified from genomic DNA, fused via splicing by overlapping extension (SOE)-PCR and cloned into pSMF2.1_Upp (Table 1). *B. megaterium* was transformed with the resulting plasmids, grown colonies streaked out on minimal medium agar plates (Gerber et al., 2015) and incubated overnight at 42 °C. Colonies were then plated on minimal medium agar plates containing 1 μM 5-fluorouracil and again incubated overnight. Resulting colonies are plasmid-free and can be screened for target gene deletion by PCR. All primers used in this study are listed in Supplemental Table 1.

2.3. *B. megaterium* cultivation conditions and in vivo substrate conversions

Precultures of *B. megaterium* were started by inoculating LB-medium (25 g/L) with cells from a glycerol stock or colony and incubated overnight at 30 °C and 180 rpm shaking. TB-medium (24 g/L yeast extract, 12 g/L tryptone, 0.4% glycerol, 100 mM potassium phosphate buffer, pH 7.4) was used for the main cultures, which were inoculated 1:100 with the precultures. For cells harboring pSMF2.1 or pUCTV2 derivatives, tetracycline was added to the cultures to a final concentration of 10 μg/mL. Recombinant protein expression was induced at an optical density of ca. 0.4 at 600 nm by adding xylose dissolved in water to a final concentration of 0.5% (w/v). For the assessment of intrinsic CYP106A1, 20αHSD and 20βHSD activities, substrates were added in the early logarithmic growth phase (approximately 2 h after inoculation of the main culture), after being dissolved in ethanol (stock solutions: 14.9 mM progesterone, 15 mM 17OH-P, final ethanol concentration in the cultures: 2%). For in vivo conversions with over-expressed FabG, the substrate was added 20 h after protein induction, to achieve highest possible product yields (stock solutions: 25 mM 17OH-P and RSS, final ethanol concentration in the cultures: 4%).

2.4. Culture sample treatment and RP-HPLC analysis

1 mL culture samples were taken at different time points and extracted twice with equal volumes of ethylacetate. In case of the 20αHSD activity measurements, progesterone was added as an internal standard prior to that to account for extraction loss. After evaporation, the extracts were dissolved in 10% (v/v) acetonitrile in water. Gradient RP-HPLC analysis was carried out using a 125/4 Nucleodur C18 column (Macherey & Nagel) on a device manufactured by Jasco with detection at a wavelength of 240 nm. 10% (v/v) acetonitrile in water was used as mobile phase on channel A, pure acetonitrile on channel B. For all HPLC measurements the following method was used: 0–3.5 min: 100% A, 3.5–22 min: 100% A to 70% A, 22–25 min: 70% A to 0% A, 25–30 min: 0% A, 30–35 min: 100% A. The flow rate was 1 mL/min, column oven temperature was 40 °C. Purification of products for NMR analysis was performed with a VP 250/8 Nucleodur 100-5 C18 ec column, applying the above method with a flow rate of 4 mL/min. In order to quantify product yields, firstly absence of steroid degradation with strain GHH8 (Supplemental Fig. 2) and similarity of molar extinction coefficients of substrates and products (Supplemental Fig. 3) were verified. For the quantification, the product peak area was divided by the sum of product and substrate areas. The resulting quotient was multiplied with the applied substrate concentration. Molar concentrations of the products were then

converted to mass concentrations (mg/L) based on their molecular weight.

2.5. NMR characterization of the metabolites P3–P5

The NMR spectra were recorded in CDCl₃ and DMSO-d₆ (only P3) with a Bruker DRX 500 or a Bruker Avance 500 NMR spectrometer at 298 K. The chemical shifts were relative to CHCl₃ at δ 7.26 (¹H NMR) and CDCl₃ at δ 77.00 (¹³C NMR) or to DMSO at δ 2.50 (¹H NMR) respectively using the standard δ notation in parts per million. The 1D NMR (¹H and ¹³C NMR) and the 2D NMR spectra (gs-HH-COSY and gs-HSQCED) were recorded using the BRUKER pulse program library.

In contrast to S2, 17α-hydroxyprogesterone, the NMR data of its conversion products P3 and P4 lacked a carbonyl function at C-20. Therefore, resonances of an additional secondary hydroxyl group appeared in both spectra sets indicating that P3 (20α-OH) and P4 (20β-OH) were epimers concerning the hydroxyl group at C-20. Their spectra matched perfectly with those of commercially available reference compounds and were in agreement with data in literature (P3: (Wishart et al., 2013), P4: (Hunter and Carragher, 2003)).

P5 was found to be the 20β-hydroxy derivative of S3, 11-deoxycortisol. Its NMR data were in good accordance with those from literature (Hunter and Carragher, 2003) and its ¹H and ¹³C NMR spectra showed identical resonances when compared with those recorded with an authentic sample.

P3: 17α, 20α-Dihydroxy-4-pregnen-3-one (17α,20α,DiOH-P).

¹H NMR (DMSO-d₆, 500 MHz): δ 0.68 (s, 3xH-18), 0.82 (ddd, *J* = 12.4, 10.9 and 4.0 Hz, H-9), 0.96 (qd, *J* = 13.0 and 4.2, H-7a), 1.02 (d, *J* = 6.3 Hz, 3xH-21), 1.09 (m, H-15a), 1.14 (s, 3xH-19), 1.32 (qd, *J* = 13.0 and 4.0 Hz, H-11a), 1.41 (dt, *J* = 13.0 and 4.0 Hz, H-12a), 1.49 (m, H-11b), 1.52 (m, H-8), 1.57 (m, H-16a), 1.58 (m, H-15b), 1.60 (m, H-1a), 1.64 (m, H-12b), 1.70 (m, H-14), 1.78 (m, H-7b), 1.91 (t, *J* = 11.8 Hz, H-16b), 1.95 (ddd, *J* = 13.5, 5.1 and 3.5 Hz, H-1b), 2.14 (dt, *J* = 16.7 and 3.5 Hz, H-2a), 2.24 (ddd, *J* = 14.5, 3.7 and 2.2 Hz, H-6a), 2.39 (ddd, *J* = 16.7, 14.5 and 5.1 Hz, H-2b), 2.40 (m, H-6b), 3.52 (s, OH-17), 3.60 (quint, *J* = 6.3 Hz, H-20), 4.15 (d, *J* = 6.3 Hz, OH-20), 5.62 (br s, H-4). ¹³C NMR (CDCl₃, 500 MHz): δ 0.77 (s, 3xH-18), 0.97 (ddd, *J* = 12.4, 10.7 and 3.8 Hz, H-9), 1.10 (dddd, *J* = 13.7, 12.9, 11.7 and 4.5, H-7a), 1.19 (s, 3xH-19), 1.20 (d, *J* = 6.3 Hz, 3xH-21), 1.22 (m, H-15a), 1.42 (m, H-11a), 1.58 (m, 2H, H-8 and H-12a), 1.61 (m, H-11b), 1.70 (m, H-12b), 1.71 (m, H-1a), 1.74 (m, H-16a), 1.77 (m, H-14), 1.78 (m, H-15b), 1.86 (dddd, *J* = 12.9, 5.6, 3.6 and 2.5, H-7b), 2.02 (ddd, *J* = 13.5, 5.1 and 3.2 Hz, H-1b), 2.06 (m, H-16b), 2.28 (ddd, 14.7, 4.5 and 2.5 Hz, H-6a), 2.35 (dddd, *J* = 17.1, 5.0, 3.2 and 1.0 Hz, H-2a), 2.39 (m, H-6b), 2.41 (ddd, *J* = 17.1, 14.5 and 5.1 Hz, H-2b), 3.86 (q, *J* = 6.3 Hz, H-20), 5.73 (br s, H-4). ¹³C NMR (CDCl₃, 125 MHz): δ 14.06 (CH₃, C-18), 17.38 (CH₃, C-19), 18.52 (CH₃, C-21), 20.49 (CH₂, C-11), 23.34 (CH₂, C-15), 30.94 (CH₂, C-12), 31.95 (CH₂, C-7), 32.87 (CH₂, C-6), 33.93 (CH₂, C-2), 35.48 (CH, C-8), 35.68 (CH₂, C-1), 37.69 (CH₂, C-16), 38.52 (C, C-10), 45.64 (C, C-13), 50.54 (CH, C-14), 53.26 (CH, C-9), 72.28 (CH, C-20), 85.37 (C, C-17), 123.89 (CH, C-4), 171.22 (C, C-5), 199.53 (C, C-3).

P4: 17α, 20β-Dihydroxy-4-pregnen-3-one (17α,20β,DiOH-P).

¹H NMR (CDCl₃, 500 MHz): δ 0.85 (s, 3xH-18), 0.97 (ddd, *J* = 12.5, 10.7 and 4.0 Hz, H-9), 1.08 (dddd, *J* = 13.7, 12.9, 11.7 and 4.5, H-7a), 1.16 (m, H-15a), 1.18 (d, *J* = 6.3 Hz, 3xH-21), 1.19 (s, 3xH-19), 1.43 (m, H-11a), 1.46 (m, H-16a), 1.55 (m, H-8), 1.58 (m, H-11b), 1.59 (m, H-12a), 1.66 (m, H-16b), 1.70 (m, H-1a), 1.71 (m, H-14), 1.74 (m, 2H, H-12b and H-15b), 1.84 (dddd, *J* = 13.0, 5.5, 3.5 and 2.5, H-7b), 2.02 (ddd, *J* = 13.5, 5.1 and 3.1 Hz, H-1b), 2.27 (ddd, 14.7, 4.5 and 2.5 Hz, H-6a), 2.35 (dddd, *J* = 17.0, 4.8, 3.2 and 1.0 Hz, H-2a), 2.40 (m, H-6b), 2.41 (ddd, *J* = 17.0, 14.5 and 5.0 Hz, H-2b), 4.04 (q, *J* = 6.3 Hz, H-20), 5.72 (br s, H-4). ¹³C NMR (CDCl₃, 125 MHz): δ 15.09 (CH₃, C-18), 17.37 (CH₃, C-19), 18.73 (CH₃, C-21), 20.58 (CH₂, C-11),

23.73 (CH₂, C-15), 31.98 (CH₂, C-7), 32.08 (CH₂, C-12), 32.89 (CH₂, C-6), 33.78 (CH₂, C-16), 33.93 (CH₂, C-2), 35.64 (CH₂, C-1), 35.74 (CH, C-8), 38.56 (C, C-10), 47.13 (C, C-13), 49.60 (CH, C-14), 53.37 (CH, C-9), 70.32 (CH, C-20), 84.97 (C, C-17), 123.79 (CH, C-4), 171.55 (C, C-5), 199.72 (C, C-3).

P5: 17 α , 20 β , 21-Trihydroxy-4-pregnen-3-one (20 β OH-RSS).

¹H NMR (CDCl₃, 500 MHz): δ 0.85 (s, 3xH-18), 0.97 (ddd, J =12.5, 10.7 and 4.2 Hz, H-9), 1.09 (dddd, J =13.8, 12.9, 11.7 and 4.5, H-7a), 1.19 (s, 3xH-19), 1.23 (m, H-15a), 1.44 (m, H-11a), 1.58 (m, 3H, H-8, H-12a and H-16a), 1.62 (m, H-11b), 1.69 (m, H-14), 1.72 (m, H-1a), 1.78 (m, H-16b), 1.79 (m, H-15b), 1.86 (m, H-12b), 1.87 (m, H-7b), 2.04 (ddd, J =13.5, 5.1 and 3.1 Hz, H-1b), 2.28 (ddd, H =14.7, 4.5 and 2.5 Hz, H-6a), 2.37 (dddd, J =17.0, 4.8, 3.2 and 1.0 Hz, H-2a), 2.41 (m, H-6b), 2.44 (ddd, J =17.0, 14.5 and 5.0 Hz, H-2b), 3.77 (m, H-21a), 3.83 (m, 2H, H-20 and H-21b), 5.73 (br s, H-4). ¹³C NMR (CDCl₃, 125 MHz): δ 15.12 (CH₃, C-18), 17.37 (CH₃, C-19), 20.65 (CH₂, C-11), 23.84 (CH₂, C-15), 31.99 (CH₂, C-7), 31.74 (CH₂, C-12), 32.87 (CH₂, C-6), 33.22 (CH₂, C-16), 33.92 (CH₂, C-2), 35.65 (CH₂, C-1), 35.75 (CH, C-8), 38.56 (C, C-10), 47.41 (C, C-13), 49.04 (CH, C-14), 53.39 (CH, C-9), 64.69 (CH₂, C-21), 73.15 (CH, C-20), 85.18 (C, C-17), 123.82 (CH, C-4), 171.46 (C, C-5), 199.72 (C, C-3).

3. Results and discussion

3.1. Gene deletion of the innate steroid hydroxylase CYP106A1

First, in order to assess a potential HSD activity of *B. megaterium* strain GHH1 towards 17OH-P and RSS, the endogenous gene encoding the cytochrome P450 CYP106A1 was deleted, to construct a strain lacking any interfering steroid hydroxylase activity. The remaining innate cytochromes P450 are CYP102A1 (BM-3), CYP109E1 and CYP109A2. Wild-type BM-3 has not been reported to be involved in steroid metabolism. CYP109E1 and CYP109A2 did not show any activity towards common C21-steroids, as measured in our laboratory (unpublished data). *B. megaterium* strain GHH1 is a derivative of strain MS941 and differs only in the knockout of the *upp* gene, which was used as a counter-selection marker during the gene deletion experiments (Table 1 lists all strains used in this study) and still contains intact ORFs encoding the aforementioned cytochromes P450.

Cultivation of strain GHH1 in presence of the model substrate progesterone led to the formation of two products, with a substrate conversion rate of up to 20 % after 24 h. This is in agreement with results of Lee et al. (Lee et al., 2015), who describe the formation of a mono- and a dihydroxyprogesterone by CYP106A1 from *B. megaterium* ATCC 14581 (98.8 % amino acid identity with CYP106A1 from GHH1) with progesterone as substrate, as determined by mass spectrometry. We deleted the gene for CYP106A1 by transforming *B. megaterium* with vector pUCTV2_Upp_Δ106A1 (plasmid maps of the backbone vectors and used restriction sites for cloning be found in Supplemental Fig. 1), containing flanking regions of the target gene fused by SOE-PCR, to subsequently undergo the knockout procedure, as described previously (Gerber et al., 2015). The promoter of the CYP106A1 gene was kept intact to avoid negative effects on the transcription of genes downstream of the ORF. The genomic DNA of resulting colonies was screened for the deletion by PCR. A colony containing the deletion was identified and designated as GHH2.

Fig. 1A displays an agarose gel with genomic DNA of GHH1 (lane 1), the knockout vector (lane 2) and genomic DNA of GHH2 (lane 3) as template. The amplified DNA produced with DNA of GHH2 as template is truncated by approximately 1000 bp compared with GHH1 and exhibits the same size as the band produced with the knockout vector as a template indicating the deletion of the CYP106A1 gene. The knockout of CYP106A1 was further verified by carrying out *in vivo* whole cell conversions with progesterone as substrate followed by a RP-HPLC analysis. As evidenced in Fig. 1B, no products of CYP106A1 catalysis were formed with strain GHH2.

3.2. *In vivo* formation of 20 α - and 20 β -reduced products with 17OH-P as substrate

The next step was to investigate the formation of products of potential endogenous HSD activities with the newly obtained strain GHH2. All HSD reactions carried out in this study are summarized in Fig. 2.

The incubation of GHH2 with 17OH-P (S2) as a model substrate resulted in the formation of a major (P3) and a minor (P4) product, as assessed by RP-HPLC (Fig. 3). Product P3 was purified and a subsequent NMR analysis (see Section 2.5) revealed it to be

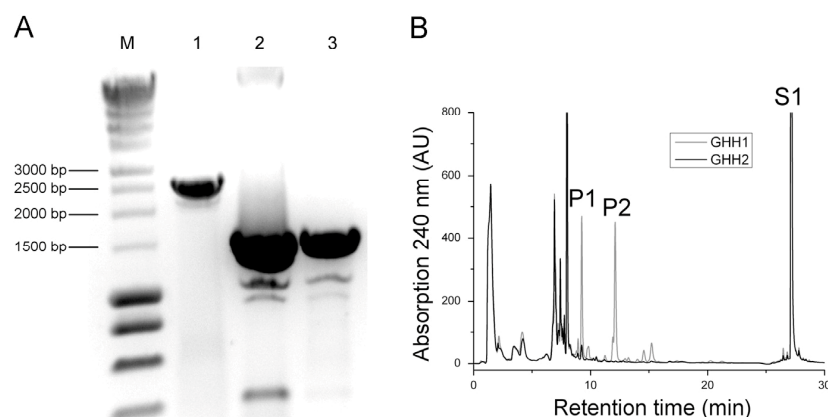


Fig. 1. Agarose gel electrophoresis and *in vivo* steroid conversion confirming CYP106A1 deletion. A: Screening for the knockout was performed by PCR, applying primers 106Afor and 106Brev which were used for the construction of the deletion construct by SOE-PCR (primer name and sequences are listed in Supplemental Table 1). A colony containing the knockout was identified and designated as strain GHH2. Agarose gel electrophoresis displaying amplified DNA with genomic DNA of the parent strain GHH1 (lane 1), knockout vector pUCTV2_Upp_Δ106A1 (lane 2) and genomic DNA of strain GHH2 (lane 3) as template (M: SmartLadder DNA marker (Eurogentec)). The major band in lane 3 is truncated by circa 1000 bp compared with the control DNA from strain GHH1 and corresponds to the band produced with the knockout vector as template (lane 2). B: RP-HPLC chromatogram showing the *in vivo* conversion of progesterone (S1, final concentration 300 μ M) with strains GHH1 (gray) and GHH2 (black) after 24 h. The two CYP106A1 products formed by strain GHH1, P1 and P2, are not present in cultures of GHH2.

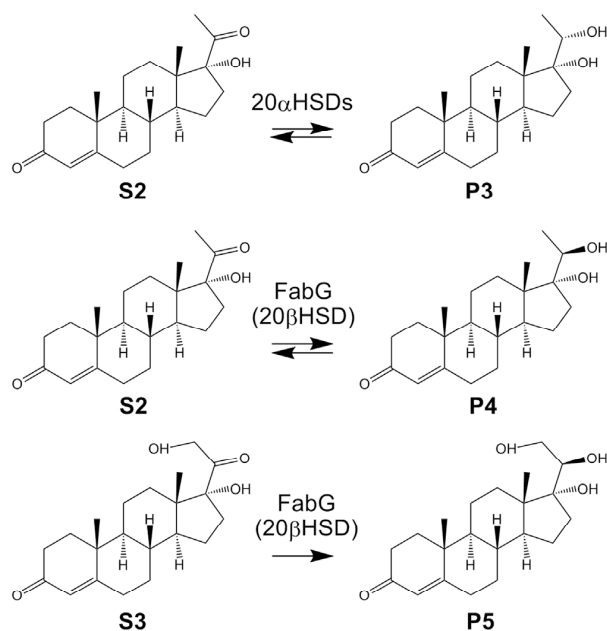


Fig. 2. Summary of the catalyzed HSD reactions described in this study. S2: 17 α -hydroxyprogesterone (17OH-P), P3: 17 α ,20 α -dihydroxy-4-pregnen-3-one (17 α ,20 α -DiOH-P), P4: 17 α ,20 β -dihydroxy-4-pregnen-3-one (17 α ,20 β -DiOH-P), S3: 11-deoxycortisol (RSS), P5: 20 β OH-RSS.

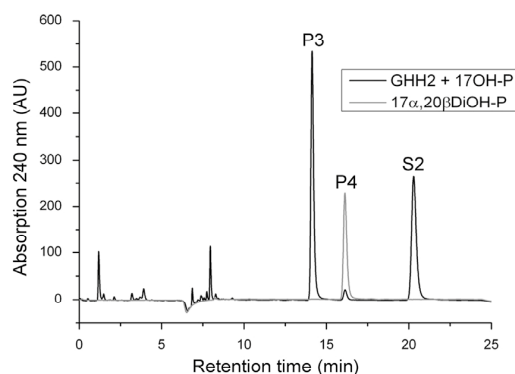


Fig. 3. Identification of 20 α - and 20 β -reduced products of 17OH-P. The HPLC-chromatogram shows the *in vivo* conversion of 17OH-P (S2, final concentration 300 μ M) with *B. megaterium* strain GHH2 (black) after 24 h. 17 α ,20 α -DiOH-P (P3) was formed as a major, 17 α ,20 β -DiOH-P (P4) as a minor product (gray: authentic 17 α ,20 β -DiOH-P standard). Additional peaks appearing during the first 10 min are also present in *B. megaterium* cultures incubated without the substrate 17OH-P (Supplemental Fig. 4).

17 α ,20 α -dihydroxy-4-pregnen-3-one (17 α ,20 α -dihydroxyprogesterone, 4-pregnene-17 α ,20 α -diol-3-one, 17 α ,20 α -DiOH-P), while the minor product P4 exhibited the same retention time as an authentic standard of the target product 17 α ,20 β -dihydroxy-4-pregnen-3-one (17 α ,20 β -dihydroxyprogesterone, 4-pregnene-17 α ,20 β -diol-3-one, 17 α ,20 β -DiOH-P) on the HPLC. The structure of P4 was further verified by NMR analysis (Section 2.5).

3.3. Identification and genomic deletion of 20 α HSDs

To remove the strain's unwanted 20 α HSD activity for the exclusive production of 20 β -reduced steroids, as a next step potential 20 α HSDs were identified from *B. megaterium*, their activities verified by overexpression experiments and their genes finally deleted.

Using MegaBac v9 (<http://megabac.tu-bs.de>), the genome of *B. megaterium* was scanned for potential HSDs applying amino acid sequences of human, bovine and rat 20 α HSDs as references. The open reading frames (ORF) of potential candidates were amplified from genomic DNA and cloned into vector pSMF2.1 under the control of a strong xylose-inducible promoter. *B. megaterium* GHH2 was transformed with the resulting plasmids. The 20 α HSD activity of the different recombinant strains towards the substrate 17OH-P was assessed by HPLC after carrying out *in vivo* whole-cell conversions. The following four ORFs were identified as putative 20 α HSD-encoding genes: BMD_0912 (annotated as oxidoreductase, aldo/keto reductase family), BMD_1068, BMD_1595 and BMD_3715 (all three annotated as a 2,5-diketo-D-gluconic acid reductase A). All four proteins belong to the NAD(P)(H) utilizing superfamily of aldo-keto reductases (AKRs) present in nearly all phyla. These mostly soluble oxidoreductases are involved in the reduction of aldehydes and ketones to primary and secondary alcohols during the phase I metabolism of xenobiotics and act on a broad range of substrates including lipids, prostaglandins, chemical carcinogens and steroids (Penning, 2015). Fig. 4A displays an exemplary HPLC chromatogram from the *in vivo* conversion of 17OH-P with strain GHH2 transformed with either pSMF2.1_0912 (expressing BMD_0912) or the empty backbone vector pSMF2.1 as a control after 20 hours. In comparison with the control strain, formation of 17 α ,20 α -DiOH-P is significantly increased while the substrate is almost completely depleted in the culture of the 20 α HSD overexpressing strain. The time course of the product formation for all 20 α HSD overexpressing strains compared with the control strain pictured in Fig. 4B demonstrates that all 20 α HSDs are able to carry out both the reduction and oxidation reaction. 17 α ,20 α -DiOH-P peak areas were normalized through multiplication by the quotients of the optical densities of the control strain, harboring the empty backbone vector pSMF2.1, and each 20 α HSD-overexpressing strain to account for negative or positive effects of the produced proteins on cell growth, which could influence 17 α ,20 α -DiOH-P production. The formation of 17 α ,20 α -DiOH-P is peaking during the first 20 h with the recombinant strains, then declines after being oxidized again, reaching a concentration comparable with the control strain. Oxidase or reductase activities of AKRs is determined by the prevailing ratio of NAD(P)H to NAD(P)⁺ in the cells (Rizner et al., 2003), which likely leads to the decrease in concentration of 17 α ,20 α -DiOH-P after 6 h in the cultures, due to NAD(P)H depletion. Each 20 α HSD was deleted stepwise in the same way as described for the knockout of CYP106A1 and the *in vivo* activity assessed after each deletion step. The final strain was designated as GHH8 (Table 1). Fig. 4C summarizes the allelic state of strain GHH8 in comparison with GHH2, as verified by PCR. All four 20 α HSDs have been truncated by approximately 800 bp. Whole-cell conversions with GHH2 and GHH8 were carried out, applying 17OH-P as substrate. After 72 h, no 17 α ,20 α -DiOH-P formation could be observed with the 20 α HSD-devoid strain (Fig. 4D). Lastly, strain GHH8 exhibited no difference in growth rate compared with the parental strain MS941, one of the most frequently applied and commercially available *B. megaterium* strains for biotechnological purposes (Fig. 4E), indicating that the combined gene deletions had no negative effects on the *B. megaterium* metabolism, at least under laboratory conditions.

3.4. Identification of *B. megaterium* 20 β HSD FabG and application for the production of 20 β -reduced steroids

After construction of strain GHH8 lacking any steroid hydroxylase and 20 α HSD activity, the responsible 20 β HSD was identified from *B. megaterium*'s genome and overexpressed to establish a whole-cell system for the production of 20 β -reduced steroids.

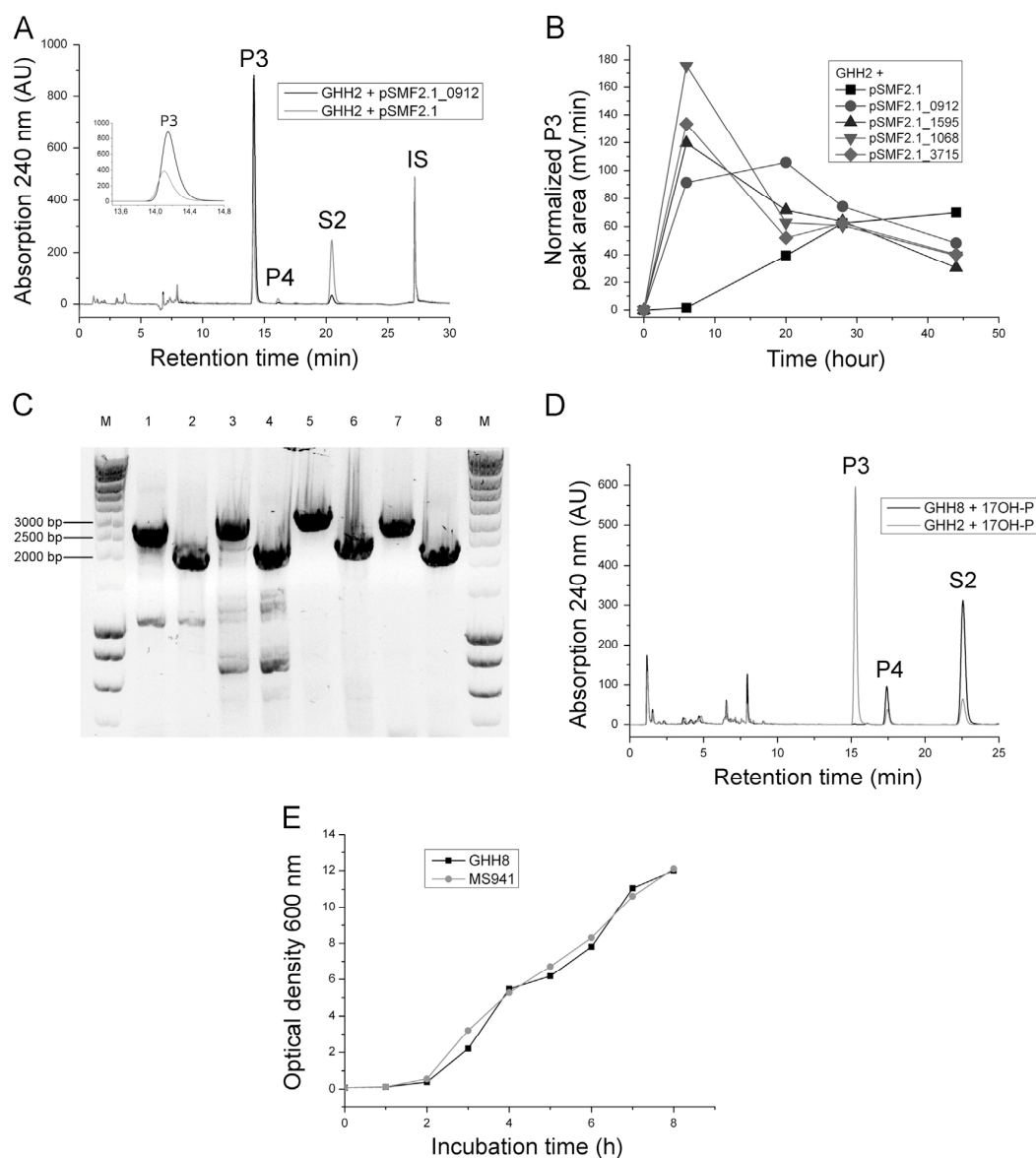


Fig. 4. Identification and genomic deletion of *B. megaterium* 20 α HSDs. A: Exemplary HPLC-chromatogram displaying the in vivo conversion of 17OH-P (S2) with strain GHH2 either transformed with pSMF2.1_0912 (black), encoding the 20 α HSD BMD_0912, or the empty vector pSMF2.1, as a control. 17 α ,20 α DiOH-P (P3) formation is significantly increased with the BMD_0912 overexpressing strain (inset: zoomed in P3 peak, IS: internal standard progesterone). B: Time course of 17 α ,20 α DiOH-P formation with the control strain and the 20 α HSD overexpressing strains. After 6 hours, the 20 α HSD overexpressing strains showed drastically increased product yields compared with the control strain. Each 20 α HSD catalyzed both the reduction and oxidation reaction, leading to a decline of 17 α ,20 α DiOH-P yield at later cultivation stages. Product areas were normalized to the optical density of the control strain, taking extraction loss into account by use of the internal standard. C: Agarose gel electrophoresis after PCR with genomic DNA of strain GHH2 and GHH8 applying primers that flank the locus of each specific 20 α HSD (0912Afor/0912Brev, 1595Afor/1595Brev, 1068Afor/1068Brev and 3715Afor/3715Brev). All four 20 α HSDs are truncated by ca. 800 bp in strain GHH8 (lanes 1, 3, 5, and 7; genomic DNA from GHH2 as template; lane 2: genomic DNA GHH5 (Δ BMD_0912); lane 4: genomic DNA GHH6 (Δ BMD_0912, Δ BMD_1068); lane 6: genomic DNA GHH7 (Δ BMD_0912, Δ BMD_1068, Δ BMD_3715) and lane 8: genomic DNA GHH8 (Δ BMD_0912, Δ BMD_1068, Δ BMD_3715, Δ BMD_1595); M: SmartLadder DNA marker). D: HPLC-chromatogram showing the conversion of 17OH-P (S2) with strain GHH2 (gray) and GHH8 (black) after 72 h. No 17 α ,20 α DiOH-P (P3) was produced with strain GHH8. E: Growth curve of strain GHH8 (black) and the parental strain MS941 (gray). In contrast to strain GHH2, MS941 still contains intact *cyp106A1* and *upp* ORFs. The combined gene deletions had no effect on growth of GHH8 under the applied cultivation conditions.

Similar genes to eukaryotic 20 β HSDs could not be found in the genome of *B. megaterium*. However, applying the amino acid sequence of a 3- α -(or 20- β)-HSD (Rv2002) from *Mycobacterium tuberculosis* as a reference (Yang et al., 2003), potential candidate genes could be obtained. These ORFs were again amplified from genomic *B. megaterium* DNA and cloned into pSMF2.1. The 20 α HSD activity-free strain GHH8 was transformed with the resulting plasmids and the in vivo activity towards 17OH-

P was measured by HPLC. FabG (BMD_4208; 3-oxoacyl-(acyl-carrier-protein) reductase) was identified as the responsible enzyme with 20 β HSD activity. FabG belongs to the family of short chain dehydrogenases involved in the type II biosynthesis of saturated and unsaturated fatty-acids, reducing the β -ketoacyl group to a β -hydroxy group during chain elongation. These enzymes exhibits a Rossmann-fold binding domain and can accept NADP(H) and/or NAD(H) as cofactor (Javidpour et al., 2014). Enzymes of that type

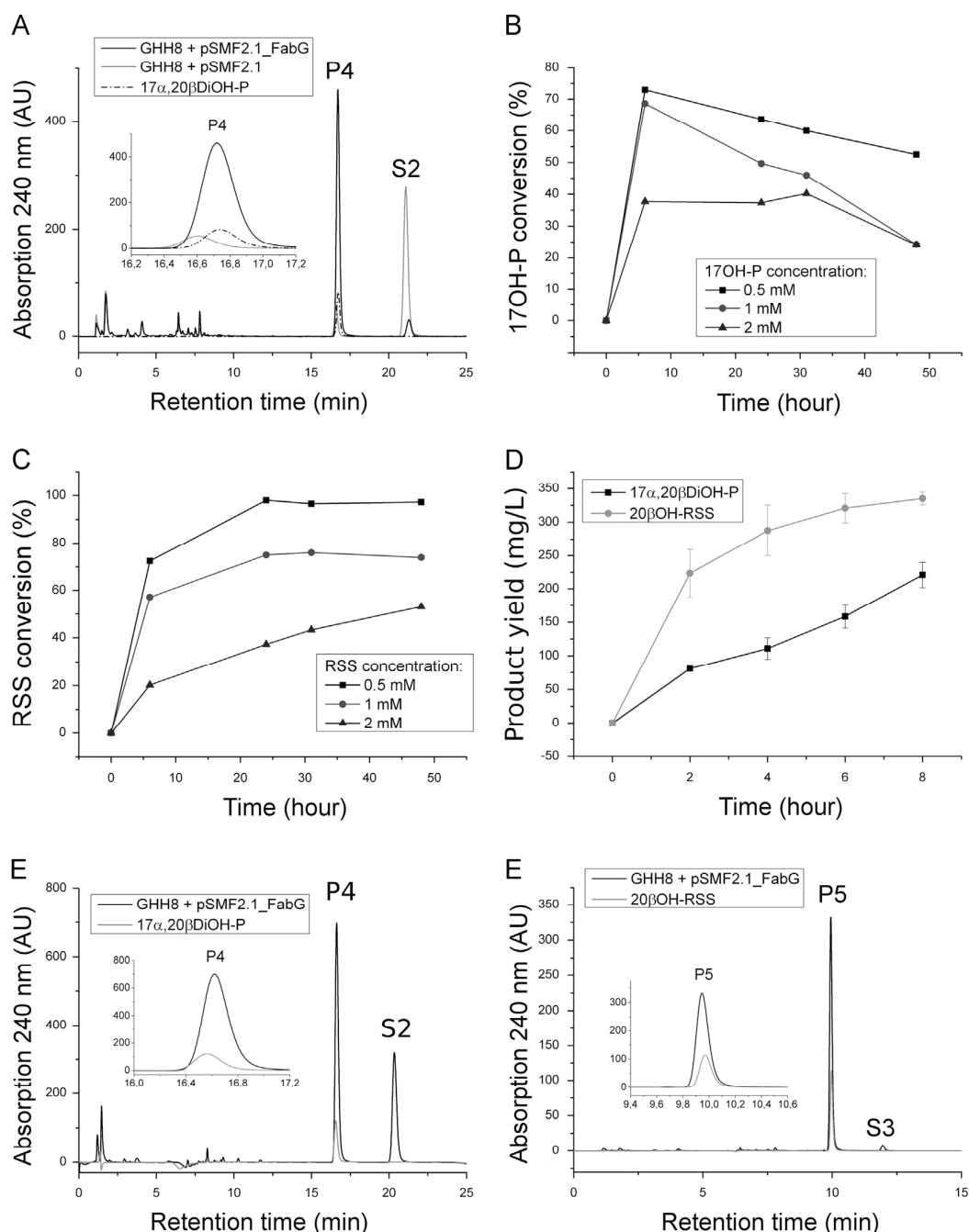


Fig. 5. Identification and overexpression of *B. megaterium* 20 β HSD FabG for 20 β -reduced steroid hormone production. A: HPLC-chromatogram showing the *in vivo* conversion of 17OH-P (S2, final concentration 300 μ M) to 17 α ,20 β DiOH-P (P4) with strain GHH8 transformed with pSMF2.1_FabG (black) or the empty pSMF2.1 vector (gray). The substrate was added simultaneously to protein induction. After 6 hours, product yield is strongly increased in the strain expressing FabG. B and C: Conversion of varying concentrations of 17OH-P and RSS after different time points (single experiments). Application of 1 mM of both substrates led to the highest absolute product yields within 10 h. D: Product yields at different time points using 17OH-P and RSS as substrates with a final concentration of 1 mM (triplicate biological experiments, error bars indicate standard deviation). Maximum yields of 17 α ,20 β DiOH-P and 20 β OH-RSS were achieved after 8 hours, with 220.9 and 335.25 mg/L, respectively. E: HPLC chromatogram showing the conversion of 17OH-P (S2, final concentration 1 mM) to 17 α ,20 β OH-P (P4) after 8 h. F: HPLC chromatogram showing the conversion of RSS (S3) (S, final concentration 1 mM) to 20 β OH-RSS (P5) after 8 h.

are also involved in bacterial PHB synthesis by reducing acetoacetyl-CoA and in the 20 β oxidation/reduction of steroids (Yang et al., 2003). After overexpression of FabG in GHH8, the 17 α ,20 β DiOH-P yield was drastically increased compared with strain GHH8 harboring the empty pSMF2.1 plasmid (Fig. 5A). In addition, RSS was found to be a substrate, being converted to

20 β OH-RSS (see Section 2.5 for NMR data). To the best of our knowledge, these steroids have not been described as substrates for any of the FabG isoform so far. Cortisol and cortisone were also identified as substrates of FabG (Supplemental Fig. 5, NMR data of the purified steroids in Supplemental Table 2), however a physiological function of the resulting products has not been

described in literature. Optimal substrate concentrations and conversion times were determined for 17OH-P and RSS (Fig. 5B and C). FabG could both reduce 17OH-P and oxidize 17 α ,20 β DiOH-P (Fig. 5B), RSS was reduced to 20 β OH-RSS with no reverse oxidation reaction of 20 β OH-RSS taking place (Fig. 5C). Whole-cell conversions with both substrates were carried out using a final concentration of 1 mM, taking samples within 10 hours (Fig. 5D). Maximum yields of 221 mg/L 17 α ,20 β DiOH-P (66.6 % relative substrate conversion) and 335 mg/L 20 β OH-RSS (96.4 %) were achieved after 8 hours. Formation rates during the first hour of incubation were as high as 40.4 mg/L/h 17 α ,20 β DiOH-P and 111.8 mg/L/h 20 β OH-RSS, as determined from Fig. 5D. Fig. 5E and F display HPLC-chromatograms for both conversions after 8 h, with the respective product standards.

4. Conclusions

17 α ,20 β DiOH-P and 20 β OH-RSS are important teleost steroid hormones and pheromones involved in oocyte maturation, spermiation, initiation of meiosis and increase of sperm motility. Their application in aquacultures has been shown to be beneficial for increasing fish spawning rates. However, both steroids are very expensive due to the lack of efficient chemical or microbial syntheses. We have established a *B. megaterium* based whole-cell system for the high-yield production of these progestogens without side-product formation. Both steroid hormones can be rapidly produced using cheap steroid precursors. Due to the high difference in price between substrates and products (compare vendor prices (Sigma-Aldrich, Steraloids) and Ouedraogo et al., 2013), carrying out the *in vivo* conversions using simple-shake flask cultivation already results in a profitable bioprocess, even when product loss from further purification steps can be expected, thus providing a good basis for large-scale fermentations.

The product of 20 α HSD activity, 17 α ,20 α -DiOH-P, occurs in higher concentrations in the human plasma than 17OH-P (Whitworth et al., 1983). It has been shown to have a hypertensive effect in sheep when injected concurrently with major corticosteroids (Butkus et al., 1982), but exhibits low affinities for classical mineralocorticoid or glucocorticoid receptors. In addition, it has been described as a competitive inhibitor of rat CYP17 lyase activity. Moreover, 17 α ,20 α DiOH-P appears to be involved in spermiation in several species of fish (Asahina et al., 1990; Tan et al., 1995). The lack of information on the specific physiological role of 17 α ,20 α DiOH-P makes it also an interesting target for a biotechnological production to study its function in more detail.

Overall, the biocatalyst is more efficient, environmentally friendly and less expensive than previously described chemical syntheses of both 20 β -reduced steroid hormones, which could open the way for the broad application of these progestogens in aquacultures.

Competing interests

The authors declare no competing interest.

Authors' contributions

A.G. designed research. A.G. performed all experiments, except for purification of 17 α ,20 α DiOH-P, carried out by P.H. M.M. provided plasmids and reagents. J.Z. performed NMR analyses. A.G. wrote the manuscript. M.M. and R.B. participated in the interpretation and discussion of experimental results and writing of the manuscript.

Acknowledgments

We would like to thank Friedhelm Meinhardt (Münster) for providing plasmid pUCTV2.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.ymben.2016.02.010>.

References

- Asahina, K., Barry, T.P., Aida, K., Fusetani, N., Hanyu, I., 1990. Biosynthesis of 17 α ,20 α -dihydroxy-4-pregnen-3-one from 17 α -hydroxyprogesterone by spermatozoa of the common carp, *Cyprinus carpio*. *J. Exp. Zool.* 255 (2), 244–249.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124 (1), 128–145.
- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98 (14), 6185–6203.
- Biedendieck, R., Borgmeier, C., Bunk, B., Stammen, S., Scherling, C., Meinhardt, F., et al., 2011. Systems biology of recombinant protein production using *Bacillus megaterium*. *Methods Enzymol.* 500, 165–195.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto-beta-boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl. Microbiol. Biotechnol.* 93 (3), 1135–1146.
- Butkus, A., Congiu, M., Scoggins, B.A., Coghlan, J.P., 1982. The affinity of 17 alpha-hydroxyprogesterone and 17 alpha, 20 alpha-dihydroxyprogesterone for classical mineralocorticoid or glucocorticoid receptors. *Clin. Exp. Pharmacol. Physiol.* 9 (2), 157–163.
- Dong, H., Zhang, D., 2014. Current development in genetic engineering strategies of *Bacillus* species. *Microb. Cell Factories* 13 (1), 1–11.
- Donova, M.V., Egorova, O.V., 2012. Microbial steroid transformations: current state and prospects. *Appl. Microbiol. Biotechnol.* 94 (6), 1423–1447.
- Ehrhardt, M., Gerber, A., Hannemann, F., Bernhardt, R., 2016. Expression of human CYP27A1 in *B. megaterium* for the efficient hydroxylation of cholesterol, vitamin D3 and 7-dihydrocholesterol. *J. Biotechnol.* 218, 34–40.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Factories.* 14 (1), 015–0300.
- Haider, S., Rao, N.V., 1994. Induced spawning of maturing Indian catfish, *Clarias batrachus* (L.), using low doses of steroid hormones and salmon gonadotropin. *Aquacult. Res.* 25 (4), 401–408.
- Hong, W.-S., Chen, S.-X., Zhang, Q.-Y., Zheng, W.-Y., 2006. Sex organ extracts and artificial hormonal compounds as sex pheromones to attract broodfish and to induce spawning of Chinese black sleeper (*Bostrichthys sinensis* Lacépède). *Aquacult. Res.* 37 (5), 529–534.
- Hunter, A.C., Carragher, N.E., 2003. Flexibility of the endogenous progesterone lactonisation pathway in *Aspergillus tamarii* KITA: transformation of a series of cortical steroid analogues. *J. Steroid Biochem. Mol. Biol.* 87 (4–5), 301–308.
- Hunter, D.J.B., Behrendorff, J.B.Y.H., Johnston, W.A., Hayes, P.Y., Huang, W., Bonn, B., et al., 2011. Facile production of minor metabolites for drug development using a CYP3A shuffled library. *Metab. Eng.* 13 (6), 682–693.
- Javidpour, P., Pereira, J.H., Goh, E.B., McAndrew, R.P., Ma, S.M., Friedland, G.D., et al., 2014. Biochemical and structural studies of NADH-dependent FabG used to increase the bacterial production of fatty acids under anaerobic conditions. *App. Environ. Microbiol.* 80 (2), 497–505.
- King, H.R., Young, G., 2001. Milt production by non-spermiating male Atlantic salmon (*Salmo salar*) after injection of a commercial gonadotropin releasing hormone analog preparation, 17 α -hydroxyprogesterone or 17 α ,20 β -dihydroxy-4-pregnen-3-one, alone or in combination. *Aquaculture* 193 (1–2), 179–195.
- Kiss, F.M., Khatri, Y., Zapp, J., Bernhardt, R., 2015. Identification of new substrates for the CYP106A1-mediated 11-oxidation and investigation of the reaction mechanism. *FEBS Lett.* 589 (18), 2320–2326.
- Kiss, F.M., Schmitz, D., Zapp, J., Dier, T.K., Volmer, D.A., Bernhardt, R., 2015. Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* - identification of a novel 11-oxidase activity. *Appl. Microbiol. Biotechnol.* 24, 24.
- Koshcheenko, K.A., Krasnova, L.A., Garsiia-Rodriges, L.K., Survtsev, V.I., Skriabin, G. K., 1976. 20alpha- and 20beta-reduction of steroids by immobilized enzyme preparations and *Bacillus megaterium* cells. *Prikl. Biokhim. Mikrobiol.* 12 (3), 322–326.
- Kovganko, N.V., Kashkan, Z.N., Shkumatov, V.M., 2001. Simple Synthesis of 17 α ,20 β -Dihydroxypreg-4-en-3-one. *Chem. Nat. Compd.* 37 (1), 55–56.
- Lee, G.Y., Kim, D.H., Kim, D., Ahn, T., Yun, C.H., 2015. Functional characterization of steroid hydroxylase CYP106A1 derived from *Bacillus megaterium*. *Arch. Pharm. Res.* 38 (1), 98–107.

- Lee, W.K., Yang, S.W., 2002. Relationship between ovarian development and serum levels of gonadal steroid hormones, and induction of oocyte maturation and ovulation in the cultured female Korean spotted sea bass *Lateolabrax maculatus* (Jeom-nong-eo). *Aquaculture* 207 (1–2), 169–183.
- Martins Pinheiro, M.F., Guimarães de Souza, S.M., Gil Barcellos, L.J., 2003. Exposure to 17 α ,20 β -dihydroxy-4-pregnen-3-one changes seminal characteristics in Nile tilapia, *Oreochromis niloticus*. *Aquacult. Res.* 34 (12), 1047–1052.
- Miwa, T., Yoshizaki, G., Naka, H., Nakatani, M., Sakai, K., Kobayashi, M., et al., 2001. Ovarian steroid synthesis during oocyte maturation and ovulation in Japanese catfish (*Silurus asotus*). *Aquaculture* 198 (1–2), 179–191.
- Nagahama, Y., Yamashita, M., 2008. Regulation of oocyte maturation in fish. *Dev. Growth Differ.* 50 (1), 13.
- Nicolaou, S.A., Gaida, S.M., Papoutsakis, E.T., 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofuels and chemicals, to biocatalysis and bioremediation. *Metab. Eng.* 12 (4), 307–331.
- Ohta, H., Kagawa, H., Tanaka, H., Okuzawa, K., Hirose, K., 1996. Changes in fertilization and hatching rates with time after ovulation induced by 17, 20 β -dihydroxy-4-pregnen-3-one in the Japanese eel, *Anguilla japonica*. *Aquaculture* 139 (3–4), 291–301.
- Ouedraogo, Y.P., Huang, L., Torrente, M.P., Proni, G., Chadwick, E., Wehmschulte, R.J., et al., 2013. A direct stereoselective preparation of a fish pheromone and application of the zinc porphyrin tweezer chiroptical protocol in its stereochemical assignment. *Chirality* 25 (9), 575–581.
- Penning, T.M., 2015. The aldo-keto reductases (AKRs): overview. *Chem. Biol. Interact.* 234, 236–246.
- Richhardt, J., Larsen, M., Meinhardt, F., 2010. An improved transconjugation protocol for *Bacillus megaterium* facilitating a direct genetic knockout. *Appl. Microbiol. Biotechnol.* 86.
- Rizner, T.L., Lin, H.K., Peehl, D.M., Steckelbroeck, S., Bauman, D.R., Penning, T.M., 2003. Human type 3 3 α -hydroxysteroid dehydrogenase (aldo-keto reductase 1C2) and androgen metabolism in prostate cells. *Endocrinology* 144 (7), 2922–2932.
- Schmitz, D., Zapp, J., Bernhardt, R., 2014. Steroid conversion with CYP106A2 – production of pharmaceutically interesting DHEA metabolites. *Microb. Cell Factories* 13 (81), 1475–2859.
- Scott, A.P., Sumpter, J.P., Stacey, N., 2010. The role of the maturation-inducing steroid, 17,20 β -dihydroxypregn-4-en-3-one, in male fishes: a review. *J. Fish Biol.* 76 (1), 183–224.
- Shkumatov, V.M., Usova, E.V., Radyuk, V.G., Kashkan, Z.N., Kovganko, N.V., Juretzek, T., et al., 2003. Oxidation of 17 α ,20 β - and 17 α ,20 α -dihydroxypregn-4-en-3-ones, side products of progesterone biotransformation with recombinant microorganisms expressing cytochrome P-45017 α . *Russ. J. Bioorg. Chem.* 29 (6), 581–587.
- Tan, A.M.C., Lee, S.T.L., Kime, D.E., Chao, T.M., Lim, H.S., Chou, R., et al., 1995. 17 α ,20 α -dihydroxy-4-pregnen-3-one, not its 20 β isomer, is produced from 17 α -hydroxyprogesterone by spermatozoa of secondary male groupers (*Epinephelus tauvina*) derived from females implanted with 17 α -methyltestosterone. *J. Exp. Zool.* 271 (6), 462–465.
- Vary, P.S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W.D., et al., 2007. *Bacillus megaterium* – from simple soil bacterium to industrial protein production host. *Appl. Microbiol. Biotechnol.* 76 (5), 957–967.
- Whitworth, J.A., Butkus, A., Coghlan, J.P., Denton, D.A., Saines, D., Scoggins, B.A., 1983. Plasma 4-pregnene-17 alpha, 20 alpha-diol-3 one (17 alpha, 20 alpha-dihydroxyprogesterone) and 17 alpha-hydroxyprogesterone in man. *Acta Endocrinol.* 102 (2), 271–276.
- Winter, J., Cerone-McLernon, A., O'Rourke, S., Ponticorvo, L., Bokkenheuser, V.D., 1982. Formation of 20 β -dihydrosteroids by anaerobic bacteria. *J. Steroid Biochem.* 17 (6), 661–667.
- Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., et al., 2013. HMDB 3.0 – the human metabolome database in 2013. *Nucleic Acids Res.* 41, 17.
- Wittchen, K.D., Meinhardt, F., 1995. Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl. Microbiol. Biotechnol.* 42 (6), 871–877.
- Yamamoto, Y., Yatabe, T., Higuchi, K., Takeuchi, Y., Yoshizaki, G., 2015. Improvement of ovulation induction by additive injection of 17,20 β -dihydroxy-4-pregnen-3-one after human chorionic gonadotropin administration in a pelagic egg spawning marine teleost, nibe croaker *Nibea mitsukurii* (Jordan & Snyder). *Aquacult. Res.* 46 (6), 1323–1331.
- Yang, J.K., Park, M.S., Waldo, G.S., Suh, S.W., 2003. Directed evolution approach to a structural genomics project: Rv2002 from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* 100 (2), 455–460.
- Yao, K., Wang, F.-Q., Zhang, H.-C., Wei, D.-Z., 2013. Identification and engineering of cholesterol oxidases involved in the initial step of sterols catabolism in *Mycobacterium neoaurum*. *Metab. Eng.* 15, 75–87.
- Yao, K., Xu, L.-Q., Wang, F.-Q., Wei, D.-Z., 2014. Characterization and engineering of 3-ketosteroid- α -1-dehydrogenase and 3-ketosteroid-9 α -hydroxylase in *Mycobacterium neoaurum* ATCC 25795 to produce 9 α -hydroxy-4-androstene-3,17-dione through the catabolism of sterols. *Metab. Eng.* 24, 181–191.

2.2 (Gerber et al., 2016)

Supplemental information

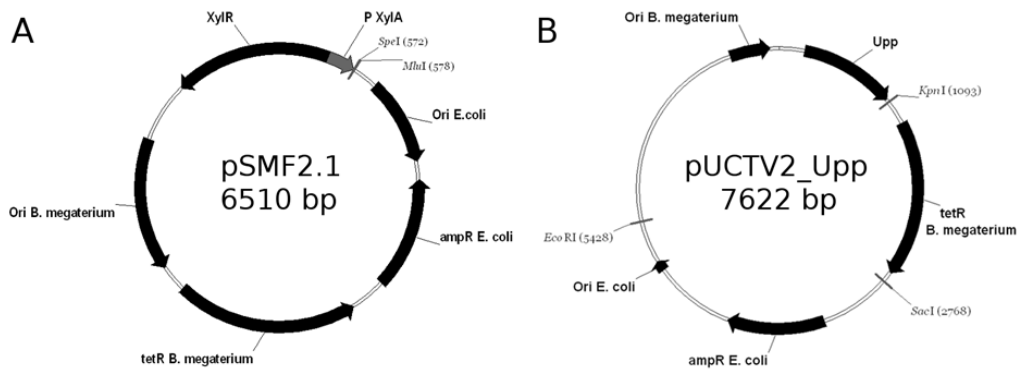
Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one.

Gerber, A., Milhim, M., **Hartz, P.**, Zapp, J. and Bernhardt, R.

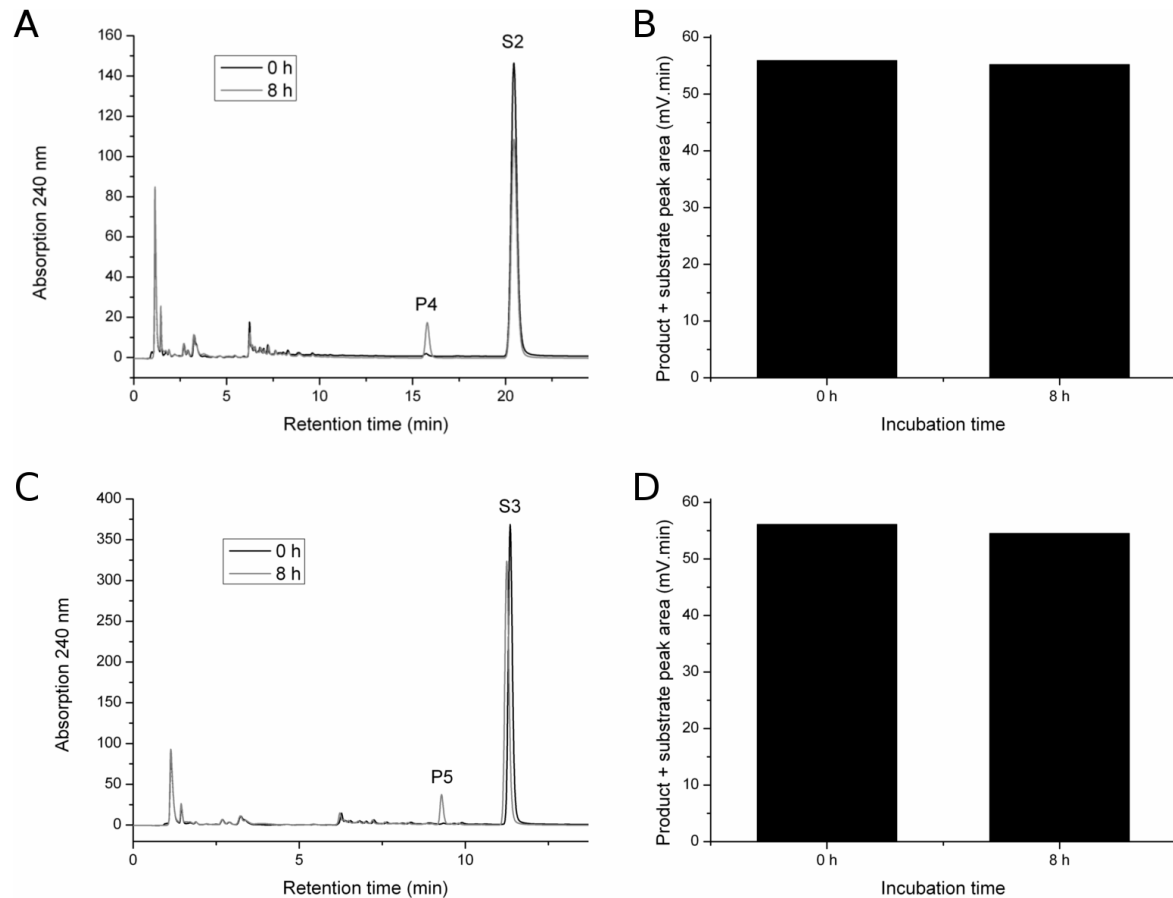
Metabolic Engineering. 2016 Feb; 36:19-27

DOI: 10.1016/j.ymben.2016.02.010

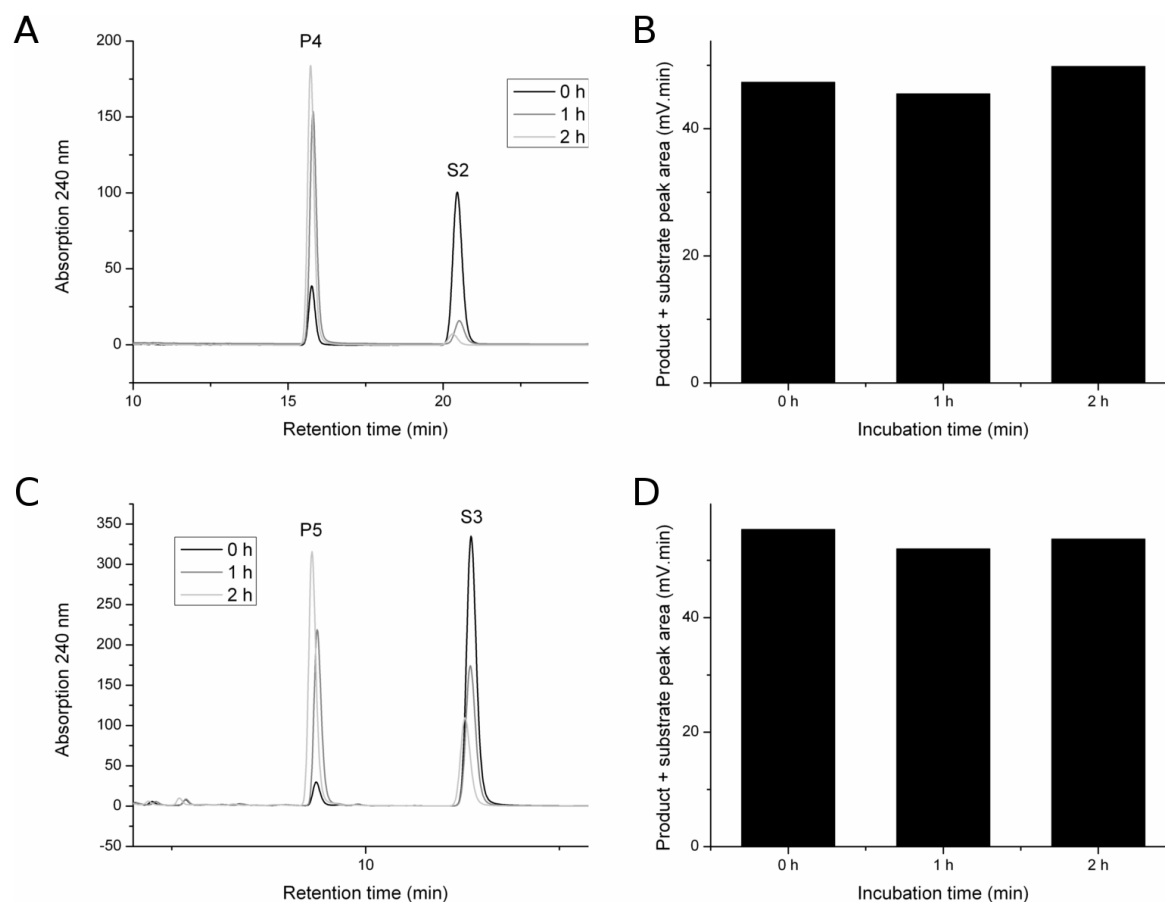
Reprinted with permission of Metabolic Engineering. All rights reserved.



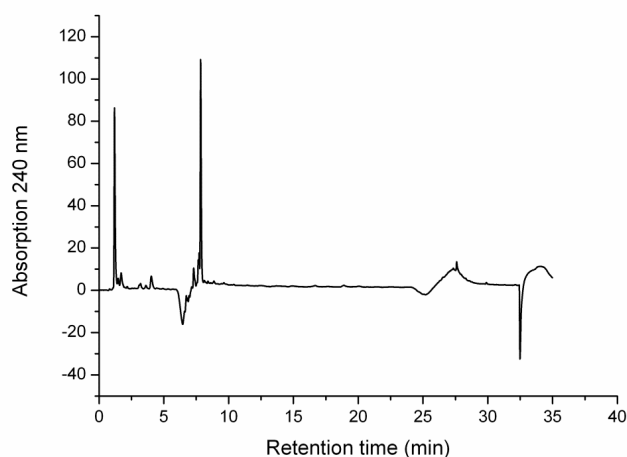
Supplemental Fig. 1. Map of backbone vectors. Plasmid pSMF2.1 was applied for overexpression experiments. SpeI and MluI restriction sites were used for cloning of BMD_0912, BMD_1595, BMD_1068, BMD_3715 and FabG resulting in plasmids pSMF2.1_0912, pSMF2.1_1595, pSMF2.1_1068, pSMF2.1_3715 and pSMF2.1_FabG, respectively (XylR: xylose repressor, PXylA: xylose isomerase promoter, Ori: origin of replication, ampR: ampicillin resistance gene, tetR: tetracycline resistance gene). Plasmid pUCTV2_Upp was applied for gene deletions. Fused flanking regions of *cyp106A1* and *BMD_3715* were cloned using the EcoRI restriction site, SacI site was used for flanking regions of *BMD_0912* and KpnI was used for the flanking regions of *BMD_1595* and *BMD_1068* resulting in plasmids pUCTV2_Upp_Δ*106A1*, pUCTV2_Upp_Δ*3715*, pUCTV2_Upp_Δ*0912*, pUCTV2_Upp_Δ*1595* and pUCTV2_Upp_Δ*1068*, respectively (Upp: uracil phosphoribosyltransferase).



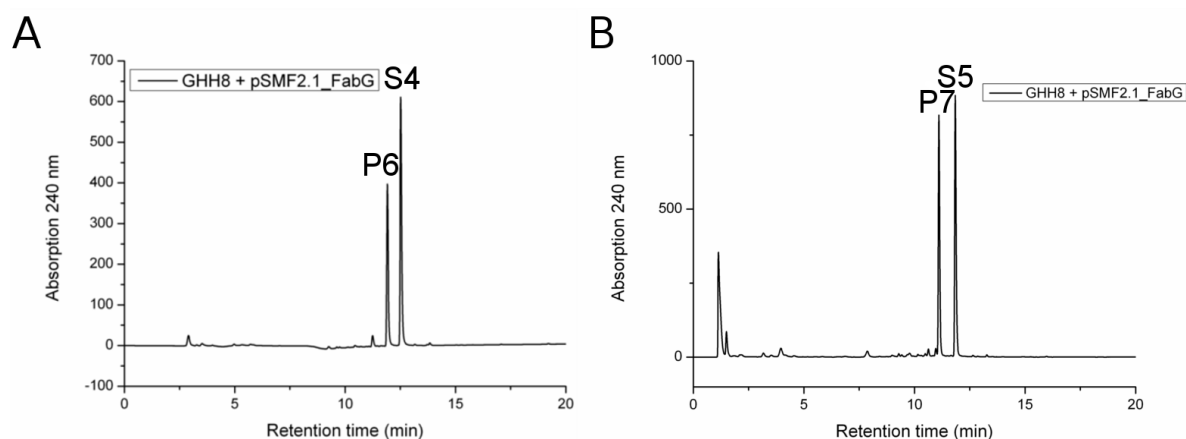
Supplemental Fig. 2. Absence of steroid degradation activity with *B. megaterium* strain GHH8. Main cultures were grown for 20 h, then substrates were added to a final concentration of 200 μ M. A and B: Addition of 17OH-P (S2). C and D: Addition of RSS (S3). No difference in the sum of substrate and product peak areas between 0 and 8 h could be observed for both substrates. 17 α ,20 α DiOH-P (P4) and 20 β OH-RSS (P5) are formed through the action of the chromosomally encoded FabG.



Supplemental Figure 3. Determination of the sum of product and substrate peak areas during 20 β HSD catalysis. A and B: Conversion of 200 μ M 17OH-P (S2) to 17 α ,20 β DiOH-P (P4). C and D: Conversion of 200 μ M RSS (S3) to 20 β OH-RSS (P5). During both reactions, the overall product and substrate peak areas remain constant, indicating that the substrates and their products exhibit similar molar extinction coefficients.



Supplemental Fig. 4. Cultivation of strain GHH8 medium with TB-medium in absence of a substrate after 24 hours. Peaks detected during the first 10 minutes are not related to conversions of the steroidal substrates.



Supplemental Fig. 5. *In-vivo* conversion of cortisol and cortisone by FabG. A: Conversion of cortisol (S, final concentration 300 μ M) to 20 β OH-cortisol (P) after 24 hours. The substrate was added simultaneously to protein induction. B: Conversion of cortisone (S, final concentration 300 μ M) to 20 β OH-cortisone (P) after 24 hours. The substrate was added simultaneously to protein induction.

Supplemental Table 1: List of primers used in this study.

Primer name	Sequence (5' -> 3')	Description
0912for	TATCA <i>ACTAGT</i> AAATCAAGGAGGTGAAT GTACA <u>ATGGAAA</u> CTTACAGTCAACTAC	Forward primer for BMD_0912 amplification (SpeI site (<i>italic</i>), followed by ribosomal binding site (RBS, bold) and start codon of ORF (<u>underlined</u>))
0912rev	TATCAACGCGTTT <u>AAAAATCAAAGTTATC</u> TGGATC	Reverse primer for BMD_0912 amplification (MluI site, stop codon of ORF (<u>underlined</u>))
1595for	TATCA <i>ACTAGT</i> AAATCAAGGAGGTGAAT GTACA <u>ATGAATATTGTTACATTA</u> AACAA	Forward primer for BMD_1595 amplification (SpeI site, RBS, start codon)
1595rev	TATCAACGCGTTT <u>ATCGGACGTT</u> CATGTC ACTTGGG	Reverse primer for BMD_1595 amplification (MluI site, stop codon)
1068for	TATCA <i>ACTAGT</i> AAATCAAGGAGGTGAAT GTACA <u>ATGAGTAATCATTTGCAAGATAC</u> AGT	Forward primer for BMD_1068 amplification (SpeI site, RBS, start codon)
1068rev	TATCAACGCGTTT <u>AAAAATCAA</u> AATTGTC CGGATC	Reverse primer for BMD_1068 amplification (MluI site, stop codon)
3715for	TATCAACGCGTTT <u>AAAAATCAAAGTTGTC</u> AGGATCT	Forward primer for BMD_3715 (MluI site, stop codon)
3715rev	TATCA <i>ACTAGT</i> AAATCAAGGAGGTGAAT GTACA <u>ATGATGAAAAATTTACAGGATAC</u> AG	Reverse primer for BMD_3715 (SpeI site, RBS, start codon)
Fabgfor	TATCAACGCGTTT <u>ACATCACCATTCCGCC</u> GTCAACG	Forward primer for FabG amplification (MluI, stop codon)
Fabgrev	TATCA <i>ACTAGT</i> AAATCAAGGAGGTGAAT GTACA <u>ATGTTACAAGGGAAAGTTGCGGT</u> TG	Reverse primer for FabG amplification (SpeI site, RBS, start codon)
106Afor	TATCAG <u>AATTCTCTGTGATCATTCCATTA</u> CTCGATTTTCT	Forward primer for amplification of flanking region downstream of CYP106A1 (EcoRI site)
106Arev	CATGTTAAACAAGTCTTGAGCGACTACG AAGGCCTTTTCTCATATCGAACCATTTGA AG	Reverse primer for amplification of flanking region downstream of CYP106A1
106Bfor	CTTCAAATGGTTCGATATGAGAAAAGGC CTTCGTAGTCGCTCAAGACTTGTTAACA TG	Forward primer for amplification of flanking region upstream of CYP106A1
106Brev	TATCAG <u>AATTCGGTTAGCAA</u> ACTATATCA CGTTTGATCTTAAGAATGA	Reverse primer for amplification of flanking region upstream of CYP106A1 (EcoRI site)
0912Afor	TATCAGAGCTCTACGTCATATATTCTCTCT TACAGG	Forward primer for amplification of flanking region downstream of BMD_0912 (SacI site)
0912Arev	CAAATACGTCTGCATTTTGGATAATATAA ACACCTAAACCTAACCAAGGC	Reverse primer for amplification of flanking region downstream of BMD_0912
0912Bfor	GCCTTGGTTAGGTTTAGGTGTTTATATTA TCCAAAATGCAGACGTATTTG	Forward primer for amplification of flanking region upstream of

0912Brev	TATCAGAGCTCGCTGTAGCTTTCTCAATT TCTTCTT	BMD_0912 Reverse primer for amplification of flanking region upstream of BMD_0912 (SacI site)
1595Afor	TATCAGGTACCGTTCGGGTATTGTTAATT CGATTGCAC	Forward primer for amplification of flanking region downstream of BMD_1595 (KpnI site)
1595Arev	CGTTCATGTCACTTGGGTGCGGATTTTTG ACACCAGACCAACTGGTAA	Reverse primer for amplification of flanking region downstream of BMD_1595
1595Bfor	TTACCAGTTGGTCTGGTGTCAAAAATCCG CACCCAAGTGACATGAACG	Forward primer for amplification of flanking region upstream of BMD_1595
1595Brev	TATCAGGTACCCTACTCCTACAGCTGTCA TTGCCTG	Reverse primer for amplification of flanking region upstream of BMD_1595 (KpnI site)
1068Afor	TATCAGGTACC CTGTCTGTAACG GCTGGAAACACA GC	Forward primer for amplification of flanking region downstream of BMD_1068 (KpnI site)
1068Arev	ATTTAATCCATCAATTTTGCTTACATCTG ATTACTCATTATAAAAAACCTCCTGCT	Reverse primer for amplification of flanking region downstream of BMD_1068
1068Bfor	AGCAGGAGGTTTTTTATAATGAGTAATCA GATGTAAGCAAATTGATGGATTAAAT	Forward primer for amplification of flanking region upstream of BMD_1068
1068Brev	TATCAGGTACCAGAGTGAGTACATTAGAC TTGCTCTTT	Reverse primer for amplification of flanking region upstream of BMD_1068 (KpnI site)
3715Afor	TATCAGAATTCTGTTGAGGCAAACATCTA ATGA	Forward primer for amplification of flanking region downstream of BMD_3715 (EcoRI site)
3715Arev	GAGTAAAGATGCCTGGCTTTGGCCGTGT AGGTCCAGATCCTGACAACCTT	Reverse primer for amplification of flanking region downstream of BMD_3715
3715Bfor	AAGTTGTCAGGATCTGGACCTACACGGC CAAAGCCAGGCATCTTTA CTC	Forward primer for amplification of flanking region upstream of BMD_3715
3715Brev	TATCAGAATTCATTTGAGAGGGTTGATTA TTTATTTTCTTA	Reverse primer for amplification of flanking region upstream of BMD_3715 (EcoRI site)

Supplemental Table 2: NMR data for 20 β OH-cortisol (P6) and 20 β OH-cortisone (P7)

Product	P6: 11 β , 17 α , 20 β , 21-Tetrahydroxy-4-pregnen-3-one (20 α OH-Cortisol)
¹H NMR (CDCl₃, 500 MHz)	¹ H NMR (CDCl ₃ , 500 MHz): δ 1.08 (s, 3xH-18), 1.02 (dd, J=11.1 and 3.4 Hz, H-9), 1.12 (m, H-7a), 1.44 (s, 3xH-19), 1.30 (m, H-15a), 1.55 (m, H-16a), 4.41 (q, 3.3 Hz, H-11), 1.65 (m, H-14), 1.87 (m, H-1a), 1.81 (m, 2H, H-15b and H-16b), 1.86 (m, H-12a), 1.98 (m, H-7b), 1.92 (m, H-1b), 2.00 (m, H-8), 2.18 (dt, J= 13.4 and 4.7, H-12a), 2.24 (ddd, 14.5, 4.5 and 2.0 Hz, H-6a), 2.35 (dt, J=16.8 and 4.3 Hz, H-2a), 2.46 (m, H-6b), 2.48 (ddd, J=16.8, 13.9 and 5.0 Hz, H-2b), 3.76 (dd, J=11.0 and 3.4 Hz, H-21a), 3.80 (dd, J=11.0 and 5.2 Hz, H-21b), 3.83 (dd, J=5.2 and 3.4 Hz, H-20), 5.67 (br s, H-4). ¹³ C NMR (CDCl ₃ , 125 MHz): δ 17.83 (CH ₃ , C-18), 20.92 (CH ₃ , C-19), 23.82 (CH ₂ , C-15), 31.12 (CH ₂ , C-6), 31.55 (CH, C-8), 32.71 (CH ₂ , C-7), 33.07 (CH ₂ , C-16), 33.81 (CH ₂ , C-2), 34.94 (CH ₂ , C-12), 39.21 (C, C-10), 41.53 (CH ₂ , C-1), 46.64 (C, C-13), 50.53 (CH, C-14), 56.00 (CH, C-9), 64.68 (CH ₂ , C-21), 68.54 (CH, C-11), 72.88 (CH, C-20), 85.15 (C, C-17), 122.26 (CH, C-4), 172.55 (C, C-5), 199.78 (C, C-3).
Product	P7: 17 α , 20 β , 21-Trihydroxy-4-pregnen-3,11-dione (20 β OH-Cortisone)
¹H NMR (CDCl₃, 500 MHz)	δ 0.77 (s, 3xH-18), 1.30 (m, H-7a), 1.32 (m, H-15a), 1.39 (s, 3xH-19), 1.63 (td, J=14.0 and 4.4 Hz, H-1a), 1.73 (m, H-16a), 1.84 (m, H-16b), 1.90 (m, H-8), 1.92 (m, H-15b), 1.93 (m, H-9), 1.96 (m, H-7b), 2.28 (m, H-6a), 2.29 (m, H-2a), 2.33 (m, H-14), 2.40 (tdd, J= 14.5, 5.2 and 1.8 Hz, H-6b), 2.47 (ddd, J=17.0, 14.0 and 5.0 Hz, H-2b), 2.49 (d, J=12.5 Hz, H-12a), 2.66 (d, J=12.5 Hz, H-12b), 2.75 (ddd, J= 14.0 5.0 and 3.2 Hz, H-1b), 3.77 (m, 3H, H-20, H-21a and H-21b), 5.72 (br s, H-4). ¹³ C NMR (CDCl ₃ , 125 MHz): δ 15.71 (CH ₃ , C-18), 17.19 (CH ₃ , C-19), 23.45 (CH ₂ , C-15), 32.31 (2xCH ₂ , C-6 and C-7), 33.33 (CH ₂ , C-16), 33.71 (CH ₂ , C-2), 34.68 (CH ₂ , C-1), 36.96 (CH, C-8), 38.15 (C, C-10), 48.35 (CH, C-14), 51.26 (CH ₂ , C-12), 51.58 (C, C-13), 62.43 (CH, C-9), 64.41 (CH ₂ , C-21), 72.57 (CH, C-20), 84.11 (C, C-17), 124.45 (CH, C-4), 169.13 (C, C-5), 200.02 (C, C-3), 210.92 (C, C-11).

2.3 (König et al., 2019)

High-yield C11-oxidation of hydrocortisone by establishment of an efficient whole-cell system in *Bacillus megaterium*.

König, L., Hartz, P., Bernhardt, R and Hannemann F.

Metabolic Engineering. 2019 Jun; 55:59-67

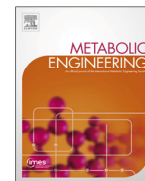
DOI: 10.1016/j.ymben.2019.06.005

Reprinted with permission of Metabolic Engineering. All rights reserved.



Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

High-yield C11-oxidation of hydrocortisone by establishment of an efficient whole-cell system in *Bacillus megaterium*

Lisa König, Philip Hartz, Rita Bernhardt, Frank Hannemann*

Department of Biochemistry, Campus B2.2, 66123, Saarland University, Saarbrücken, Germany

ARTICLE INFO

Keywords:

Bacillus megaterium
 11 β -hydroxysteroid dehydrogenase
 Biotransformation
 Recombinant protein expression
 Cell factories
 Cortisone
 Hydrocortisone
 11 β -dehydrogenation

ABSTRACT

Steroidal compounds are one of the most widely marketed pharmaceutical products. Chemical synthesis of steroidal compounds faces many challenges, including the requirement for multiple chemical steps, low yield and selectivity in several synthesis steps, low profitability and the production of environmental pollutants. Consequently, in recent decades there has been growing interest in the use of microbial systems to produce pharmaceutical compounds. Several microbial systems have recently been developed for the microbial synthesis of the glucocorticoid hydrocortisone, which serves as a key intermediate in the production of several other pharmaceutically important steroidal compounds.

In this study, we sought to establish an efficient, microbial-based system, for the conversion of hydrocortisone into cortisone. To this end, we developed a strategy for high-yield cortisone production based on ectopic expression of the guinea-pig 11 β -Hydroxysteroid dehydrogenase type 1 (11 β -HSD1) in *Bacillus megaterium*. We screened different constructs, containing a variety of promoters tailored for *B. megaterium*, and created modified versions of the enzyme by protein engineering to optimize cortisone yield. Furthermore, we utilized co-expression of an alcohol dehydrogenase to promote NADP⁺ regeneration, which significantly improved 11 β -HSD1 activity.

The process thereby developed was found to show a remarkably high regioselectivity of > 95% and to generate cortisone yields of up to 13.65 g L⁻¹ d⁻¹, which represents a ~1000-fold improvement over the next-best reported system. In summary, we demonstrate the utility of *B. megaterium* MS941 as a suitable host for recombinant protein production and its high potential for industrial steroid production.

1. Introduction

The synthesis of steroidal drugs represents one of the most important fields in the pharmaceutical industry. Since the first report of a successful total synthesis of the sex hormone equilenin in 1939, remarkable progress has been made (Bachmann and Cole, 1939). However, obtaining a satisfactory reaction performance that results in high product titers and selectivity in all stages of the reaction remains an envisaged goal (Donova and Egorova, 2012). Due to numerous complex reaction steps, the application of synthetic chemistry for the catalysis of steroid producing reactions is challenging, for example in terms of cost-efficiency and performance under demanding and harsh conditions. Particular problematic are limitations in regio- and stereoselectivity, and the use of toxic reagents, which often results in decreased profitability and ecologically unsustainable productions (Hollmann et al.,

2011). In contrast, biotechnological processes, especially the implementation of whole-cells to perform multi-step reactions for steroid modification, offer numerous advantages, including (i) cheap production of active enzymes, with increased stability provided by the protective cellular environment, (ii) the possibility for simultaneous, and therefore economic, co-factor regeneration and (iii) the ability for sustainable and selective production resulting in steroidal compounds of high purity, using organic and non-toxic media (Nikolova and Ward, 1993; Schmid et al., 2001). The shift from conventional towards biotechnological bioprocesses for manufacturing fine chemicals has thus been a focus of growing interest in the pharmaceutical industry. This is particularly true for carbon–hydrogen bond functionalization reactions, which are frequently replaced by enzymatic or microbial processes to produce commercially valuable compounds (Lednicer, 2011; Ortiz de Montellano, 2010).

Abbreviations: 11 β -HSD1, 11 β -Hydroxysteroid dehydrogenase; FabG, 3-oxoacyl-[acyl-carrier-protein] reductase; KPi, potassium phosphate buffer; LB, lysogenic broth; LbADH, *Lactobacillus brevis* alcohol dehydrogenase; RP-HPLC, reverse phase high performance liquid chromatography; TB, terrific broth; WCW, wet cell weight

* Corresponding author. Department of Biochemistry, Saarland University, Campus B 2.2, 66123, Saarbrücken, Germany.

E-mail address: f.hannemann@mx.uni-saarland.de (F. Hannemann).

<https://doi.org/10.1016/j.ymben.2019.06.005>

Received 6 February 2019; Received in revised form 31 May 2019; Accepted 14 June 2019

Available online 15 June 2019

1096-7176/ © 2019 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

One prominent example is the integration of a functional group at position C11 of steroids. Such a modification significantly affects steroid polarity, toxicity and bioreactivity (Funder, 2010). To date, only a few microbial systems have been established that enable the formation of industrially and pharmaceutically-relevant 11 β -hydroxylated steroidal compounds: all of these are based on the activity of cytochromes P450 (Bernhardt and Urlacher, 2014; Bureik and Bernhardt, 2007). For example, Suzuki et al. utilized purified P450_{11 β} of *C. lunata* to convert a range of 11-deoxy steroids into their 11 β -hydroxylated derivatives (Suzuki et al., 1993) and Schiffer et al. demonstrated the efficient 11 β -hydroxylation of 11-deoxycortisol, utilizing a modified human CYP11B1 heterologous expressed in *E. coli* (Schiffer et al., 2015). Remarkably, Sczcebara and co-workers recently established the total production of hydrocortisone in a recombinant *Saccharomyces cerevisiae* strain, using glucose or ethanol as simple raw materials, by co-expression of the relevant natural steroid biosynthetic enzymes (Sczcebara et al., 2003).

To complement the systems described above, we have established a high-yield, cell-based system, for the selective downstream-processing of C11-hydroxylated steroids to yield further compounds of high pharmaceutical interest. Specifically, oxidation of the steroidal 11 β -hydroxy group allows for the generation of related 11-keto derivatives, as exemplified in this study by the efficient production of the biotechnologically relevant metabolite cortisone from hydrocortisone. This type of reaction represents an important link to connect different biocatalytic approaches and to enhance their versatility and product range.

To date, the most common multi-step production of cortisone, known as “UpJohn synthesis” (Peterson and Murray, 1952), involves a 10-step synthesis based on progesterone as the initial compound. Due to the complex and non-profitable performance of asymmetric β -hydroxylation and thus, functionalization of progesterone at position C11 by chemical methods, this step has been conducted using biocatalysis in *Rhizopus nigricans*. This process represents one of the first applications of microbial systems in commercial high-yield steroid production. For almost 40 years, no alternative approach was available until recently, when Kiss and co-workers developed a microbial whole-cell based system for the C11-oxidation of hydrocortisone to yield cortisone. For this purpose, *Bacillus megaterium* MS941 was engineered to over-produce its endogenous cytochrome P450, CYP106A1 (Kiss et al., 2015). Nonetheless, this system suffers from low product yield, and poor selectivity in hydrocortisone oxidation, which makes it a non-competitive system compared to the previously available semi-biotechnological process. To date, no further progress has been made that would enable a commercially viable, industrial scale, biotechnological production of cortisone. For this purpose, we utilized *B. megaterium* MS941 for heterologous production of an alternative enzyme, capable to perform C11-oxidation of hydrocortisone.

The application of this gram positive expression host is reasonable due to the lack of endotoxin production. This is advantageous,

especially for applications in food and pharmaceutical industry, since these compounds harbor severe health risks for human, if they are not completely removed (Liebers et al., 2006). Importantly, *B. megaterium* shows high protein production capacity, plasmid stability and is used in industrial applications for a few decades (de Carvalho, 2017; Korneli et al., 2013). Furthermore, the complete genomes of the biotechnologically important *B. megaterium* strains MS941 and QM B1551 have already been sequenced in 2011 (Eppinger et al., 2011). Very recently, this data has been exploited for the conduction of metabolomic engineering (Biedendieck et al., 2017; Korneli et al., 2013, 2012) and for the identification of new functional promoters in this host, in order to enhance the *B. megaterium* expression system (Hartz et al., 2019).

In this study, we report on the development of an efficient whole-cell system in *B. megaterium* MS941, based on the expression of a modified guinea pig (*Cavia porcellus*) 11 β -hydroxysteroid dehydrogenase 1 (11 β -HSD1). By optimization of enzyme expression and culture conditions we were able to increase cortisone yields more than 1000-fold compared to the next best biotechnological based conversion system (Kiss et al., 2015), whilst maintaining high product purity. Our study thus paves the way for a commercially viable, biotechnological production of C11-oxidized steroidal compounds and further downstream-processing of the generated substances.

2. Materials and methods

2.1. Chemicals and enzymes

Cortisone and hydrocortisone were obtained from Sigma-Aldrich in highest purity available. Restriction enzymes were obtained from New England Biolabs (Ipswich, MA, USA) and Fast-Link™ Ligase was purchased from Lucigen Corporation. The PCRs were performed with Phusion® High-Fidelity DNA Polymerase and dNTP's, both obtained from New England Biolabs (Ipswich, MA, USA). Further chemicals and reagents were purchased from standard resources.

2.2. Bacterial strains and cultivation

Plasmid construction was conducted using *E. coli* Top10F^c (F-mcrA Δ (mrr-hsdRMS-mcrBC) f80lacZDM15 DlacX74 deoR recA1 araD139 Δ (ara-leu)7697 galU galK rpsL (StrR) endA1 nupG). For experiments including gene expression and whole-cell conversion recombinant *B. megaterium* MS941 was used. This mutant strain is a derivative of DSM319, lacking the major extracellular protease gene nprM (Wittchen and Meinhardt, 1995). Protoplasts of the strain were transformed with the described plasmids (Fig. 1) by polyethylene glycol-mediated method (Chang and Cohen, 1979). As control, *B. megaterium* MS941 was used containing an empty pSMF2.1 vector (Bleif et al., 2012) for verification that the bioconversion of hydrocortisone is catalyzed solely by the recombinant 11 β -HSD1 enzyme. The strain MS941 containing

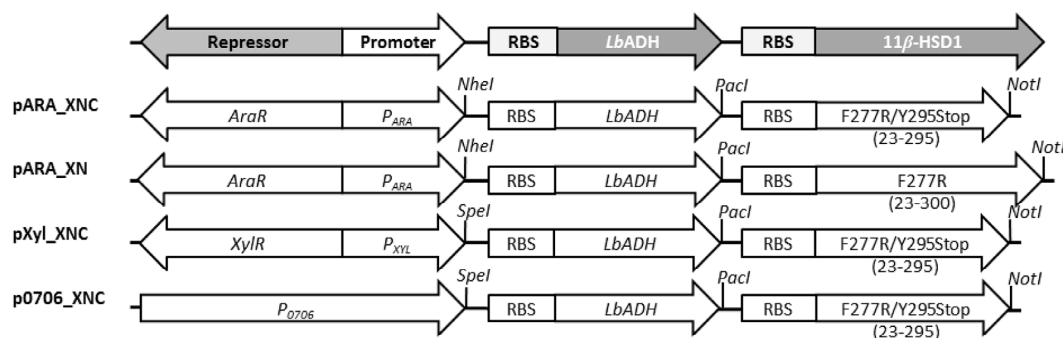


Fig. 1. Vector maps containing genes for the alcohol dehydrogenase of *Lactobacillus brevis* (LbADH) and the modified codon-optimized guinea pig 11 β -HSD1 gene sequences (11 β -HSD1) adapted to *B. megaterium*.

the empty vector did not catalyze the desired reaction (data not shown), demonstrating that the overexpressed enzymes are responsible for catalytic activity.

2.3. Plasmid construction

The sequence of the guinea pig 11 β -HSD1 gene (PDB code: 1XSE) was codon-optimized to the *B. megaterium* genome utilizing the online tool Jcat (<http://www.jcat.de>) and synthesized by GeneArt AG (Regensburg, Germany). The vector pBTacLbADH was used as template for cloning of the *Lactobacillus brevis* alcohol dehydrogenase (*LbADH*) gene to enable NADP⁺ regeneration during conversion experiments. This vector was kindly provided by Dr. Lütz (Institut für Biotechnologie 2, Forschungszentrum Jülich). The genes were amplified via PCR and all of the resulting products contain their own RBS and *NheI* or *SpeI*/*PacI* and *PacI*/*NotI* restriction sites. The codon-optimized guinea pig 11 β -HSD1 sequences contain the mutation F255R and are N-terminally truncated (24-300; XN; F255R). Furthermore, an N- and C-terminal truncated version of the insert (24-273; XNC; F255R/274Stop) was generated by PCR. The genes were cloned into the basic shuttle vector pSMF2.1 constructing different derivatives comprising the promoters P₀₇₀₆, P_{ARA} and P_{XYL}. Relevant sections of the constructed vectors are depicted schematically in Fig. 1. All utilized primers are assigned in the Supplementary Material Table S1.

2.4. Co-expression in shake flasks

Seed cultures were inoculated from a –80 °C glycerol stock using 50 mL LB medium supplemented with 10 μ g/mL tetracycline and incubated overnight at 37 °C and 150 rpm. The main cultures were supplemented with the overnight culture of recombinant *B. megaterium* MS941 cells (1:100) and were grown at 37 °C and 150 rpm. The synthesis of 11 β -HSD1 and *LbADH* in *B. megaterium* was carried out in 300 mL baffled Erlenmeyer flasks containing either 50 mL TB medium or 50 mL modified M9CA medium (Hartz et al., 2018) supplemented with 10 μ g/mL tetracycline. When an OD₅₇₈ of 0.4 was reached, expression was initiated by the addition of L-arabinose (0.4–2% w/v) or D-xylose (0.5% w/v) dissolved in distilled water. Cultures were further incubated at 30 °C or 37 °C and 150 rpm for 24 h.

2.5. In vivo steroid conversions in test tubes

Following the expression period, cultures were harvested by centrifugation (450g, 15 min, RT), washed and then suspended in 200 mM potassium phosphate buffer (pH 7.4). Alternatively the harvested cells were suspended in the already used expression medium. In both cases a defined cell density was adjusted for steroid conversion. The reaction was performed in 2 mL final volume, containing 10–25% DMSO (or 5% ethanol), 2.5% acetone and hydrocortisone dissolved in DMSO or ethanol. The reactions were carried out at 30 or 37 °C, 200 rpm in test tubes for a defined time.

2.6. Reversed phase HPLC analysis (RP-HPLC)

For product quantification via RP-HPLC, the samples were extracted twice with seven volumes of chloroform. The organic solvent was evaporated and the remaining steroids were suspended in 50% acetonitrile and separated on a Jasco reversed phase HPLC system of the LV2000 series using a reversed-phase ec MN NucleoDur C₁₈ (4.0 \times 125 mm) or reversed-phase ec MN NucleoDur C₁₈ Isis (4.6 \times 125mm) columns (Macherey-Nagel, Betlehem, PA, USA). The column was kept at an oven temperature of 40 °C. For measurements of the samples an acetonitrile/water gradient was applied (phase A: 10% acetonitrile; phase B: 100% acetonitrile; 0 min 20% B, 5 min 20% B, 13 min 40% B, 15 min 80% B, 18 min 80% B, 19 min 20% B, 22 min 20% B) with a flow rate of 0.8 mL/min and absorbance of the substances were detected at

240 nm wavelength. Differences in retention times are due to the use of different columns, validation of the substances was conducted by comparison with respective standards. The isolation of conversion products was performed using an ec MN NucleoDur C₁₈ VP (5 μ m, 8 \times 250mm) (Macherey-Nagel, Betlehem, PA, USA) column. The products were separated by preparative HPLC, according to their retention time. The maximum injectable amount of sample could reach 250 μ L and the flow rate 4 mL/min. The collected fractions were evaporated to dryness and analyzed by mass spectrometry (MS) using a QTRAP[®] 5500 LC-MS/MS System via Electrospray ionization (ESI).

2.7. Determination of the relative NADP⁺ content

The NADP⁺ concentrations were determined using an NADP/NADPH Assay Kit (Merck KGaA, Darmstadt, Germany). For fluorometric NADP⁺ detection, biological triplicates of the conversion cultures were harvested after 6 and 24 h, washed with PBS (137 mM NaCl, 10 mM KH₂PO₄/Na₂HPO₄, 2.7 mM KCl, pH 7.4) and adjusted to an OD₆₀₀ of 50 using deionized water. 10 μ L of this suspension was used for co-factor extraction. Following steps were performed according to the manufacturer's user manual and measured at $\lambda_{\text{ex}} = 530$ nm/ $\lambda_{\text{em}} = 585$ nm using the CLARIOstar[®] Plus Multi-mode Microplate Reader (BMG LABTECH GmbH, Ortenberg, Germany). The fluorometric signals were averaged and standardized to the mean values of the control samples.

3. Results

3.1. Establishment of the expression system

A modified guinea pig (*Cavia porcellus*) 11 β -hydroxysteroid dehydrogenase 1 (11 β -HSD1) has previously been shown to have high catalytic activity and selectivity in reduction of cortisone at position C11 in *E. coli* (Zhang et al., 2014). We thus sought to enable cortisone synthesis in *B. megaterium* by expressing an N-terminally truncated codon-optimized guinea pig 11 β -HSD1. Previously, it was shown that the removal of the N-terminal residues of 11 β -HSD1 is essential for soluble protein production, due to diminished formation of aggregates in a recombinant bacterial system (Elleby et al., 2004). Our 11 β -HSD1 sequence also contains the additional modification F277R. The substitution of phenylalanine by a charged amino acid (here arginine) serves to promote repulsion between protein surface domains and thus, to increase protein solubility, as previously described (Lawson et al., 2009).

To further support the biotransformation of hydrocortisone, we utilized an NADP⁺ regeneration system based upon the co-expression of an alcohol dehydrogenase from *Lactobacillus brevis* (*LbADH*) in the bicistronic expression vector. This enzyme accepts ketones as substrates and reduces them to the corresponding alcohols with the simultaneous conversion of NADPH to NADP⁺ (Fig. 2). Generating a higher NADP⁺/NADPH ratio is intended to support NADP⁺-dependent oxidation of hydrocortisone. The genes for both 11 β -HSD1 and *LbADH* were cloned into the expression vector pARA, which enables arabinose-inducible protein expression (Hartz et al., 2019).

3.2. Effect of temperature and acetone supplementation on hydrocortisone conversion

After establishing our system, we first sought to study the impact of different conversion temperatures. We found enhanced conversion efficiency of hydrocortisone to cortisone at 37 °C compared to 30 °C (Fig. 3A) and thus chose to perform all subsequent reactions at 37 °C.

HPLC chromatograms shown in Fig. 3A and B also revealed the production of an additional unknown compound at a retention time of 4.5 min. By using cells containing an empty vector, we found the 11 β -HSD1-independent side-product to be 20 β -dihydrocortisone, a modified

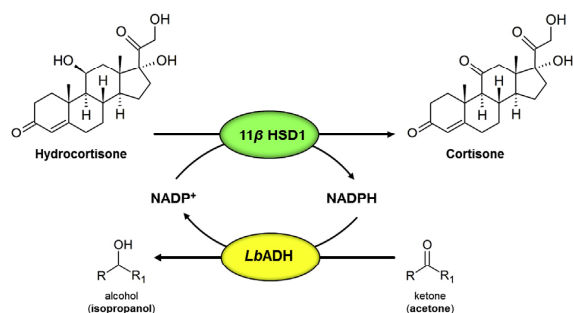


Fig. 2. Scheme of the 11 β -HSD1 and LbADH catalyzed reactions yielding cortisone and propanol. 11 β -HSD1 converts hydrocortisone into cortisone in the presence of NADP⁺. LbADH serves as a NADP⁺ regeneration system by converting acetone into propanol using NADPH as a co-factor.

product of cortisone (Fig. S1). In order to improve hydrocortisone bioconversion, we sought to activate LbADH activity by adding 2.5% (v/v) acetone to the reaction mixture before conversion, which mediates NADP⁺ regeneration both in resting and hydrocortisone-converting cells (Fig. S2). We found that acetone enhanced the hydrocortisone transformation and brought the additional advantage of limiting the formation of the undesired secondary product (Figs. 3B and S1).

3.3. C-terminal truncation of the 11 β -HSD1 improves cortisone production

Next, we sought to investigate the impact of deleting the C-terminal, hydrophobic fragment of 11 β -HSD1, which was previously shown to further increase the solubility of the enzyme without any obvious detrimental impact on enzyme activity (Zhang et al., 2014). Consequently, we truncated the four C-terminal amino acids of 11 β -HSD1 by insertion of the mutation Y295Stop to yield the construct pARA_XNC. We found that the N- and C-terminal truncated version of the protein (F277R/Y295Stop) led to a faster and more efficient hydrocortisone conversion (factor of 2.8) compared to the exclusively N-terminally truncated version (F277R) (Fig. 4). Thus, for all further experiments 11 β -HSD1 F277R/Y295Stop was used.

3.4. Optimization of the *B. megaterium* whole-cell system

3.4.1. Optimization of conversion conditions with respect to substrate solubility and cell density

Hydrocortisone displays a limited solubility in EtOH ($c_{\max} = 40$ mM). By contrast, DMSO as solvent shows a dissolving

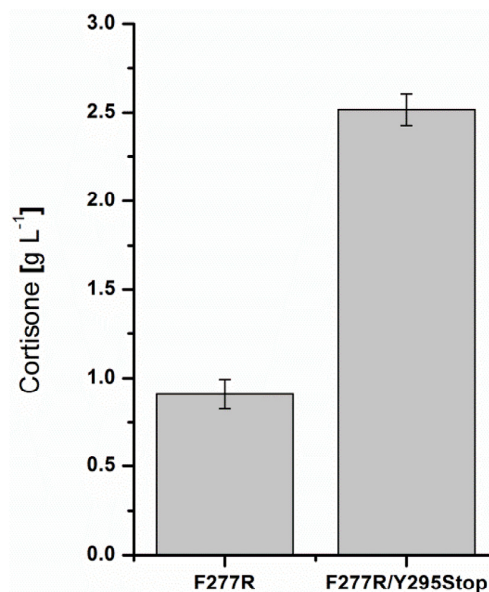


Fig. 4. Comparison of the whole-cell substrate conversion by the 11 β -HSD1 enzyme variants containing F277R substitution either with (pARA_XNC) or without C-terminal truncation (pARA_XN) at position Y295. The 24 h expression period took place in TB medium after induction with 0.4% arabinose. The reactions were carried out in 2 mL of 100 mM potassium phosphate buffer (pH 7.4) for 24 h, containing 25% DMSO, 2.5% acetone final concentration and 100 g L⁻¹ wet cell weight (WCW) in the presence of 10 mM hydrocortisone. Extracted steroids were quantified via RP-HPLC. Values represent the mean of three conversion experiments. Error bars indicate respective standard deviations.

capacity up to 400 mM. Thus, we quantified the correlation between substrate solubility and cortisone production efficiency by using varying final concentrations of DMSO in conversion cultures ranging from 0% to 35% (v/v) (Fig. 5).

We found the conversion efficiency to increase monotonically from 0.53 ± 0.005 g L⁻¹ d⁻¹ at 0% DMSO up to 2.61 ± 0.039 g L⁻¹ d⁻¹ at 25% DMSO. At higher DMSO concentrations, however, the conversion efficiency decreased. We checked for alternatives to DMSO and substituted the solvent partially with the non-ionic surfactant Tween-20 at varying concentrations, while keeping DMSO constant at 10%. Nonetheless, the addition of Tween-20 was found to inhibit cortisone production (Fig. S3). Thus, for all following experiments, the DMSO concentration was kept at 25%. We further examined the effect of cell

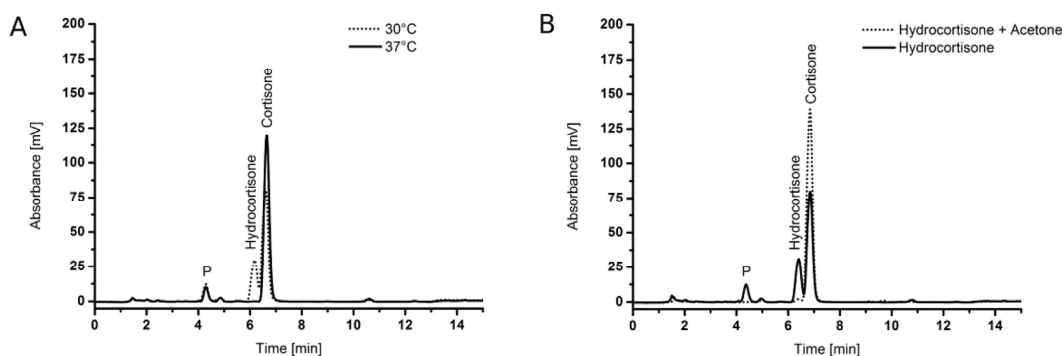


Fig. 3. HPLC chromatograms of hydrocortisone bioconversions using *B. megaterium* MS941 overexpressing LbADH and the N-terminally truncated version of the modified guinea pig 11 β -HSD1 F277R (sequences contained in pARA_XN). Resting cells in 100 mM potassium phosphate buffer (pH 7.4) were used for whole-cell biotransformation of 100 μ M hydrocortisone, supplied in 5% EtOH (v/v). P indicates a side-product formed by cortisone conversion. A: temperature dependent hydrocortisone bioconversion for 1 h at 30 °C and 37 °C, respectively; B: hydrocortisone conversion at 30 °C for 2 h, with or without supplementation of 2.5% acetone.

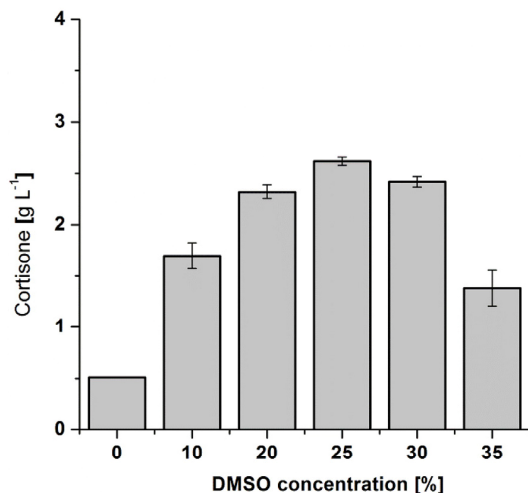


Fig. 5. Effect of DMSO concentration on 11β -HSD1 dependent cortisone formation in *B. megaterium* after 24 h. Protein production using the vector pARA_XNC was conducted in TB medium for 24 h after induction with 0.4% arabinose. The biotransformations of hydrocortisone were carried out in test tubes containing 2 mL reaction volume. The harvested cells were suspended in 100 mM potassium phosphate buffer (pH 7.4) by adjustment of 100 g L^{-1} WCW. The reaction mixtures comprise varying concentrations of DMSO, 2.5% acetone and 15 mM hydrocortisone. Steroid quantification was conducted via RP-HPLC analysis. The results represent the mean of three conversion experiments, performed simultaneously. Error bars indicate respective standard deviations.

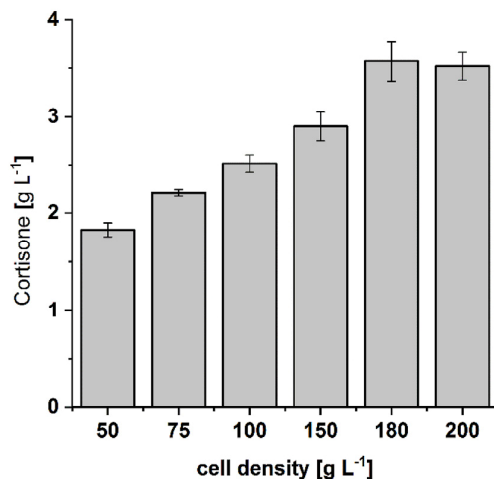


Fig. 6. Effect of increasing cell densities (WCW) on cortisone formation with *B. megaterium* MS941. Protein production using vector pARA_XNC was conducted in TB medium for 24 h after induction of protein expression with 0.4% arabinose. The whole-cell biotransformation of hydrocortisone was carried out in test tubes containing 2 mL reaction volume. The harvested cells were suspended in 100 mM potassium phosphate buffer (pH 7.4) by adjustment of varying cell densities (WCW). The reaction mixtures contain 25% DMSO, 2.5% acetone and 15 mM hydrocortisone, respectively. Steroid quantification was conducted via RP-HPLC analysis. The results represent the mean of three conversion experiments, performed simultaneously. Error bars indicate respective standard deviations.

density on cortisone productivity. To this end, we adjusted varying cell densities in the range of $50\text{--}200 \text{ g L}^{-1}$ wet cell weight (WCW) (Fig. 6). Cortisone yield was found to increase linearly from $1.83 \pm 0.07 \text{ g L}^{-1}$ at 50 g L^{-1} WCW to $3.57 \pm 0.21 \text{ g L}^{-1}$ at 180 g L^{-1} WCW. Further

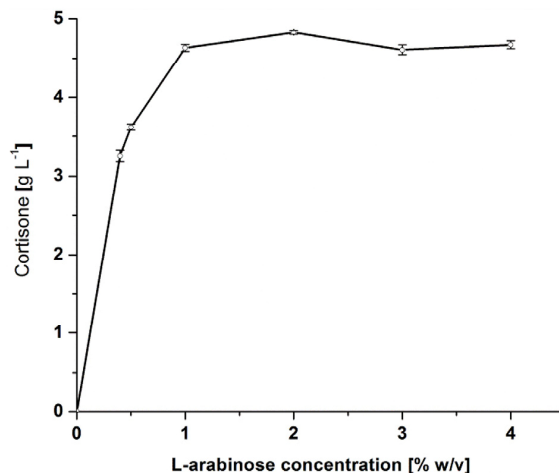


Fig. 7. Bioconversion of hydrocortisone using increasing arabinose concentrations for induction of gene expression. Protein production using the vector pARA_XNC was induced by arabinose in different concentrations and was conducted in TB medium for 24 h. The biotransformation of 15 mM hydrocortisone was performed in test tubes in 2 mL final volume. The harvested cells were suspended in 200 mM potassium phosphate buffer (pH 7.4), 25% DMSO, 2.5% acetone and cell density was adjusted to 180 g L^{-1} WCW. After a 24 h reaction period, the steroids were quantified via RP-HPLC. The results represent the mean of three conversion experiments. Error bars indicate respective standard deviations.

increase in cell density up to 200 g L^{-1} resulted in insignificant changes in cortisone production ($3.52 \pm 0.14 \text{ g L}^{-1}$). Therefore, cell density was maintained at 180 g L^{-1} WCW for all further experiments.

3.4.2. Expression conditions: inducer concentration

We further improved hydrocortisone conversion by studying efficient induction of protein expression. For this, we varied the supply of arabinose in concentrations ranging from 0 to 4% (w/v) and assessed arabinose-dependent induction efficiency by product quantification of following conversion reactions (Fig. 7). The highest yields of cortisone were observed using 2% arabinose ($4.83 \pm 0.02 \text{ g L}^{-1}$) showing an 1.5 fold increase in conversion compared with conversions using 0.4% arabinose for expression induction as applied in previous experiments. Cortisone yields remained constant at higher concentrations, consequently, 2% arabinose was used for all following experiments.

3.4.3. Influence of expression and conversion media on reaction selectivity

We next asked about the impact of expression and conversion media on the reaction selectivity and productivity. In Fig. 8, the activity of this reaction is demonstrated with regard to hydrocortisone consumption and cortisone production for determination of the respective reaction selectivity.

When complex medium (TB medium) was used, either for expression or for both expression and conversion, no significant difference in cortisone yield was observed, whilst the reaction selectivity was slightly increased from 88% to 93% when the conversion was carried out in recycled TB medium compared to conversions conducted in potassium phosphate buffer (Figs. 8 and 9A). In contrast, when M9CA minimal medium was used for expression followed by conversion in potassium phosphate buffer a marked loss in selectivity to 47% was observed (Fig. 9B). However, the use of M9CA minimal medium for both expression and conversion strongly increased reaction selectivity to 96% (Fig. 9C).

Nonetheless, in comparison with previous conditions, only 83% efficiency in hydrocortisone transformation was observed when expression and conversion were conducted both in M9CA medium. This

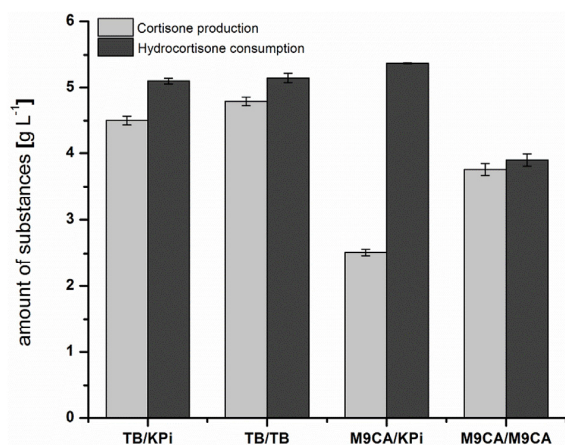


Fig. 8. Productivity and selectivity of cortisone formation dependent on expression and conversion medium. Protein production using vector pARA_XNC was carried out using TB or M9CA medium after supplementation of 2% arabinose for induction. The whole-cell biotransformation of hydrocortisone was performed in used expression medium (TB or M9CA) or 200 mM potassium phosphate buffer (pH 7.4). The reactions were carried out using 25% DMSO and 2.5% acetone final concentration and 180 g L⁻¹ WCW in the presence of 15 mM hydrocortisone. Light bars represent the obtained cortisone yield and dark bars indicate hydrocortisone consumption after 24 h. The steroids were quantified via RP-HPLC. The values represent the mean of three conversion experiments. Error bars indicate respective standard deviations.

effect might arise from a decelerated growth of the cells and thus, leads to a delay in protein production. On this account, the expression temperature was raised to 37 °C. This change would additionally facilitate an industrial fermenter process by maintaining the temperature during the expression and conversion period. Indeed, the elevation in expression temperature increased cortisone production in minimal medium equal to expression in complex media at 30 °C and hydrocortisone conversion in buffer of 4.76 g L⁻¹ d⁻¹ (Fig. 10).

3.4.4. Influence of different promoters for protein expression

Finally, we investigated the consequence of utilizing different promoter constructs (pARA_XNC, pXYL_XNC and p0706_XNC) for the comparison of time-dependent cortisone productions. For these experiments we additionally increased the hydrocortisone concentration up to 50 mM, since we previously observed elevated cortisone yields (1.75 fold) by MS941/p0706_XNC if substrate concentrations were doubled from 25 to 50 mM (Fig. S4). Additionally, we monitored enzyme activity utilizing MS941/p0706_XNC without feeding acetone. Fig. 11 shows conversion efficiencies of the different constructs as function of time. When 11 β -HSD1 was expressed from the strong

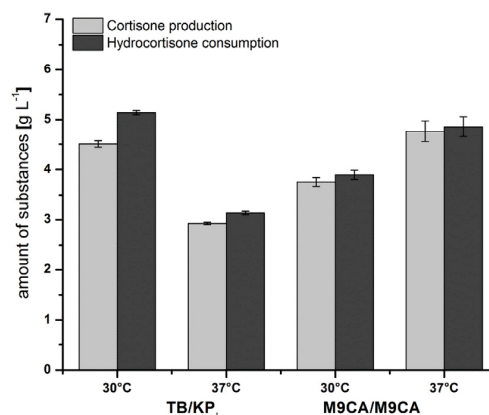


Fig. 10. Efficiency and selectivity of cortisone formation dependent on expression temperature and medium for protein expression. The expression periods of 24 h (MS941/pARA_XNC) were carried out at 30 °C or 37 °C expression temperature in TB or M9CA medium. The whole-cell biotransformation of hydrocortisone was performed at 37 °C in potassium phosphate buffer (pH 7.4) or recycled M9CA expression medium. The reactions were performed in presence of 25% DMSO, 2.5% acetone and 15 mM hydrocortisone. Light bars represent the cortisone yield and dark bars indicate the respective hydrocortisone consumption. Extracted steroids were analyzed via RP-HPLC. Values represent the mean of three conversion experiments. Error bars indicate respective standard deviations.

constitutive P₀₇₀₆ promoter, we were able to achieve cortisone yields of 13.65 ± 0.64 g L⁻¹ d⁻¹. This represents a 1.7- and even 2-fold increase in conversion efficiency compared with MS941/pARA_XNC (8.01 ± 0.32 g L⁻¹ d⁻¹) and MS941/pXYL_XNC conversions (6.66 ± 0.11 g L⁻¹ d⁻¹), respectively. Further increase of incubation duration did not lead to a significantly elevated level of cortisone when MS941/p0706_XNC is utilized, although cortisone yield continued to increase over time in reactions in which 11 β -HSD1 was expressed from the other two promoters. The addition of acetone led to a two-fold enhanced cortisone production, which highlights the supporting property of acetone in effective bioconversions. Additionally, the use of P₀₇₀₆ and P_{ARA} promoters (Hartz et al., 2019) led to a 2-fold and 1.3-fold higher conversion efficiency respectively compared with P_{XYL} under the same conditions.

4. Discussion

In this study we report on the development of a whole-cell based system for the efficient oxidation of hydrocortisone position C11 to yield cortisone. Specifically, we expressed the guinea pig 11 β -hydroxysteroid dehydrogenase 1 (11 β -HSD1) in *B. megaterium* as a biocatalyst

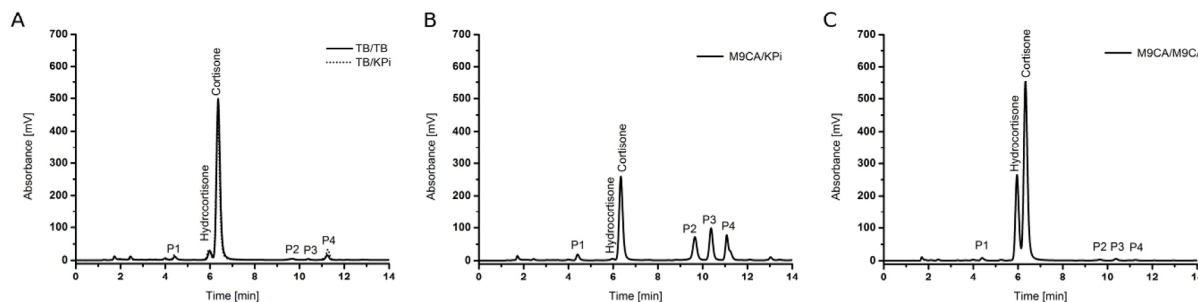


Fig. 9. HPLC chromatograms of hydrocortisone conversion in *B. megaterium* MS941 overexpressing guinea pig 11 β -HSD1 and LbADH. The whole-cell biotransformation of hydrocortisone was performed in different expression and conversion media. A. TB medium B. expression in M9CA medium and conversion in 200 mM potassium phosphate buffer, (pH 7.4) (KPi); C. expression in M9CA medium and reused for conversion. The reactions were carried out using 25% DMSO, 2.5% acetone and 180 g L⁻¹ WCW in the presence of 15 mM hydrocortisone.

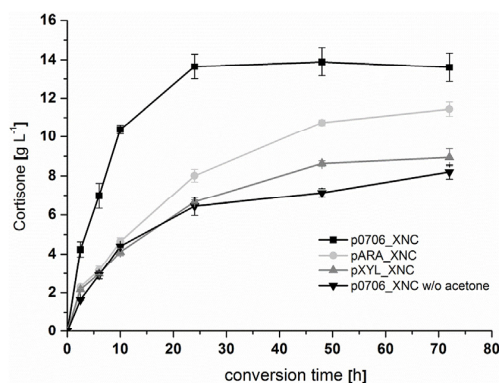


Fig. 11. Time-dependent hydrocortisone bioconversion utilizing different promoters for 11 β -HSD1-expression in *B. megaterium* MS941. The expression periods of 24 h (pARA_XNC, pXYL_XNC) and 27 h (p0706_XNC) were carried out at 37 °C in M9CA medium. The whole-cell biotransformation of hydrocortisone was performed at 37 °C in reused M9CA expression medium. Reactions containing 25% DMSO and 0 or 2.5% acetone (p0706_XNC w/o acetone and p0706_XNC, respectively) in the presence of 50 mM hydrocortisone. Extracted steroids were analyzed by RP-HPLC. Values represent the mean of three conversion experiments. Error bars indicate respective standard deviations.

to promote a highly efficient steroid conversion. Our system exhibits a 1000-fold improvement in cortisone yield compared with the CYP106A1-based starting system (Kiss et al., 2015). Due to a high reaction selectivity (> 95%) and volumetric production (13.65 g L⁻¹ d⁻¹) this system has high potential to be employed in the industrial production of steroidal compounds.

4.1. Enhancement in reaction selectivity

By the addition of acetone to the whole-cell-based reaction we inhibited a subsequent modification of cortisone (Fig. 3B) and enabled the generation of higher product concentrations (Fig. 11). The supplementation with acetone as we implemented here has three effects: (i) acetone increases membrane fluidity and permeability and thus, putatively enhances substrate uptake into the cell and facilitates substrate-protein and protein-protein interactions (Ingram, 1977; Palamanda et al., 2000); (ii) acetone inhibits the formation of the undesired side-product P (Figs. 3B and S1), most likely 20 β -dihydrocortisone. The production of this compound by the NADPH-consuming enzyme FabG was previously observed in *B. megaterium* MS941 under similar cultivation conditions, confirmed by NMR (Gerber et al., 2016; Putkaradze et al., 2017). Reduced 20 β -dihydrocortisone formation by *B. megaterium* MS941 (containing an empty vector) as a consequence of acetone addition (Fig. S1) suggests that the solvent directly affects FabG stability or activity. A similar conclusion was reached by Carrea et al. by comparison of different solvents (Carrea et al., 1988); (iii) acetone influences the cellular NADP⁺/NADPH ratio due to activation of the co-expressed *LbADH*. Acetone supplementation increased NADP⁺ generation 1.82-fold in non-converting cells and as much as 3.8-fold in hydrocortisone-converting cells, which consume NADP⁺ to a greater extent, after 24 h (Fig. S2). Consequently, on the one hand, the ratio of the relating co-factors play a major role in 11 β -HSD1 reaction activity and direction, which is reflected here in doubled cortisone yields when acetone was added to the conversion reaction (Fig. 11). On the other hand, the acetone-induced increase in the NADP⁺/NADPH ratio decreases NADPH availability for the FabG-mediated reaction (Fig. 12).

Therefore, the addition of acetone is an important step for increasing regioselectivity (up to 88%) towards cortisone and for an elevated conversion efficiency resulting in maximal yields of 4.5 g L⁻¹ d⁻¹ of cortisone in TB medium (Figs. 8 and 9A). As second step we conducted conversions in recycled M9CA medium which further

enhanced the selectivity up to 96% (Figs. 8 and 9C). By contrast, we observed a massive decline in the selectivity if cultures were previously cultivated in M9CA medium for protein production and subsequently transferred into buffer for conversion. Changing the conversion condition results in a drop of the pH from 7.0 to 6.8 within 24 h of expression in M9CA medium (data not shown) and is further reduced to pH 6 during hydrocortisone conversion in recycled M9CA medium, while the performance in buffer showed a constant pH of 8.4 (Fig. S5). Presumably, the change in this parameter causes the promotion of undefined enzymatic reactions, which lead to the formation of the observed hydrophobic side-products P2, P3 and P4 (Fig. 9B). In order to demonstrate if the side-products are formed by host-related reactions using either hydrocortisone or cortisone as substrate, we performed ESI MS analysis of the products P2, P3 and P4. This analysis revealed that the corresponding HPLC fractions contain a complex mixture of several compounds which are not all visible in the chromatogram (Fig. 9B). According to their determined molecular masses some compounds can be allocated to cortisone or hydrocortisone derivatives formed by esterification with differently sized short-chain fatty acids and corresponding dehydrated products. For instance, there is a possible esterification of hydrocortisone with propionic acid [415+H]⁺, 3 β -hydroxyhexanoic acid [475+H]⁺ and/or hexenoic acid [457+H]⁺ or esterification of cortisone with isobutyric acid or butyric acid [429+H]⁺. The latter is the dehydrated product of 3 β -hydroxybutyric acid, a compound which is predominantly present in hydrophobic granules (poly(3-hydroxybutyrate) granules; PHB granules) in *B. megaterium* serving as energy storage located in the cytosol (Jawed et al., 2016; Jendrossek, 2009). During the cultivation in nutrient-poor buffer, the mobilization of short-chain fatty acids is putatively induced for metabolic homeostasis (Handrick et al., 2000; James B.W. et al., 1999).

Furthermore, we found that expression and conversion in the same medium is possible without major changes in conversion efficiency (Figs. 8 and 9). Since a minimal medium is the growth medium of choice in fermenter systems, this experiment serves as proof-of-concept by imitating fermenter conditions. Thereby the utilization of additional buffer for implementation was avoided, resulting in a more sustainable process, and further, required overall preparation steps can be reduced. Moreover, the additional replacement of cultivation media by the M9CA medium resulted in an enhanced reaction selectivity of 96% (Figs. 8 and 9A) caused by impaired side-product formation, leading to a highly efficient cortisone formation of 8 g L⁻¹ d⁻¹ under optimal cultivation conditions using the pARA_XNC construct (Fig. 11).

4.2. Elevating the productivity of the system

To obtain enhanced catalytic activity and a further increase of the hydrocortisone conversion, we first created a mutated variant of 11 β -HSD1. Specifically, we removed the N-terminal hydrophobic domain and modified the C-terminal region of the enzyme by the deletion of four terminal residues as well as introducing a positively charged residue at position F277R. These modifications were previously shown to enhance protein solubility upon overexpression in bacterial systems due to limiting hydrophobic protein-protein or protein-membrane interactions which can result in the formation of inactive protein aggregates (Blum et al., 2000; Lawson et al., 2009; Walker et al., 2001; Zhang et al., 2014). Indeed, cortisone production was observed to be 2.8-fold higher upon truncation of the C-terminus (pARA_XNC) compared with the full-length version (pARA_XN; Figs. 4 and 13), which is consistent with previous observations (Zhang et al., 2014).

We further optimized cortisone production up to 5-fold by using DMSO at 25% (v/v) (Fig. 5). This phenomenon is probably mainly caused by enhancing hydrocortisone solubility and thus increasing substrate bioavailability (Katsu and Iguchi, 2016). Additionally, DMSO may protect the biological system from damage by free radicals (Barker et al., 1965) and increases membrane permeability, which putatively

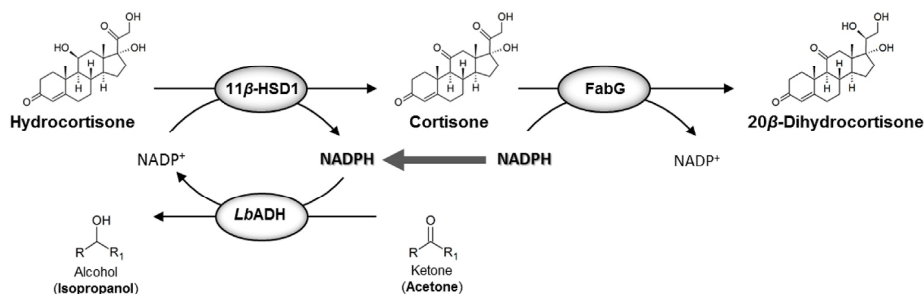


Fig. 12. Scheme of the 11β -HSD1, *LbADH* and *FabG* catalyzed reactions yielding cortisone, isopropanol and 20β -dihydrocortisone. The *FabG* related formation of the unwanted side-product 20β -dihydrocortisone occurs in dependence of NADPH.

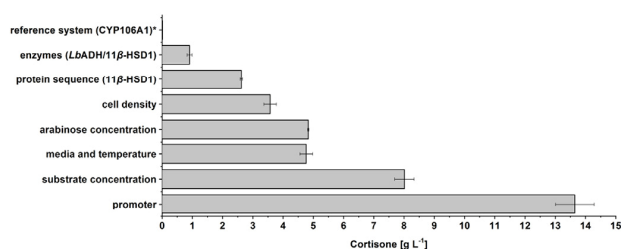


Fig. 13. Summary of the milestones in the enhancement of the presented biotechnological process. The stages of process-optimization were visualized in chronological order showing their impact on cortisone production per 24 h. The cortisone productivity using the 11β -HSD1 assumes the preliminary adjustment of the DMSO and acetone concentration. *The value of $14 \text{ mg L}^{-1} \text{ d}^{-1}$ refers to the process established by Kiss et al. in a *B. megaterium* whole-cell system (Kiss et al., 2015).

leads to an enhanced hydrocortisone uptake (Kametani and Umezawa, 1966). Although the previously described optimizations considerably increased the efficiency of cortisone formation, the major improvement in the productivity of the process was obtained by the replacement of the inducible promoters (P_{ARA} , P_{XYL}) by a constitutive promoter enabling a fully autonomous expression process, without the requirement of manual intervention for changing cultivation conditions. We realized this by substitution of P_{ARA} with P_{0706} within the pSMF2.1 backbone, which both represent two novel promoters, recently identified. These promoters have been found to be equal or even more efficient in lacZ-expression compared with the xylose-inducible system (Hartz et al., 2019). In contrast to P_{XYL} and P_{ARA} , the constitutive promoter P_{0706} displayed a 1.7- to 2.0-fold increase in cortisone productivity, elevating the yield from $6.66 \text{ g L}^{-1} \text{ d}^{-1}$ or $8.01 \text{ g L}^{-1} \text{ d}^{-1}$, respectively, up to $13.65 \text{ g L}^{-1} \text{ d}^{-1}$ under optimal cultivation conditions (Figs. 11 and 13). Furthermore, the use of a constitutive promoter provides the possibility to conduct multi-enzyme reactions, without having to consider possible catabolite repression as side-effect of the addition of sugars for induction (e.g. sucrose or arabinose) (Rygu and Hillen, 1992). These favorable properties make the system presented here ideally suited for biotechnological purposes.

4.3. Perspective

The efficient performance of C11-oxidation of hydrocortisone or other 11β -hydroxylated steroids could open a broad field of applications, yielding further substances of analytical and pharmaceutical interest. For instance, the complementation of the oxidation reaction, demonstrated in this study, by a subsequent Δ^1 -dehydrogenation could enable the generation of the synthetic derivative prednisone, which possesses elevated glucocorticoid receptor affinity and selectivity compared with naturally occurring steroids. These properties result in higher drug potency and in elimination of physiological side-effects

such as corticoid-induced salt-retention (Schacke, 2002; Zhang et al., 2013). Furthermore, exploitation of this system is conceivable for 11β -dihydrogenation in prednisolone yielding prednisone, producing an alternative drug with anti-inflammatory properties and increased value. Moreover, the isolated or additional C21-acetylation (Mosa et al., 2015) could enable the production of cortisone acetate or prednisone acetate, which are also substances used for therapeutic purposes.

5. Conclusions

Herein, we (i) established a functional and self-sufficient whole-cell system in *B. megaterium* serving as a model for the efficient C11-oxidation of C11-hydroxylated steroids. The co-expression of *LbADH*, which represents a supporting NADP^+ regeneration system, enhanced the efficiency of the biotransformation. (ii) We optimized the reaction conditions for biocatalysis by inhibition of the *B. megaterium* *FabG*-mediated modification of cortisone and by replacement of expression and conversion media from TB medium to M9CA minimal medium, which resulted in a remarkably high selectivity of the reaction (> 95%); (iii) We established a sustainable high-yield cortisone production of $13.65 \text{ g L}^{-1} \text{ d}^{-1}$ by utilizing an efficient host-derived constitutive promoter for gene expression. Its combination with the established optimized conditions resulted in a process independent of inducer supplementation or changes in fermentation conditions during cultivation, making the presented whole-cell system applicable for up-scaling and deployment in industrial fermenter processes. The presented reaction shows high potential for future studies to enable and expand its application for subsequent down-stream processing (yielding e.g. prednisone, prednisone acetate, or cortisone acetate) and for the production of further 11-keto derivatives of C19 or C21 steroids, which can be produced for pharmaceutical or analytical purposes.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LK conducted the presented experiments and drafted the manuscript. PH assisted in the conduction of experiments and interpretation of the results. FH participated in the design of the study, interpretation of the results and manuscript drafting. RB participated in the interpretation of the results and in manuscript drafting.

Funding

This work was supported by a grant of the Deutsche Bundesstiftung Umwelt (DBU).

Acknowledgements

We thank Prof. Dr. Bruce Morgan for helpful advice in writing the manuscript, Yannik Zimmermann for his assistance in conducting the fluorescence measurements and Dr. Klaus Hollemeyer (Institute of Bioanalytical Chemistry) for mass spectrometry.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jymben.2019.06.005>.

References

- Bachmann, W.E., Cole, W., 1939. The total synthesis of the sex hormone Equilenin. *J. Am. Chem. Soc.* 61, 974–975. <https://doi.org/10.1021/ja01873a513>.
- Barker, S., Crews, S., Marsters, J., Stacey, M., 1965. Inhibition of hyaluronic acid degradation by dimethyl sulphoxide. *Nature* 207 (5004), 1388–1389.
- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98, 6185–6203. <https://doi.org/10.1007/s00253-014-5767-7>.
- Biedendieck, R., Borgmeier, C., Bunk, B., 2017. Systems Biology of Recombinant Protein Production Using *Bacillus Megaterium*. <https://doi.org/10.1016/B978-0-12-385118-5.00010-4>.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- β -boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl. Microbiol. Biotechnol.* 93, 1135–1146. <https://doi.org/10.1007/s00253-011-3467-0>.
- Blum, A., Martin, H.J., Maser, E., 2000. Human 11 β -hydroxysteroid dehydrogenase 1/carbonyl reductase: recombinant expression in the yeast *Pichia pastoris* and *Escherichia coli*. *Toxicology* 144, 113–120. [https://doi.org/10.1016/S0300-483X\(99\)00197-3](https://doi.org/10.1016/S0300-483X(99)00197-3).
- Bureik, M., Bernhardt, R., 2007. Steroid hydroxylation: microbial steroid biotransformations using cytochrome P450 enzymes. *Mod. Biooxidation Enzym. React. Appl.* 155–176. <https://doi.org/10.1002/9783527611522.ch6>.
- Carrea, G., Riva, S., Bovara, R., Pasta, P., 1988. Enzymatic oxidoreduction of steroids in two-phase systems: effects of organic solvents on enzyme kinetics and evaluation of the performance of different reactors. *Enzym. Microb. Technol.* 10, 333–340. [https://doi.org/10.1016/0141-0229\(88\)90011-7](https://doi.org/10.1016/0141-0229(88)90011-7).
- Chang, S., Cohen, S.N., 1979. High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *MGG Mol. Gen. Genet.* 168, 111–115. <https://doi.org/10.1007/BF00267940>.
- de Carvalho, C.C.C.R., 2017. Whole cell biocatalysts: essential workers from Nature to the industry. *Microb. Biotechnol.* 10, 250–263. <https://doi.org/10.1111/1751-7915.12363>.
- Donova, M.V., Egorova, O.V., 2012. Microbial steroid transformations: current state and prospects. *Appl. Microbiol. Biotechnol.* 94, 1423–1447. <https://doi.org/10.1007/s00253-012-4078-0>.
- Elleby, B., Svensson, S., Wu, X., Stefansson, K., Nilsson, J., Hallén, D., Oppermann, U., Abrahmsén, L., 2004. High-level production and optimization of monodispersity of 11 β -hydroxysteroid dehydrogenase type 1. *Biochim. Biophys. Acta - Proteins Proteomics* 1700, 199–207. <https://doi.org/10.1016/j.bbapap.2004.05.003>.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin, M., Elbourne, L.D.H., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukyanchuk, V., Ball, B., Ul-Qamar, R., Seibel, J., Bremer, E., Jahn, D., Ravel, J., Vary, P.S., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J. Bacteriol.* 193, 4199–4213. <https://doi.org/10.1128/JB.00449-11>.
- Funder, J.W., 2010. Minireview: aldosterone and mineralocorticoid receptors: past, present, and future. *Endocrinology* 151, 5098–5102. <https://doi.org/10.1210/en.2010-0465>.
- Gerber, A., Milhim, M., Hartz, P., Zapp, J., Bernhardt, R., 2016. Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one. *Metab. Eng.* 36, 19–27. <https://doi.org/10.1016/j.jymben.2016.02.010>.
- Handrick, R., Reinhardt, S., Jendrossek, D., 2000. Mobilization of poly(3-hydroxybutyrate) in *Ralstonia eutropha*. *J. Bacteriol.* 182, 5916–5918. <https://doi.org/10.1128/JB.182.20.5916-5918.2000>.
- Hartz, P., Milhim, M., Trenkamp, S., Bernhardt, R., Hannemann, F., 2018. Characterization and engineering of a carotenoid biosynthesis operon from *Bacillus megaterium*. *Metab. Eng.* 49, 47–58. <https://doi.org/10.1016/j.jymben.2018.07.017>.
- Hartz, P., Mattes, C., Schad, M., Bernhardt, R., Hannemann, F., 2019. Expanding the promoter toolbox of *Bacillus megaterium*. *J. Biotechnol.* 294, 38–48. <https://doi.org/10.1016/j.jbiotec.2019.01.018>.
- Hollmann, F., Arends, I.W.C.E., Buehler, K., Schallmeyer, A., Bühler, B., 2011. Enzyme-mediated oxidations for the chemist. *Green Chem.* 13, 226–265. <https://doi.org/10.1039/C0GC00595A>.
- Ingram, L., 1977. Changes in lipid composition of *Escherichia coli* resulting from growth with organic solvents and with food additives EConcn. *Appl. Environ. Microbiol.* 33, 1233–1236.
- James, B.W., Mauchline, W.S., Dennis, P.J., Keevil, C.W., Wait, R., 1999. Poly-3-hydroxybutyrate in *Legionella pneumophila*, an energy source for survival in low-nutrient environments. *Appl. Environ. Microbiol.* 65, 822–827.
- Jawed, K., Mattam, A.J., Fatma, Z., Wajid, S., Abdin, M.Z., Yazdani, S.S., 2016. Engineered production of short chain fatty acid in *Escherichia coli* using fatty acid synthesis pathway. *PLoS One* 11, 1–20. <https://doi.org/10.1371/journal.pone.0160035>.
- Jendrossek, D., 2009. Polyhydroxyalkanoate granules are complex subcellular organelles (carbonosomes). *J. Bacteriol.* 191, 3195–3202. <https://doi.org/10.1128/JB.01723-08>.
- Kamctani, T., Umeczawa, O., 1966. NII-electronic library service. *Chem. Pharm. Bull.* 14, 369–375. <https://doi.org/10.1248/cpb.37.3229>.
- Katsu, Y., Iguchi, T., 2016. Cortisol, *Handbook of Hormones*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-801028-0.00476-1>.
- Kiss, F.M., Schmitz, D., Zapp, J., Dier, T.K.F., Volmer, D.A., Bernhardt, R., 2015. Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* – identification of a novel 11-oxidase activity. *Appl. Microbiol. Biotechnol.* 99, 8495–8514. <https://doi.org/10.1007/s00253-015-6563-8>.
- Korneli, C., Bolten, C.J., Godard, T., Franco-lara, E., 2012. Debottlenecking Recombinant Protein Production in *Bacillus Megaterium* under Large-Scale Conditions — Targeted Precursor Feeding Designed from Metabolomics, vol 109. pp. 1538–1550. <https://doi.org/10.1002/bit.24434>.
- Korneli, C., David, F., Biedendieck, R., Jahn, D., Wittmann, C., 2013. Getting the big beast to work-Systems biotechnology of *Bacillus megaterium* for novel high-value proteins. *J. Biotechnol.* 163, 87–96. <https://doi.org/10.1016/j.jbiotec.2012.06.018>.
- Lawson, A.J., Walker, E.A., White, S.A., Dafforn, T.R., Stewart, P.M., Ride, J.P., 2009. Mutations of key hydrophobic surface residues of 11 β -hydroxysteroid dehydrogenase type 1 increase solubility and monodispersity in a bacterial expression system. *Protein Sci.* 18, 1552–1563. <https://doi.org/10.1002/pro.150>.
- Lednicer, D., 2011. Steroid Chemistry at a Glance. John Wiley & Sons.
- Liebers, V., Bruning, T., Raulf-Heimsoth, M., 2006. Occupational endotoxin-exposure and possible health effects on humans. *Am. J. Ind. Med.* 49, 474–491. <https://doi.org/10.1002/ajim.20310>.
- Mosa, A., Hutter, M.C., Zapp, J., Bernhardt, R., Hannemann, F., 2015. Regioselective acetylation of C21 hydroxysteroids by the bacterial chloramphenicol acetyltransferase I. *Chembiochem* 16, 1670–1679. <https://doi.org/10.1002/cbic.201500125>.
- Nikolova, P., Ward, O.P., 1993. Whole cell biocatalysis in nonconventional media. *J. Ind. Microbiol.* 12, 76–86. <https://doi.org/10.1007/BF01569905>.
- Ortiz de Montellano, P.R., 2010. Hydrocarbon hydroxylation by cytochrome P 450 enzymes. *Chem. Rev.* 110, 932–948. (Washington, DC, United States). <https://doi.org/10.1021/cr9002193>.
- Palamanda, J., Feng, W.W., Lin, C.C., Nomeir, A.A., 2000. Stimulation of tolbutamide hydroxylation by acetone and acetonitrile in human liver microsomes and in a cytochrome P-450 2C9-reconstituted system. *Drug Metab. Dispos.* 28, 38–43.
- Peterson, D.H., Murray, H.C., 1952. Microbiological oxygenation of steroids at carbon 11. *J. Am. Chem. Soc.* 74, 1871–1872. <https://doi.org/10.1021/ja01127a531>.
- Putkaradze, N., Kiss, F.M., Schmitz, D., Zapp, J., Hutter, M.C., Bernhardt, R., 2017. Biotransformation of prednisone and dexamethasone by cytochrome P450 based systems – identification of new potential drug candidates. *J. Biotechnol.* <https://doi.org/10.1016/j.jbiotec.2016.12.011>.
- Rygu, T., Hillen, W., 1992. Catabolite repression of the xyl operon in *Bacillus megaterium*. *J. Bacteriol.* 174, 3049–3055. <https://doi.org/10.1128/jb.174.9.3049-3055.1992>.
- Schacke, H., 2002. Mechanisms involved in the side effects of glucocorticoids. *Pharmacol. Ther.* 96, 23–43. [https://doi.org/10.1016/S0163-7258\(02\)00297-8](https://doi.org/10.1016/S0163-7258(02)00297-8).
- Schiffner, L., Anderko, S., Hobler, A., Hannemann, F., Kagawa, N., Bernhardt, R., 2015. A recombinant CYP11B1 dependent *Escherichia coli* biocatalyst for selective cortisol production and optimization towards a preparative scale. *Microb. Cell Fact.* <https://doi.org/10.1186/s12934-015-0209-5>.
- Schmid, A., Dordick, J.S., Hauer, B., Kiener, A., Wubbolt, M., Witholt, B., 2001. Industrial biocatalysis and tomorrow. *Nature* 409, 258–268.
- Suzuki, K., Sanga, K., Chikaoka, Y., Itagaki, E., 1993. Purification and properties of cytochrome P-450 (P-4501u n) catalyzing steroid 11/3-hydroxylation in *Curvularia lunata*. 1203, 215–223.
- Szczębara, F.M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Dupont, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Pompon, D., Dumas, B., 2003. Total Biosynthesis of Hydrocortisone from a Simple Carbon Source in Yeast. <https://doi.org/10.1038/nbt775>.
- Walker, E.A., Clark, A.M., Hewison, M., Ride, J.P., Stewart, P.M., 2001. Functional expression, characterization, and purification of the catalytic domain of human 11 β -Hydroxysteroid dehydrogenase type 1. *J. Biol. Chem.* 276, 21343–21350. <https://doi.org/10.1074/jbc.M011142200>.
- Wittchen, K.D., Meinhardt, F., 1995. Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl. Microbiol. Biotechnol.* 42, 871–877. <https://doi.org/10.1007/BF00191184>.
- Zhang, D., Zhang, R., Zhang, J., Chen, L., Zhao, C., Dong, W., Zhao, Q., Wu, Q., Zhu, D., 2014. Engineering a hydroxysteroid dehydrogenase to improve its soluble expression for the asymmetric reduction of cortisone to 11 β -hydrocortisone. *Appl. Microbiol. Biotechnol.* 98, 8879–8886. <https://doi.org/10.1007/s00253-014-5967-1>.
- Zhang, H., Tian, Y., Wang, J., Li, Y., Wang, H., Mao, S., Liu, X., 2013. Construction of Engineered *Arthrobacter Simplex* with Improved Performance for Cortisone Acetate Biotransformation 9503–9514. <https://doi.org/10.1007/s00253-013-5172-7>.

2.3 (König et al., 2019)

Supplemental information

High-yield C11-oxidation of hydrocortisone by establishment of an efficient whole-cell system in *Bacillus megaterium*.

König, L., Hartz, P., Bernhardt, R and Hannemann F.

Metabolic Engineering. 2019 Jun; 55:59-67

DOI: 10.1016/j.ymben.2019.06.005

Reprinted with permission of Metabolic Engineering. All rights reserved.

Table S1: List of used primers for cloning in this work

Primer name	Sequence (5' → 3')	Description
11 β HSD_PacI_for	GTACTTAATTA AAAAATCAAG GAGGTGATGTACAATGAA CGAAAAATTCCGTCAG	Forward primer for 11 β -HSD1 amplification (PacI site (<i>italic</i>), followed by RBS (bold) and start codon of ORF (<u>underlined</u>))
11 β HSD_NotI_rev	GTACGCGGCCGCTTAAGCC CAACGACCG	Reverse primer for 11 β -HSD1 amplification (NotI site (<i>italic</i>), and stop codon of ORF (<u>underlined</u>))
HSD_Y274stop_NotI	CATGGCGGCCGCTTATAAT TTTTCGTTAGATAATACGT TATCCC	Reverse primer for 11 β -HSD1 amplification (C-terminally truncated version, NotI site (<i>italic</i>) and stop codon of ORF (<u>underlined</u>))
SpeI_LbADH_for	GTACACTAGTAAAATCAAG GAGGTGAATATACAATGT CTAACCGTTTGGATGGT	Forward primer for <i>LbADH</i> amplification (SpeI site (<i>italic</i>), followed by RBS (bold) and start codon of ORF (<u>underlined</u>))
LbADH_NheI_for	GTACGCTAGCAAATCAAGG AGGTGAATATACAATGTC TAACCGTTTGGATGG	Forward primer for <i>LbADH</i> amplification (NheI site (<i>italic</i>), followed by RBS (bold) and start codon of ORF (<u>underlined</u>))
LbADH_PacI_rev	CATGTTAATTA ACTATTGAG CAGTGTAGCCACC	Reverse primer for <i>LbADH</i> amplification (PacI site (<i>italic</i>), and stop codon of ORF (<u>underlined</u>))

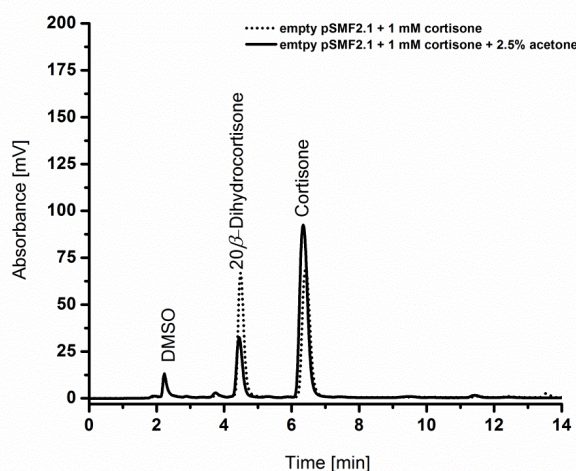


Figure S1: HPLC chromatogram of cortisone conversion in *B. megaterium* MS941 containing an empty pSMF2.1 expression vector. Following a 24 h expression period in M9CA medium at 37°C, the biotransformation of cortisone was performed in M9CA medium at the same temperature after adjustment of 180 g L⁻¹ cell density (WCW). The reaction was conducted for 24 h in 2 mL reaction volume containing 25% DMSO and 1 mM cortisone with or without supplementation of 2.5% acetone. Extracted steroids were quantified by RP-HPLC.

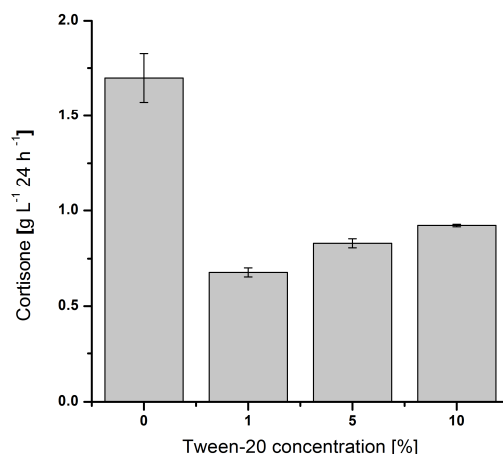


Figure S2: Effect of Tween-20 on 11 β -HSD1 dependent cortisone formation in *B. megaterium* after 24 h. Gene expression in MS941/pARA_XNC was conducted in TB medium for 24 h after induction with 0.4% arabinose. The biotransformations of hydrocortisone were carried out in test tubes containing 2 mL reaction volume. The cells were suspended in 100 mM potassium phosphate buffer (pH 7.4) by adjustment of 100 g L⁻¹ WCW. The reactions contain varying concentrations of Tween-20, 10% DMSO, 2.5% acetone and 15 mM hydrocortisone. Steroid quantification was conducted by RP-HPLC analysis. The results represent the mean of three conversion experiments, performed in simultaneously. Error bars indicate respective standard deviations.

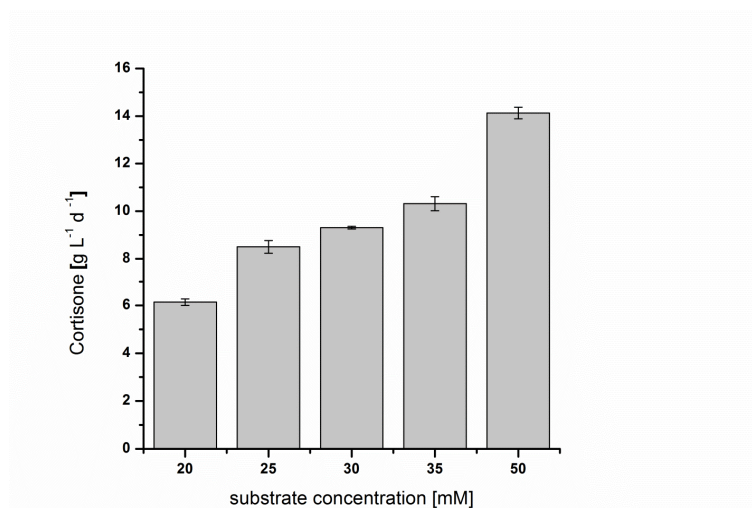


Figure S3: Effect of increasing hydrocortisone concentrations on 11 β -HSD1 dependent cortisone formation in *B. megaterium* after 24 h. Gene expression in MS941/p0706_XNC was conducted in M9CA medium for 24 h. The biotransformations of hydrocortisone were carried out in test tubes containing 2 mL reaction volume. The cells were suspended in previously used medium by adjustment of 180 g L⁻¹ WCW. The reactions contain varying concentrations of the substrate hydrocortisone using the same final concentrations of DMSO (25 %) and acetone (2.5%). Steroid quantification was conducted by RP-HPLC analysis. The results represent the mean of three conversion experiments, performed simultaneously. Error bars indicate respective standard deviations.

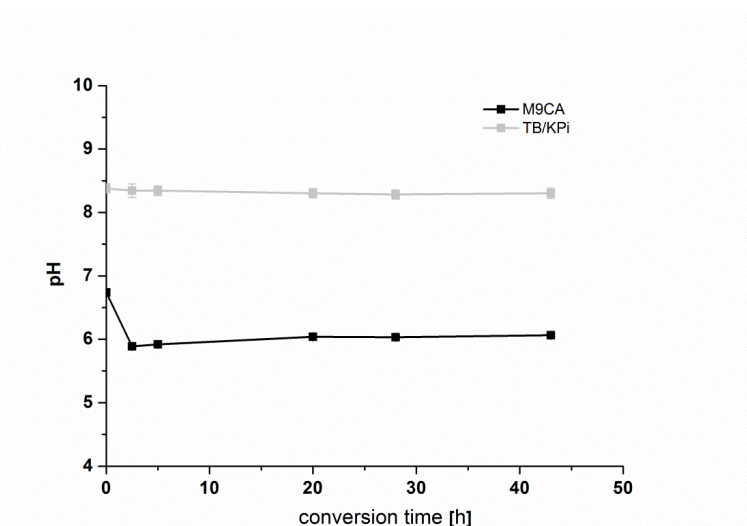


Figure S4: pH monitoring during hydrocortisone bioconversion in *B. megaterium* MS941. The whole-cell biotransformation of hydrocortisone was performed in different expression and conversion media. The grey line shows development of pH of hydrocortisone conversion reaction potassium phosphate buffer when cells were previously cultivated in TB; The black line represents values of pH monitoring in M9CA medium, recycled after expression period for conversion. The reactions were carried out using 25% DMSO, 2.5% acetone final concentration and 180 g L⁻¹ WCW in the presence of 15 mM hydrocortisone.

2.4 (Milhim et al., 2019)

A novel short chain dehydrogenase from *Bacillus megaterium* for the conversion of the sesquiterpene nootkatol to (+)-nootkatone.

Milhim, M., Hartz, P., Gerber, A. and Bernhardt R.

Journal of Biotechnology. 2019 May; 301:52-55

DOI: 10.1016/j.jbiotec.2019.05.017

Reprinted with permission of Journal of Biotechnology. All rights reserved.



Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Short communication

A novel short chain dehydrogenase from *Bacillus megaterium* for the conversion of the sesquiterpene nootkatol to (+)-nootkatone

Mohammed Milhim, Philip Hartz, Adrian Gerber, Rita Bernhardt*

Institute of Biochemistry, Saarland University, 66123, Saarbrücken, Germany



ARTICLE INFO

Keywords:

Bacillus megaterium
Short-chain dehydrogenases/reductases (SDR)
(+)-nootkatone
Nootkatol

ABSTRACT

(+)-Nootkatone is a natural ingredient that occurs in grapefruit and certain other plants and is responsible for the characteristic smell of grapefruit. Due to its versatile applications in the flavor and fragrance industry as well as its application in some medical uses it recruits the interests of academic research along with industrial biotechnology. In the current work we present the application of a novel short chain dehydrogenase from *Bacillus megaterium* in an *in vivo* whole-cell biocatalyst system for the conversion of the intermediate nootkatol into the industrially valuable (+)-nootkatone. The newly identified dehydrogenase converted nootkatol selectively and efficiently into the final product. The conversion ratio of about 100% was achieved within 40 min yielding about 44 mg/L (+)-nootkatone. Furthermore, the herein identified dehydrogenase provides a new tool to overcome the limitation of the two-step enzymatic biotechnological process for the production of (+)-nootkatone.

1. Introduction

Terpenes are a large group of naturally occurring hydrocarbons, which have valuable applications in the flavor and fragrance industry. The sesquiterpene (+)-nootkatone (C₁₅H₂₂O) has a characteristic grapefruit flavor and was first isolated from the cedar wood *Chamaecyparis nootkatensis*. It is found in trace amounts in some citrus fruits such as grapefruit, mandarin and pomelo (Leonhardt and Berger, 2014). For commercial applications (+)-nootkatone is synthesized from (+)-valencene, an abundant constituent of various citrus species (e.g. Valencia orange), which are convenient and favorable bioresources. The conversion of (+)-valencene to (+)-nootkatone comprises two steps (Scheme 1) proceeding via a regioselective hydroxylation of (+)-valencene at C-2 producing the (trans)-nootkatol, which is then oxidized to (+)-nootkatone (Fratz et al., 2009). Little is known about the enzymatic step/s involved in this biosynthesis. Therefore, many methods have been developed trying to synthesize (+)-nootkatone. A chemosynthesis method, which is based on the allylic oxidation of the precursor (+)-valencene using toxic hazardous oxidants such as chromate, manganite and *tert*-butyl hydroperoxide has been reported. More recently, biotransformations using whole-cells methods (bacteria, fungi or cell culture) expressing different enzymes such as lipoxygenase, laccase or cytochrome P450 monooxygenases have also been described (Leonhardt and Berger, 2014). While the first step can be catalyzed by a cytochrome P450 (Cankar et al., 2011;

Gavira et al., 2013), an enzyme which catalyzes the second step is still required. In the present study, we identified for the first time a novel short chain dehydrogenase (SDR) from the Gram positive bacterium *Bacillus megaterium*, which possesses activity towards (trans)-nootkatol converting it to the industrially valuable (+)-nootkatone. In view of the biotechnological application, the new dehydrogenase is very promising, as it was successfully cloned and overexpressed in *Bacillus megaterium*, and subsequently used efficiently for the oxidation of nootkatol.

2. Materials and methods

2.1. Strains, expression vectors, enzymes, and chemicals

E. coli TOP10 from Invitrogen (Karlsruhe, Germany) was used for cloning experiments. *Bacillus megaterium* GHH2 (Gerber et al., 2016) was used for *in vivo* biotransformation. All standards, and other chemicals used were purchased from Sigma-Aldrich (Schnelldorf, Germany). (Trans)-nootkatol was a gift from the Institute of Plant Molecular Biology of CNRS UPR2357, Strasbourg. All reagents and solvents were of analytical grade.

2.2. Cloning of the gene encoding BMD₂₀₉₄

Genomic DNA was prepared using a genomic DNA isolation kit (nextec). The DNA fragment encoding the full length BMD₂₀₉₄ was

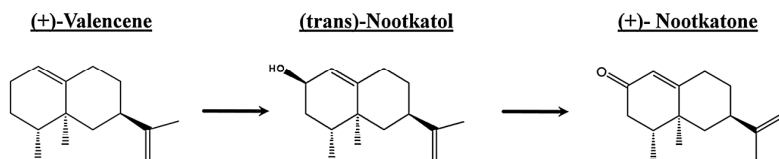
* Corresponding author at: Institute of Biochemistry, Campus B 2.2, Saarland University, 66123, Saarbrücken, Germany.
E-mail address: ritabern@mx.uni-saarland.de (R. Bernhardt).

<https://doi.org/10.1016/j.jbiotec.2019.05.017>

Received 11 December 2018; Received in revised form 9 May 2019; Accepted 27 May 2019

Available online 28 May 2019

0168-1656/ © 2019 Elsevier B.V. All rights reserved.



Scheme 1. The two-step enzymatic oxidation of (+)-valencene via (trans)-nootkatol into (+)-nootkatone.

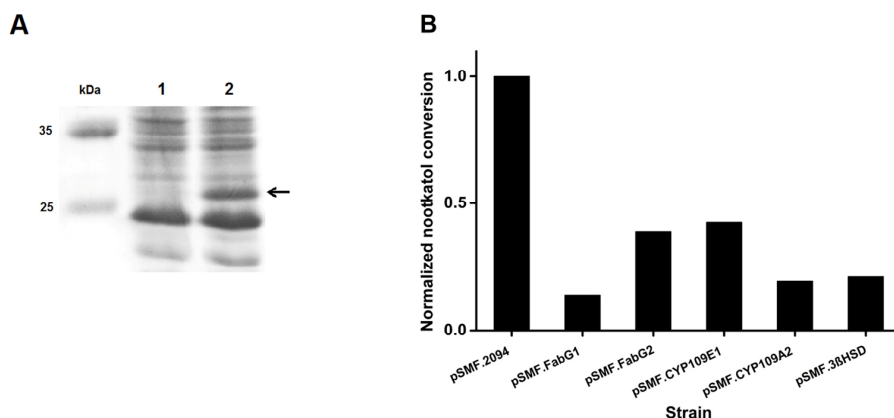


Fig. 1. (A) 15% SDS-PAGE analysis of cell lysate from *B. megaterium* transformed with the empty pSMF2.1 vector (Lane 1) or the pSMF.2094 vector (lane 2) expressing BMD_2094. The Page Ruler Prestained Protein Ladder (Thermo Fisher Scientific, USA) was used as protein standard (the size of the bands is presented in kDa). The arrow indicates BMD_2094 band. (B) *B. megaterium* whole-cell conversion of (trans)-nootkatol. Whole-cell biotransformation was performed using *B. megaterium* cells transformed with pSMF2.1 vector carrying the corresponding ORF for the genes described. Conversion was carried out by using resting cells in potassium phosphate buffer (50 mM KPi pH 7.4 + 2% glycerol) and substrate concentration of 200 μ M. Samples were taken after 5 h conversion and analyzed

via GC–MS. The data were normalized by setting BMD_2094 to 1. ORF description: FabG1 [BMD_0641: 3-oxoacyl-(acyl-carrier-protein) reductase], FabG2 [BMD_4208: 3-oxoacyl-(acyl-carrier-protein) reductase], CYP109E1 [BMD_3874: cytochrome P450], CYP109A2 [BMD_2035: cytochrome P450], 3 β HSD [Uniprot acc. No. P26439: human 3 beta-hydroxysteroid dehydrogenase].

PCR amplified from the genomic DNA of *B. megaterium* strain GHH2 using the forward primer [CGCTTAATTAaaaatcaaggagggtgaatgtaca ATGTCACAGCATTATGCGC] (the small letter represent RBS sequence) and the reverse primer [TTATCAACTAGTTTATTTAATTTCTTGTGTTT TGGGACAACGG] followed by cloning in the xylose-inducible shuttle vector pSMF2.1 (Bleif et al., 2012) with the *PacI/SpeI* restriction sites, resulting in plasmid pSMF.2094. The constructed vector was verified by sequencing.

2.3. Cultivation conditions of *B. megaterium* and in vivo substrate conversion

B. megaterium protoplast preparation as well as the subsequent PEG-mediated transformation was carried out as described elsewhere (Biedendieck et al., 2011).

Transformed *B. megaterium* cells were grown overnight in 50 ml Luria-Bertani (LB) broth medium supplemented with 10 μ g/ml tetracycline at 37 $^{\circ}$ C and shaking at 140 rpm. For the expression of proteins, 50 ml Terrific broth (TB) medium containing 10 μ g/ml tetracycline were inoculated (1:100) with the transformed cells and cultivated in baffled erlenmeyer flask at 37 $^{\circ}$ C with rotary shaking at 140 rpm. Protein expression was induced at OD600 = 0.4 – 0.6 with 4 mg/ml xylose. The temperature was then reduced to 30 $^{\circ}$ C. After incubation for 24 h, cells were harvested by centrifugation (4000 g) for 10 min at 4 $^{\circ}$ C and washed once with 1 vol of conversion buffer (50 mM potassium phosphate buffer (pH 7.4) supplemented with 2% glycerol). After a second centrifugation, the cell pellets were resuspended in conversion buffer to an end cell suspension concentration of 40 g wet cell weight (wcw) / L buffer. The substrate was added to a final concentration of 200 μ M and the culture was incubated for the indicated time at 30 $^{\circ}$ C and 140 rpm.

2.4. Metabolite extraction and GC–MS analysis

Culture sample/s (1 ml) were taken at different time points and extracted twice with equal volumes of ethylacetate followed by evaporating of the organic phase using a rotary evaporator. After that, the residues were dissolved in ethylacetate and analyzed with gas

chromatography mass spectrometry (GC–MS) system consisting of a DSQII quadrupole, a Focus GC column oven (Thermo scientific, Waltham, USA), and a DB-5 column (Agilent) with a length of 25 m, 0.32 mm ID, and 0.52 μ m film thickness. The metabolites were analyzed in a m/z range of 20–350. The starting oven temperature was 50 $^{\circ}$ C for 1 min and then the temperature was ramped to 310 $^{\circ}$ C by 10 $^{\circ}$ C/min and held for 3 min with a flow rate of 1 ml/min. The EI-mass spectra were compared with the NIST mass spectral library (version 2.0).

3. Results and discussion

The suitability and efficiency of *B. megaterium* based whole cell systems for the production of terpenoids and steroids has been shown before in our laboratories (Bleif et al., 2012; Brill et al., 2013; Gerber et al., 2015). Here, we report on the ability of a *B. megaterium* based whole cell system to convert the sesquiterpene (trans)-nootkatol into the valuable ketone form (+)-nootkatone (scheme 1). After demonstrating the ability of *B. megaterium* cells to convert (trans)-nootkatol into (+)-nootkatone (data not shown), we aimed to identify and characterize the responsible enzyme for this reaction, which is thought to be a dehydrogenation reaction. Therefore, we searched for potential enzyme candidate/s in our collection of cloned dehydrogenases and cytochromes P450 originating from *B. megaterium* strain GHH2 [unpublished data, (Gerber et al., 2016)].

The activity of the candidate enzymes towards (trans)-nootkatol was verified by overexpression approach in *B. megaterium* GHH2. To achieve this, we used the xylose inducible vector pSMF2.1 as a backbone for the overexpression of the genes. Remarkably, none of the screened strains showed high activity except for the one transformed with the pSMF.2094. The expression of BMD_2094 and its contribution to the observed activity towards (trans)-nootkatol was confirmed by SDS-PAGE electrophoresis (Fig. 1A) demonstrating a correlation between the expression of BMD_2094 and the observed efficient (trans)-nootkatol conversion (Fig. 1B) in the recombinant and induced strain. The background activity detected in the other transformed strains in Fig. 1B resulted from the endogenous (genomic) BMD_2094 or from unspecific reactions (data not shown) leading to a decrease of (trans)-nootkatol (without nootkatone formation) catalyzed by the expressed proteins.

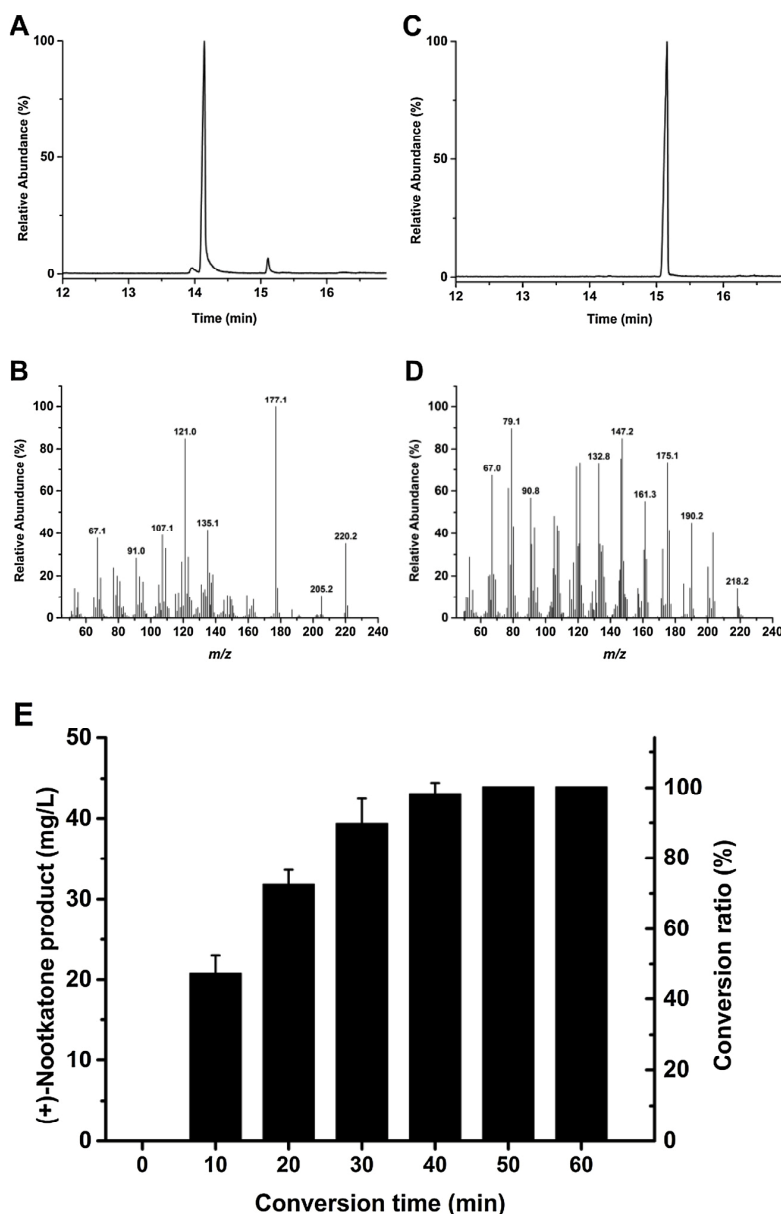


Fig. 2. *B. megaterium* whole-cell conversion of (trans)-nootkatol. (A–D) GC–MS analysis of the *in vivo* conversion of 200 μ M (trans)-nootkatol. Whole-cell biotransformation was performed using *B. megaterium* cells carrying the pSMF2.1 empty vector (A) or the pSMF.2094 vector expressing BMD_2094 (C) for 5 h. (B) and (D) display the product identification by mass spectrometry of (trans)-nootkatol and (+)-nootkatone, respectively. (E) Time dependent *in vivo* conversion of 200 μ M (trans)-nootkatol by *B. megaterium* cells transformed with pSMF.2094 vector. Conversion was carried out by using resting cells in potassium phosphate buffer (50 mM KPi pH 7.4 + 2% glycerol) and samples were taken at the indicated times and analyzed via GC–MS. The data is represented as mean \pm SD of three separate measurements.

BMD_2094 ORF [Uniprot acc. No. D5DFE7] encodes a 259 amino acid polypeptide with a calculated molecular mass of the deduced amino acid sequence of 28.85 kDa and a pI of 9.40. The BMD_2094 protein belongs to the short chain dehydrogenase (SDR) superfamily which is found in all domains of life (Kallberg et al., 2010). SDRs are a functionally diverse family of NAD(P)(H)-dependent oxidoreductases and are known for a broad substrate spectrum ranging from alcohols, sugars, steroids and aromatic compounds to xenobiotics. The N-terminal region binds the coenzymes NAD(H) or NADP(H), while the C-terminal region constitutes the substrate binding part (Kallberg et al., 2002). Previously, many publications highlighted the participation of the dehydrogenases in the biosynthetic pathway of different terpenes such as the synthesis of zerumbone from 8-hydroxy- α -humulenein the ginger plant *Zingiber zerumbet* (Okamoto et al., 2011) and the germacrene-derived sesquiterpene lactones from the vegetable *Cichorium intybus* (de Kraker et al., 2001). As shown in Fig. 2, the *B. megaterium* strain expressing the BMD_2094 (Fig. 2B) was able to convert (trans)-nootkatol

(RT = 14.14 min) completely to (+)-nootkatone (RT = 15.16 min), whereas in the control strain (Fig. 2A), just a minimal conversion was detected (~6%). Time-dependent conversion supported the high efficiency of the BMD_2094. The conversion ratio of about 100% was achieved within 40 min yielding about 44 mg/L (+)-nootkatone (Fig. 2E). Most of the enzymatic based-methods, described so far for (+)-nootkatone production still have some limitations restricting their use in industrial processes, including the substrate specificity, the need of auxiliary redox partners, the difficulties of expression of membranous enzymes in prokaryotic microorganism and the accumulation of the intermediate nootkatol (Fraatz et al., 2009). Most of these limitations have been overcome by expressing highly selective plant cytochromes P450 in different yeast species or by the engineering of bacterial cytochromes P450 to improve the formation of (trans)-nootkatol (Schulz et al., 2015). Information about the enzymatic step leading from (trans)-nootkatol to (+)-nootkatone is, however, very scarce. The first attempt for the use of cytochromes P450 in the production of

(+)-nootkatone was described by Sowden et al. (2005). P450cam from *Pseudomonas putida* and P450BM3 from *Bacillus megaterium* have been engineered for the oxidation of (+)-valencene. P450cam WT was not able to oxidize the (+)-valencene. Therefore, a number of mutations were introduced. The P450cam mutants oxidize (+)-valencene with high regioselectivity for C2 oxidation. The relative proportions of nootkatol and nootkatone vary from 86% (trans)-nootkatol and 4% nootkatone for the F87A/Y96 F/L244A/V247 L mutant, to 38% nootkatol and 47% nootkatone for the F87 V/Y96 F/L244A variant. On the other hand, the P450BM-3 WT showed activity towards (+)-valencene but with much lower regioselectivity (Sowden et al., 2005). CYP109B1 from *Bacillus subtilis* (Girhard et al., 2009), CYP71AV8 from chicory (*Cichorium intybus*) (Cankar et al., 2011), and the tobacco CYP71D51v2 (Gavira et al., 2013) were also applied for the oxidation of (+)-valencene. However, all these attempts resulted in multiple products, whereby (trans)-nootkatol was the main one. To overcome this problem, Wriessnegger et al (2014) identified and overexpressed an endogenous alcohol dehydrogenase (ADH) to enhance the final step of (+)-nootkatone formation in *Pichia pastoris*, which lead to > 20- fold increase yielding 7 mg/L of the desired product. This amount could be improved to 17 mg /L within 48 h by coexpressing truncated hydroxymethylglutaryl-CoA reductase (tHmg1p) of *Saccharomyces cerevisiae* (Wriessnegger et al., 2014).

To the best of our knowledge, the BMD_2094 is the first identified dehydrogenase of bacterial origin, which can efficiently oxidize (trans)-nootkatol to (+)-nootkatone. Moreover, in view of the biotechnological application, the observed production of 44 mg/L (+)-nootkatone after only less than 1 h impressively underlines the value of this novel dehydrogenase for the development of efficient whole-cell systems for the production of (+)-nootkatone.

Author contributions

M.M. carried out all experiments, analyzed and interpreted the data and drafted the manuscript. P.H. and A.G. carried out part of the cloning experiments and participated in the analysis and discussion of experimental results. R.B. participated in the interpretation and discussion of experimental results and writing of the manuscript.

Conflict of interest

The authors declare that they have no competing interest.

Acknowledgment

The authors would like to gratefully acknowledge Prof. Dr. Danièle Werk-Reichhart (Institute of Plant Molecular Biology of CNRS UPR2357, Strasbourg) for kind supply of (trans)-nootkatol. The authors

would also like to acknowledge Dr. Varun Giri (Biological Experimental Physics, Department of Physics, Saarland University) for assistance with GC-MS measurements.

References

- Biedendieck, R., Borgmeier, C., Bunk, B., Stammen, S., Scherling, C., Meinhardt, F., Wittmann, C., Jahn, D., 2011. Systems biology of recombinant protein production using *Bacillus megaterium*. *Methods in Enzymology*. Elsevier, pp. 165–195.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- β -boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl. Microbiol. Biotechnol.* 93, 1135–1146.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2013. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1703–1717.
- Cankar, K., van Houwelingen, A., Bosch, D., Sonke, T., Bouwmeester, H., Beekwilder, J., 2011. A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. *FEBS Lett.* 585, 178–182.
- de Kraker, J.-W., Franssen, M.C.R., Dalm, M.C.F., de Groot, A., Bouwmeester, H.J., 2001. Biosynthesis of Germacrene A Carboxylic Acid in Chicory Roots. Demonstration of a cytochrome P450 (+)-germacrene a hydroxylase and NADP+ -dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. *Plant Physiol.* 125, 1930–1940.
- Fraatz, M.A., Berger, R.G., Zorn, H., 2009. Nootkatone - a biotechnological challenge. *Appl. Microbiol. Biotechnol.* 83, 35–41.
- Gavira, C., Höfer, R., Lesot, A., Lambert, F., Zucca, J., Werck-Reichhart, D., 2013. Challenges and pitfalls of P450-dependent (+)-valencene bioconversion by *Saccharomyces cerevisiae*. *Metab. Eng.* 18, 25–35.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Factor.* 14, 107.
- Gerber, A., Milhim, M., Hartz, P., Zapp, J., Bernhardt, R., 2016. Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one. *Metab. Eng.* 36, 19–27.
- Girhard, M., Machida, K., Itoh, M., Schmid, R.D., Arisawa, A., Urlacher, V.B., 2009. Regioselective biooxidation of (+)-valencene by recombinant *E. coli* expressing CYP109B1 from *Bacillus subtilis* in a two-liquid-phase system. *Microb. Cell Factor.* 8, 36.
- Kallberg, Y., Oppermann, U., Jörnval, H., Persson, B., 2002. Short-chain dehydrogenases/reductases (SDRs). *Eur. J. Biochem.* 269, 4409–4417.
- Kallberg, Y., Oppermann, U., Persson, B., 2010. Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models: SDR classification using HMM. *FEBS J.* 277, 2375–2386.
- Leonhardt, R.-H., Berger, R.G., 2014. Nootkatone. In: Schrader, J., Bohlmann, J. (Eds.), *Biotechnology of Isoprenoids*. Springer International Publishing, Cham, pp. 391–404.
- Okamoto, S., Yu, F., Harada, H., Okajima, T., Hattan, J., Misawa, N., Utsumi, R., 2011. A short-chain dehydrogenase involved in terpene metabolism from Zingiber zerumbet. *FEBS J.* 278, 2892–2900.
- Schulz, S., Girhard, M., Gaßmeyer, S.K., Jäger, V.D., Schwarze, D., Vogel, A., Urlacher, V.B., 2015. Selective enzymatic synthesis of the grapefruit flavor (+)-nootkatone. *ChemCatChem* 7, 601–604.
- Sowden, R.J., Yasmin, S., Rees, N.H., Bell, S.G., Wong, L.-L., 2005. Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450cam and P450BM-3. *Org. Biomol. Chem.* 3, 57–64.
- Wriessnegger, T., Augustin, P., Engleder, M., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Zellnig, G., Schwab, H., Pichler, H., 2014. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. *Metab. Eng.* 24, 18–29.

2.5 (Hartz et al., 2018)

Characterization and engineering of a carotenoid biosynthesis operon from *Bacillus megaterium*.

Hartz, P., Milhim, M., Trenkamp, S., Bernhardt, R. and Hannemann, F.

Metabolic Engineering. 2018 Jul; 49:77-58

DOI: 10.1016/j.ymben.2018.07.017

Reprinted with permission of Metabolic Engineering. All rights reserved.



Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

Characterization and engineering of a carotenoid biosynthesis operon from *Bacillus megaterium*

Philip Hartz^a, Mohammed Milhim^a, Sandra Trenkamp^b, Rita Bernhardt^a, Frank Hannemann^{a,*}^a Department of Biochemistry, Saarland University, Campus B 2.2, 66123 Saarbrücken, Germany^b Metabolomic Discoveries GmbH, Am Mühlberg 11, 14476 Potsdam, Germany

ARTICLE INFO

Keywords:

Bacillus megaterium
Carotenoid biosynthesis
Prenyl diphosphate synthase
Diapophytoene synthase
Diapophytoene desaturase
Diaponeurosporene

ABSTRACT

Bacillus megaterium belongs to the group of pigmented bacilli producing carotenoids that ensure self-protection from UV radiation-induced and collateral oxidative damage. Metabolite profiling of strain MS941 revealed the presence of the C30 carotenoids 4,4'-diapophytofluene and 4,4'-diaponeurosporenic acid. A gene function analysis demonstrated the presence of a corresponding C30 carotenoid biosynthetic pathway with pharmaceutical importance. We identified a gene cluster comprising putative genes for a farnesyl diphosphate synthase (IspA), a diapophytoene synthase (CrtM) and three distinct diapophytoene desaturases (CrtN1–3). Intriguingly, *crtM* was organized in an operon together with two of the identified *crtN* genes. The individual activities of the encoded enzymes were determined by heterologous expression and product analysis in the non-carotenogenic model organism *Escherichia coli*. Our experimental data show that the first catalytic steps of C30 carotenoid biosynthesis in *B. megaterium* share significant similarity to the corresponding biosynthetic pathway of *Staphylococcus aureus*. The biosynthesis of farnesyl diphosphates and their subsequent condensation to form 4,4'-diapophytoene are catalyzed by the identified IspA and CrtM, respectively. The following desaturation reactions to form 4,4'-diaponeurosporene, however, require the activities of multiple diapophytoene desaturases. A biosynthetic operon was engineered and successfully expressed in an *E. coli* whole-cell system creating a cell factory for a high-yield production of the C30 carotenoid 4,4'-diaponeurosporene which has promising potential in the treatment of various inflammatory diseases.

1. Introduction

Since the first carotenoid was isolated in 1831, the number of identified carotenoid structures has continuously increased to a total of over 1100 structures (Wackenroder, 1831; Yabuzaki, 2017). Carotenoids are ubiquitously found in all domains of life and thus considered to form the most diverse and widespread class of natural pigments. The physiological functions of carotenoids are as diverse as their structures but can principally be assigned to three categories: photosynthesis, photoprotection and nutrition (Kirti et al., 2014).

During photosynthesis, carotenoids mainly serve as auxiliary pigments in the light harvesting complex, where they help to extend the spectral range that can be used for photosynthesis. But they are also involved in the protection of the photosystems against excessive radiation and the resulting oxidative damage (Frank and Cogdell, 1996).

Non-photosynthetic organisms, including humans, can also benefit from these antioxidant properties. The accumulation of carotenoids in the lipid bilayer represents an important factor for the protection of unsaturated fatty acids and lipoproteins against oxidative damage, originating from reactive oxygen species (ROS) (Britton, 1995). Furthermore this incorporation significantly affects the fluidity, permeability as well as the stability of biological membranes (Gruszecki and Strzałka, 2005). Since humans lack the ability for carotenoid production, they are dependent on dietary supply of essential carotenoids like provitamin A. After metabolic activation, these carotenoids are involved in many important physiological processes including the vision, reproduction, cell growth and cell differentiation as well as immune modulation (Chew and Park, 2004; McDevitt et al., 2005; Zile and Cullum, 1983). Numerous scientific studies furthermore confirm the inhibitory effect of carotenoids on the progress of cancer,

Abbreviations: BmFDS, *Bacillus megaterium* farnesyl diphosphate synthase; CrtM, diapophytoene synthase; CrtN, diapophytoene desaturase; CrtP, diaponeurosporene oxygenase; DNSP, diaponeurosporene; DPFL, diapophytofluene; DPHY, diapophytoene; FDP, farnesyl diphosphate; GDP, geranyl diphosphate; GGDP, geranylgeranyl diphosphate; IDP, isopentenyl diphosphate; PDA, photodiode array

* Corresponding author.

E-mail address: f.hannemann@mx.uni-saarland.de (F. Hannemann).<https://doi.org/10.1016/j.ymben.2018.07.017>

Received 26 April 2018; Received in revised form 11 July 2018; Accepted 24 July 2018

Available online 26 July 2018

1096-7176/ © 2018 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

cardiovascular and neurodegenerative diseases (Gerster, 1993; Rao and Agarwal, 2000; Obulesu et al., 2011; Miyake et al., 2011). Due to these extraordinary properties, the economic and commercial interest in carotenoids is high (<http://www.bccresearch.com/report/FOD025E.html>). Various strategies have been developed to identify new carotenoids and enable their biosynthesis in carotenogenic as well as non-carotenogenic microorganisms (Rodríguez-Sáiz et al., 2010; Yoshida et al., 2009; Misawa et al., 1990; Garrido-Fernández et al., 2010). In the course of these studies, a plethora of mostly endospore forming bacteria has been found to be able to synthesize different C30 carotenoids and corresponding C30 apocarotenoids with promising properties concerning stability and anti-inflammatory potential (Perez-Fons et al., 2011; Steiger et al., 2012; Takaichi et al., 1997; Köcher et al., 2009). Treatment of mice with the C30 carotenoid 4,4'-diaponeurosporene, for example, had a beneficial effect on inflammatory diseases of the gastrointestinal tract (Jing et al., 2017). Furthermore 4,4'-diaponeurosporene was demonstrated to effectively stimulate immune response, thereby increasing resistance to Salmonella infections (Liu et al., 2016, 2017).

Although various C30 carotenoids have been identified so far, the lack of genomic data limited the elucidation of the underlying biosynthetic pathways for the most time to the staphyloxanthin producing pathogen *Staphylococcus aureus* (Marshall and Wilmoth, 1981; Wieland et al., 1994; Pelz et al., 2005; Steiger et al., 2012). Recently, functional assignment of carotenogenic genes from *Bacillus firmus* and *Bacillus indicus* provided new insights into the biosynthetic pathways for C30 carotenoids in *Bacillus* species (Steiger et al., 2015). There is evidence from prior studies that also the non-pathogenic Gram-positive bacterium *Bacillus megaterium* is able to accumulate different pigments during spore formation but these pigments have not been identified yet (Racine and Vary, 1980; Mitchell et al., 1986).

In this work, we used the published whole genome sequences of the *B. megaterium* strain DSM319 not only to identify, but also to elucidate the underlying biosynthetic pathway for carotenoid production in the biotechnologically important *B. megaterium* strain MS941 (DSM319 $\Delta nprM$) (Eppinger et al., 2011; Wittchen and Meinhardt, 1995). The genetically engineered carotenoid operons were successfully applied to establish an *Escherichia coli* based microbial cell factory for the selective production of 4,4'-diaponeurosporene which has promising potential as novel drug against various inflammatory diseases.

2. Materials and methods

2.1. Bacterial strains, expression vectors, chemicals and enzymes

Bacterial strains and expression vectors used in this study were listed in Table 1. *B. megaterium* strain MS941 was used for the preparation of genomic DNA and subsequent amplification of the ORFs located in the putative carotenoid operon. *E. coli* strain TOP10 (Invitrogen, Karlsruhe, Germany) was used for the assembly of all expression vectors. *E. coli* strains BL21 (DE3) (Novagen, Darmstadt, Germany), C43 (DE3) (Lucigen, Heidelberg, Germany) and JM109 (Promega, Mannheim, Germany) were used for the heterologous protein expression and protein purification.

The isolation kit for bacterial DNA was acquired from nextec Biotechnologie GmbH (Hilgertshausen, Germany). All restriction enzymes were obtained from New England Biolabs (Ipswich, USA). Fast-Link DNA Ligation Kit was purchased from Epicentre (Madison, USA). LB broth, yeast extract and tryptone were obtained from Becton Dickinson (Franklin Lakes, USA). Isopropyl- β -D-thiogalactopyranoside (IPTG) was acquired from Carbolution Chemicals GmbH (Saarbrücken, Germany). HPLC grade acetonitrile, methanol and isopropanol were obtained from VWR. All other chemicals were purchased from Sigma-Aldrich (St. Louis, USA).

2.2. Molecular cloning

The assembly of all s was performed with standard cloning procedures using restriction enzyme digestion and ligation. Primers that were used for the amplification of the putative carotenogenic ORFs are listed in Table 2.

For the heterologous expression and characterization of the putative farnesyl diphosphate synthase BMD_4442, the corresponding DNA fragment was PCR amplified from genomic DNA of *B. megaterium* and inserted into pET17b expression vector using *NdeI* and *NotI* restriction sites. The reverse primer was used to extend the 3' end of the cDNA with a hexahistidine coding sequence which enabled the subsequent protein purification via Immobilized Metal Ion Affinity Chromatography (IMAC). For the characterization of the predicted carotenogenic open reading frames, the putative diapophytoene synthase (BMD_0659), the putative diapophytoene desaturases (BMD_0657 and BMD_0658) as well as the putative diaponeurosporene oxygenase, were assembled in different combinations resulting in eight synthetic operons covering the aforementioned ORFs.

Each synthetic operon was assembled in a pSMF.1 vector, which allows constitutive protein expression in *E. coli* cells. The farnesyl diphosphate synthase BMD_4442 was the first DNA fragment that was PCR-amplified from genomic DNA of *B. megaterium* and inserted into pSMF.1 vector using *PacI* and *SpeI* restriction sites. The forward primer furthermore included a modified *xylA* ribosomal binding site (RBS, underlined letters) (AATCAAGGAGGTGAATGTACA) plus linker (Stammen et al., 2010). The resulting vector pSMF.A was used as backbone for all consecutive cloning steps. The carotenogenic open reading frames were amplified individually from genomic DNA of *B. megaterium* strain MS941 and inserted into pSMF.A vector using primers with the modified *xylA* RBS as well as specific combinations of restriction enzymes. The putative diapophytoene synthase BMD_0659 was inserted using *SpeI* and *MluI* restriction sites. The predicted diapophytoene desaturases BMD_0657 and BMD_0658 were inserted using *MluI/SacI* and *SacI/NotI* restriction sites, respectively. *NotI* and *AgeI* restriction sites were used for the insertion of the putative diaponeurosporene oxygenase BMD_1977.

2.3. Heterologous expression and purification of putative prenyl transferase BMD_4442

For the heterologous protein expression of the putative prenyl transferase BMD_4442, *E. coli* BL21 (DE3) cells were transformed with the expression vector pET17b.4442. A preculture was started from a single colony in LB medium (25 g/L), supplemented with ampicillin (100 μ g/ml) and incubated overnight at 37 °C and 140 rpm. A total volume of 200 ml TB medium (24 g/L yeast extract, 12 g/L tryptone, 0.4% glycerol, 100 mM potassium phosphate buffer, pH 7.4) was supplemented with appropriate antibiotics and inoculated with 1:100 of the preculture. The main culture was grown in a 2 L baffled flask at 37 °C and 140 rpm to an optical density of 0.8 at 600 nm. Protein expression was induced by addition of 1 mM IPTG and temperature was shifted to 30 °C. Cells were further grown at 30 °C and 140 rpm for 20 h. After heterologous protein expression, cells were harvested by centrifugation at 4000g and 4 °C for 15 min. For purification, the cell pellet was suspended in 25 ml lysis buffer (50 mM potassium phosphate, pH 7.4, 10% glycerol) and sonicated on ice. After centrifugation of the lysate at 30,000g and 4 °C for 30 min, the supernatant was loaded on a Ni-NTA agarose column that was equilibrated with lysis buffer. The column was washed twice with 50 ml of lysis buffer supplemented with 20 mM imidazole and 40 mM imidazole, respectively. The protein was eluted with lysis buffer supplemented with 200 mM imidazole. The eluted protein was dialyzed against lysis buffer without imidazole and stored at –80 °C. Homogeneity of the purified protein was checked by SDS-PAGE. Protein concentration was determined with bicinchoninic acid (BCA) assay (Interchim, Montluçon, France).

Table 1
List of bacterial strains and vectors used in this study.

Bacterial strain	Description	Reference
<i>E. coli</i> TOP10	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80 <i>lacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>recA1</i> <i>araD139</i> Δ(<i>ara leu</i>) 7697 <i>galU</i> <i>galK</i> <i>rpsL</i> (StrR) <i>endA1</i> <i>nupG</i>	Invitrogen (Karlsruhe, Germany)
<i>E. coli</i> JM109	<i>endA1</i> , <i>recA1</i> , <i>gyrA96</i> , <i>thi</i> , <i>hsdR17</i> (<i>r_K⁻</i> , <i>m_K⁺</i>), <i>relA1</i> , <i>supE44</i> , Δ(<i>lac-proAB</i>), [<i>F'</i> <i>traD36</i> , <i>proAB</i> , <i>laq1</i> Δ <i>M15</i>]	Promega (Mannheim, Germany)
<i>E. coli</i> BL21 (DE3)	F ⁻ <i>ompT</i> <i>gal</i> <i>dcm</i> <i>lon</i> <i>hdsS_B</i> (<i>r_B⁻</i> <i>m_B⁻</i>) λ(DE3 [<i>lacI</i> <i>lacUV5</i> - <i>T7p07</i> <i>ind1</i> <i>sam7</i> <i>nin5</i>)] [<i>malB⁺</i>] _{K-12} (λ ^S)	Novagen (Darmstadt, Germany)
<i>E. coli</i> C43 (DE3)	F ⁻ <i>ompT</i> <i>hdsS_B</i> (<i>r_B⁻</i> <i>m_B⁻</i>) <i>gal</i> <i>dcm</i> (DE3)	Lucigen (Heidelberg, Germany)
<i>B. megaterium</i> MS941	Mutant of DSM319; Δ <i>nprM</i> (extracellular protease)	Wittchen and Meinhardt (1995)
name	Description	Reference
pET17b	Vector with IPTG inducible promoter system for protein expression in <i>E. coli</i>	Novagen (Darmstadt, Germany)
pSMF2.1	Vector for constitutive protein expression in <i>E. coli</i> ; vector with xylose inducible promoter system for protein expression in <i>B. megaterium</i>	Bleif et al. (2012)
pET17b.4442	Expression and purification of BMD_4442	This study
pSMF.A	Expression of BMD_4442	This study
pSMF.AM	Expression of BMD_4442 + BMD_0659	This study
pSMF.AM ₁	Expression of BMD_4442 + BMD_0659 + BMD_0657	This study
pSMF.AM ₂	Expression of BMD_4442 + BMD_0659 + BMD_0658	This study
pSMF.AM ₁₂	Expression of BMD_4442 + BMD_0659 + BMD_0657 + BMD_0658	This study
pSMF.AM ₁₂₃	Expression of BMD_4442 + BMD_0659 + BMD_0657 + BMD_0658 + BMD_1977	This study
pSMF.AM ₃	Expression of BMD_4442 + BMD_0659 + BMD_1977	This study
pSMF.AM ₁₃	Expression of BMD_4442 + BMD_0659 + BMD_0657 + BMD_1977	This study

Table 2
List of primers used in this study. Restriction sites are shown in bold letters. Ribosomal binding sites are underlined.

Name	Sequence (5'–3')	Description
<i>NdeI</i> – 4442-for	GTGCATATGGTGAATGAAGTGAATTAAAGAG	Amplification of BMD_4442 for insertion in pET17b vector
<i>NotI</i> – 4442-rev	GTGGCGGCCGCTTACTGATGGTGTATGATGACTATCACGTTTTGCTATATATAAC	Amplification of BMD_4442 for insertion in pET17b and addition of hexahistidine tag for purification
<i>PacI</i> – 4442-for	GTGTTAATTAATAATCAAGGAGGTGAATGTACAATGGTGAATGAAGTGAATTTAAAGAG	Amplification of BMD_4442 for insertion in pSMF2.1 vector
<i>SpeI</i> – 4442-rev	GTGACTAGTTTAACTATCACGTTTTGC	Amplification of BMD_4442 for insertion in pSMF2.1 vector
<i>SpeI</i> – 0659-for	GTGACTAGTAATCAAGGAGGTGAATGTACAATGATGAACATTTCTGTAGACAAAAGC	Amplification of BMD_0659 for insertion in pSMF2.1 vector
<i>MluI</i> – 0659-rev	GTGACGCGTTTACCCCGTAGAAATAGATGTC	Amplification of BMD_0659 for insertion in pSMF2.1 vector
<i>MluI</i> – 0657-for	GTGACGCGTAATCAAGGAGGTGAATGTACAATGCAAAAAATGTAATCGTGATTGG	Amplification of BMD_0657 for insertion in pSMF2.1 vector
<i>SacI</i> – 0657-rev	GTGGAGCTCTTATAATGCACTCTCCATTTC	Amplification of BMD_0657 for insertion in pSMF2.1 vector
<i>SacI</i> – 0658-for	GTGGAGCTCAATCAAGGAGGTGAATGTACAATGAGAACAGCAATTGTTGG	Amplification of BMD_0658 for insertion in pSMF2.1 vector
<i>NotI</i> – 0658-rev	GTGGCGGCCGCTTAAAGAAAGCTTTGCAACC	Amplification of BMD_0658 for insertion in pSMF2.1 vector
<i>NotI</i> – 1977-for	GTGGCGGCCGCAATCAAGGAGGTGAATGTACAATGAAAAAGAAAGTAATTGTTGTAGG	Amplification of BMD_1977 for insertion in pSMF2.1 vector
<i>AgeI</i> – 1977-rev	GTGACCGGTTTATCCAATATCTTTTGCATACG	Amplification of BMD_1977 for insertion in pSMF2.1 vector

2.4. Prenyl transferase assay

The prenyl transferase activity of the purified protein BMD_4442 was determined in a reconstituted *in vitro* assay using IDP and GDP as substrates. The conversion was carried out in 200 μl conversion buffer (50 mM potassium phosphate buffer, pH 7.4, 10% glycerol) supplemented with 5 mM MgCl₂, 50 μM IDP and 50 μM GDP. The reaction was initiated by addition of 10 μg of the purified BMD_4442 protein. The reaction mixture was incubated at 30 °C and 1000 rpm for 60 min. Hydrolysis of the acid-labile prenyl diphosphates to their corresponding alcohols was achieved by addition of 10 μl hydrochloric acid (3 N). Samples were overlaid with 500 μl hexane and incubated at 42 °C for additional 20 min. After hydrolysis, the prenyl alcohols were extracted into the hexane phase by vigorous shaking and organic supernatants were subsequently evaporated to dryness. Authentic standards of geranyl diphosphate (GDP), farnesyl diphosphate (FDP) and geranylgeranyl diphosphate (GGDP) were prepared as described above without addition of purified BMD_4442 protein.

Samples were suspended in 10% acetonitrile. Product analysis was carried out by liquid chromatography (LC). Adequate separation of the substrates IDP and GDP as well as potential products FDP and GGDP was achieved by RP-HPLC using an EC 125/4 Nucleodur C18 column (Macherey & Nagel) on a Jasco system. The mobile phase consisted of 10% acetonitrile in water on channel A and pure acetonitrile on channel

B. The following gradient was applied for HPLC measurements: 0–15 min: 100% A to 0% A; 15.1–20 min: 0% A; 20.1–25 min: 0% A to 100% A; 25.1–30 min 100% A. The flow rate of 0.5 ml/min was temporarily raised to 1 ml/min between minutes 15.1 and 25. Column oven temperature was set to 40 °C. The injection volume of all standards and samples was 50 μl. Detection of hydrolyzed isoprenoids was carried out at a wavelength of 214 nm.

2.5. Functional characterization of the carotenoid biosynthesis operon from *B. megaterium* in *E. coli*

The *E. coli* strain TOP10 was used for the biosynthesis of carotenoids. Cells were transformed with the expression vector pSMF2.1 carrying the engineered carotenoid genes in different combinations. The cultivation was performed in LB medium supplemented with appropriate antibiotics for 20 h at 37 °C and 140 rpm. Induction of protein expression was not necessary due to the constitutive promoter.

Prior to carotenoid extraction, bacterial cell pellets were washed twice with potassium phosphate buffer (50 mM, pH 7.4). After each washing step, cells were centrifuged for 5 min at 8000g. The supernatant was discarded and the carotenoids were extracted twice with 2 ml methanol. The carotenoid extracts were evaporated to dryness and dissolved in methanol and isopropanol (1:1) for subsequent RP-HPLC analysis. Separation of carotenoid intermediates was carried out by RP-

Table 3
Protein sequence identities and similarities of putative carotenogenic ORFs of *B. megaterium* with corresponding proteins of *S. aureus*.

Protein in <i>Staphylococcus aureus</i>	Homologous ORF in <i>B. megaterium</i>	Identity/Similarity [%]
Farnesyl diphosphate synthase (IspA)	BMD_4442	50/66
Diapophytoene synthase (CrtM)	BMD_0659	31/52
Diapophytoene desaturase (CrtN)	BMD_0657	32/54
	BMD_0658	31/55
	BMD_1977	30/56
Diaponeurosporene oxygenase (CrtP)	BMD_1977	45/63

HPLC (Jasco) using an EC 125/4 Nucleodur C18 column (Macherey & Nagel). Extracts were loaded on the column and eluted by methanol and isopropanol (1:1) in isocratic mode using a flow rate of 0.5 ml/min for 15 min. The injection volume was 100 µl and the detection and identification of different carotenoid intermediates was carried out on basis of their unique absorption spectra, described by Takaichi, using a photo diode array (PDA) detector (Takaichi, 2000).

2.6. Carotenoid profiling of *Bacillus megaterium* MS941

For carotenoid profiling, 500 ml of modified M9CA medium were inoculated 1:100 with a *B. megaterium* preculture in a Minifors (Infors HT) bioreactor. M9CA medium (pH = 7.0) consisted of 6 g Na₂HPO₄, 3 g KH₂PO₄, 0.5 g NaCl, 1 g NH₄Cl, 2 g casamino acids, 30 g glycerol, 100 µl 1 M CaCl₂, 2 ml 1 M MgSO₄, 100 µl 10 mM thiamine hydrochloride and 1 ml of 1000 x trace elements solution per liter H₂O. 1000 x trace elements solution contained 2.5 g EDTA, 695 mg FeSO₄·7H₂O, 99 mg MnCl₂·4H₂O, 68 mg ZnCl₂, 24 mg CoCl₂·6H₂O, 17 mg CuCl₂·2H₂O, 24 mg NiCl₂·6H₂O, 24 mg Na₂MoO₄·2H₂O, 26 mg Na₂SeO₃·5H₂O and 6 mg H₃BO₃ per 50 ml H₂O. *B. megaterium* cells were cultured for 24 h at 37 °C and constant aeration. Samples were taken at different stages of bacterial growth.

For metabolite extraction, bacteria pellets were suspended in 80% methanol (−20 °C containing internal standards) to a final concentration of 100 mg pellet (wet weight) per ml. Glass beads were added and cells were disrupted by 5 cycles of 30 s vortexing and 30 s cooling on ice. Samples were centrifuged at 4 °C at 4000g for 15 min. An aliquot of 300 µl was taken and stored at −80 °C until LC-MS analysis. All subsequent steps were carried out at Metabolomic Discoveries GmbH (Potsdam, Germany; www.metabolomicdiscoveries.com). Metabolites were identified in comparison to Metabolomic Discoveries' database entries of authentic standards. UPLC-QTOF/MS analysis was performed using a iHILIC®-Fusion HILIC Column, 100 × 2.1 mm, 3.5 µm, 100 Å (DiChrom), operated by an Agilent 1290 UPLC system (Agilent, Santa Clara, USA). The LC mobile phase was A) 10 mM Ammoniumacetat (Sigma-Aldrich, USA), pH 6 in water (Thermo, USA) with 5% Acetonitrile (Thermo, USA) and B) Acetonitrile (Thermo, USA) with 5% 10 mM Ammoniumacetat (Sigma-Aldrich, USA), pH 6 in water (Thermo, USA) with a gradient from 5% A to 35% from 1 min to 8.5 min, to 95% at 9.5 min, isocratic at 95% to 12 min, to 5% at 12.01 min until 15 min. The flow rate was 300 µl/min with an injection volume of 1 µl.

3. Results and discussion

3.1. The putative farnesyl diphosphate synthase BMD_4442 is able to synthesize farnesyl diphosphate but not geranylgeranyl diphosphate

Carotenoids are predominantly found in plants as well as fungi and selected bacteria (Nisar et al., 2015). The first committed step in carotenoid biosynthesis is the condensation of two molecules of prenyl

diphosphates to form the characteristic polyene structure of carotenoids (Norris et al., 1995). The number of carbon atoms in the resulting carotenoid structure is strongly dependent on the chain lengths of the prenyl diphosphates that are used for this condensation reaction. C40 carotenoids, for example, are formed by head-to-head condensation of two molecules of geranylgeranyl diphosphate (C20) (Wieland et al., 1994). Biosynthesis of C30 carotenoids, however, can take place in two different ways: symmetric C30 carotenoids, also called 4,4'-diapocarotenoids, are synthesized by head-to-head condensation of two molecules of farnesyl diphosphate (C15), whereas asymmetric C30 carotenoids, like apo-8'-carotenoids, are formed accordingly by condensation of geranyl diphosphate (C10) and geranylgeranyl diphosphate (C20) (Perez-Fons et al., 2011). Furthermore, unusual carotenoids, like C35 carotenoids, are built from farnesyl diphosphate (C15) and geranylgeranyl diphosphate (C20) (Umeno and Arnold, 2003).

Consequently, chain length determination of prenyl diphosphates is crucial to predict the spectrum of carotenoids that are expected to be formed in a specific organism. For this reason we aimed to identify and characterize the prenyl diphosphate synthase from the industrially important *B. megaterium* strain MS941 in order to evaluate the uncharacterized carotenoid pigments that have been described before (Racine and Vary, 1980; Mitchell et al., 1986).

Prenyl diphosphate synthases from different organisms share several conserved domains, most notably two characteristic aspartate-rich amino acid motifs. The first aspartate-rich motif (FARM) is involved in the chain length determination of prenyl diphosphate products, while the second aspartate-rich motif (SARM) is additionally associated with substrate orientation and catalytic activity (Szkopińska and Plochocka, 2005). Both characteristic motifs were found to be present in the open reading frame BMD_4442, which shares significant sequence identity and similarity to prenyl diphosphate synthases, particularly to the well characterized farnesyl diphosphate synthase (IspA) of *S. aureus* (Table 3). Furthermore, the putative prenyl diphosphate synthase BMD_4442 is located upstream of the 1-deoxy-D-xylulose-5-phosphate synthase gene (*dxs*) which catalyzes the initial step of isoprenoid biosynthesis, ultimately yielding the substrates for the prenyl diphosphate synthase reaction itself (Fig. 4A).

The open reading frame of the putative farnesyl diphosphate synthase BMD_4442 was inserted into the expression vector pET17b and heterologously expressed in *E. coli* BL21 (DE3). As shown in Fig. 1, a single step purification procedure using a Ni-NTA column was sufficient to purify a soluble form of the recombinant BMD_4442 protein with an apparent molecular mass of 36 kDa which is slightly above the

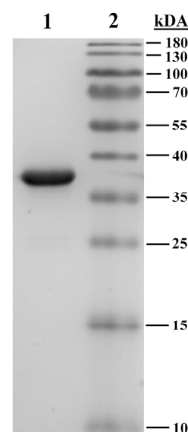


Fig. 1. 15% SDS-PAGE of purified BMD_4442. Lane 1 contains 5 µg of the purified protein and lane 2 shows the band profile of the PageRuler prestained protein ladder (Thermo Fisher Scientific, USA).

calculated molecular mass of approximately 34 kDa. According to the BCA assay, a total of 75 mg recombinant protein was purified from one liter expression culture.

The prenyl transferase assay was performed according to Schmidt and coworkers with slight modifications as described in the material and method section (Schmidt et al., 2010). Hydrolysis of the authentic standards GDP, FDP as well as GGDP was used to transform the different prenyl diphosphates to their corresponding prenyl alcohols. Since prenyl diphosphates may undergo isomerization during this hydrolyzation reaction, multiple hydrolyzation products can be formed: hydrolyzation of GDP was described to preferentially yield geraniol and its linear structural isomer linalool, whereas α -terpineol (a cyclic isomer of geraniol) is formed in only very small amounts, probably below detectable levels regarding RP-HPLC analysis (Alarcon et al., 1992; Cramer and Rittersdorf, 1967). Similarly, hydrolyzation of FDP was reported to generate farnesol and racemic nerolidol as main products (George-Nascimento et al., 1971). To the best of our knowledge, there is no comparable data for the hydrolyzation reaction of GGDP described in literature. But, equivalently to GDP as well as FDP, GGDP is supposed to form predominantly geranyl geraniol and geranyl linalool. Good separation of all standards was achieved by RP-HPLC with retention times of 8–9 min for GDP, 15–16 min for FDP and 17.5 min for GGDP hydrolyzation products, respectively (Fig. 2).

Incubation of the purified prenyl transferase BMD_4442 with GDP and IDP as substrates showed a comparable product pattern to the hydrolyzed FDP standard, indicating that BMD_4442 exhibited FDP synthase activity. A GGDP synthase activity for BMD_4442 can be excluded, since no GGDP product was detectable (Fig. 2). Consequently, the protein encoded by the open reading frame BMD_4442 will be referred to as *Bacillus megaterium* Farnesyl Diphosphate Synthase (BmFDS). These results suggest that carotenoid formation in *B. megaterium* seems to be limited to the condensation of two molecules of farnesyl diphosphate yielding the symmetric C30 carotenoid structure of 4,4'-diapocarotenoids. Due to the fact that no GGDP was detectable, condensation of GDP and GGDP to form asymmetric C30 carotenoid structures of apo-8'-carotenoids can be excluded. Similarly, formation of C35 carotenoids as well as C40 carotenoids is not possible, since both biosynthetic pathways also rely on GGDP as substrate for the initial condensation reaction.

3.2. The C30 carotenoids 4,4'-diapophytofluene and 4,4'-diaponeurosporenic acid can be detected in the metabolome of *B. megaterium* MS941

Previous studies have shown that various spore forming bacteria from wildly diverse habitats are able to produce pigments of different colors. In most cases, these pigments were identified as C30 diapocarotenoids and corresponding C30 apocarotenoids (Khaneja et al., 2010; Perez-Fons et al., 2011). Their physiological functions in non-photosynthetic microorganisms, such as *Bacillus* species, are mostly limited to photoprotection, providing extraordinary resistance to excessive radiation and the resulting oxidative damage (Moeller et al., 2005). However, current investigations postulate that C30 carotenoids may also play a crucial role as virulence factors in the life cycle of the pathogen *Staphylococcus aureus*, modulating the degree of the host's immune response (Gao et al., 2017). Scientific reports from the early 1980s by Racine and Vary as well as Mitchell and coworkers also describe the presence of natural pigments in cells of *B. megaterium* (Racine and Vary, 1980; Mitchell et al., 1986). Encouraged by their findings, we were aiming not only to identify these carotenoid pigments in the industrially important *B. megaterium* strain MS941, but also to characterize the involved enzymes, thereby elucidating the underlying biosynthetic pathway.

For the purpose of metabolic profiling of carotenoid accumulation in *B. megaterium*, cells were cultured in a Minifors (Infors HT) bioreactor. Samples representing different stages of bacterial growth were

prepared as described in the material and method section and subjected to MS analysis. Carotenoids were identified according to their characteristic m/z ratios. Although the overall abundance of carotenoids seemed to be very low, it was possible to detect two distinct C30 carotenoid compounds in the metabolome of *B. megaterium* MS941 (Fig. 3).

Based on the results of the prenyl transferase assay, carotenoid biosynthesis in *B. megaterium* MS941 was predicted to be limited to the formation of symmetric C30 structures of 4,4'-diapocarotenoids. Combined analyses of mass to charge ratios and retention times allowed to identify these C30 carotenoids as 4,4'-diapophytofluene and 4,4'-diaponeurosporenic acid. 4,4'-diapophytofluene showed a retention time of 50–55 s and a characteristic m/z of 406.36 in positive ion mode. The oxygenated derivative 4,4'-diaponeurosporenic acid showed a delayed retention time of 54–59 s, most likely due to the stronger interaction with the UPLC column, and a typical m/z of 432.29 in negative ion mode. Both carotenoids have been described as intermediates of the C30 carotenoid pathway of many unrelated bacteria, however, most prominent as precursors for the biosynthesis of staphyloxanthin in *Staphylococcus aureus* (Pelz et al., 2005; Kim and Lee, 2012). Even though there was no indication of *B. megaterium* MS941 being able to produce additional carotenoids under the tested growth conditions, glycosylation and subsequent esterification of 4,4'-diaponeurosporenic acid towards the biosynthesis of staphyloxanthin related compounds should not be excluded, particularly with regards to studies that postulate a strong association of final carotenoid product formation with sporulation in pigmented *Bacillus* species (Duc et al., 2006).

3.3. Identification of a putative pathway for C30 carotenoid biosynthesis

Based on the finding, that *B. megaterium* MS941 is able to produce the C30 carotenoids 4,4'-diapophytofluene and 4,4'-diaponeurosporenic acid under the tested growth conditions, we wanted to identify the corresponding genes that are involved in the biosynthetic pathway. Bioinformatic analyses were performed using the deduced amino acid sequences of the well characterized staphyloxanthin biosynthetic operon from *Staphylococcus aureus* as templates for homology searches [GenBank accession number: X97985]. This biosynthetic operon comprises five genes that are necessary for the production of the carotenoid pigment staphyloxanthin, which is essentially a glycosylated form of 4,4'-diaponeurosporenic acid (Scheme 1).

The biosynthesis of 4,4'-diaponeurosporenic acid, however, requires the enzymatic activity of just three of the five genes in the staphyloxanthin operon, starting with a diapophytoene synthase (CrtM) [UniProt accession number: A9JQL9] that catalyzes the symmetric head-to-head condensation of two molecules of farnesyl diphosphate to form the C30 carotenoid 4,4'-diapophytoene. This condensation is followed by a stepwise dehydrogenation reaction to form 4,4'-diaponeurosporene (Wieland et al., 1994). All dehydrogenation steps are catalyzed by the same diapophytoene desaturase (CrtN) [UniProt accession number: O07855]. The subsequent terminal oxidation of 4,4'-diaponeurosporene to form 4,4'-diaponeurosporenic acid requires the activity of a mixed diaponeurosporene oxidase (CrtP) [UniProt accession number: Q2FDU3] (Pelz et al., 2005). Based on sequence similarities, corresponding homologous to the deduced amino acid sequences of *S. aureus* CrtM, CrtN as well as CrtP were identified in the genome of the *B. megaterium* strain MS941 (Table 3).

The open reading frame BMD_0659 has a calculated molecular weight of 32.9 kDa and is considered to be a putative diapophytoene synthase because it shares significant sequence identity of 31% and sequence similarity of 52% to the *S. aureus* diapophytoene synthase CrtM. The predicted protein furthermore shows multiple conserved aspartate rich amino acid motifs [D(D)xxD] that were shown to be crucial for binding and orientation of farnesyl diphosphates in the active site of different prenyl synthases (Liu et al., 2008). Most striking, however, is the fact that this open reading frame is arranged in an operon together with two of the identified putative diapophytoene

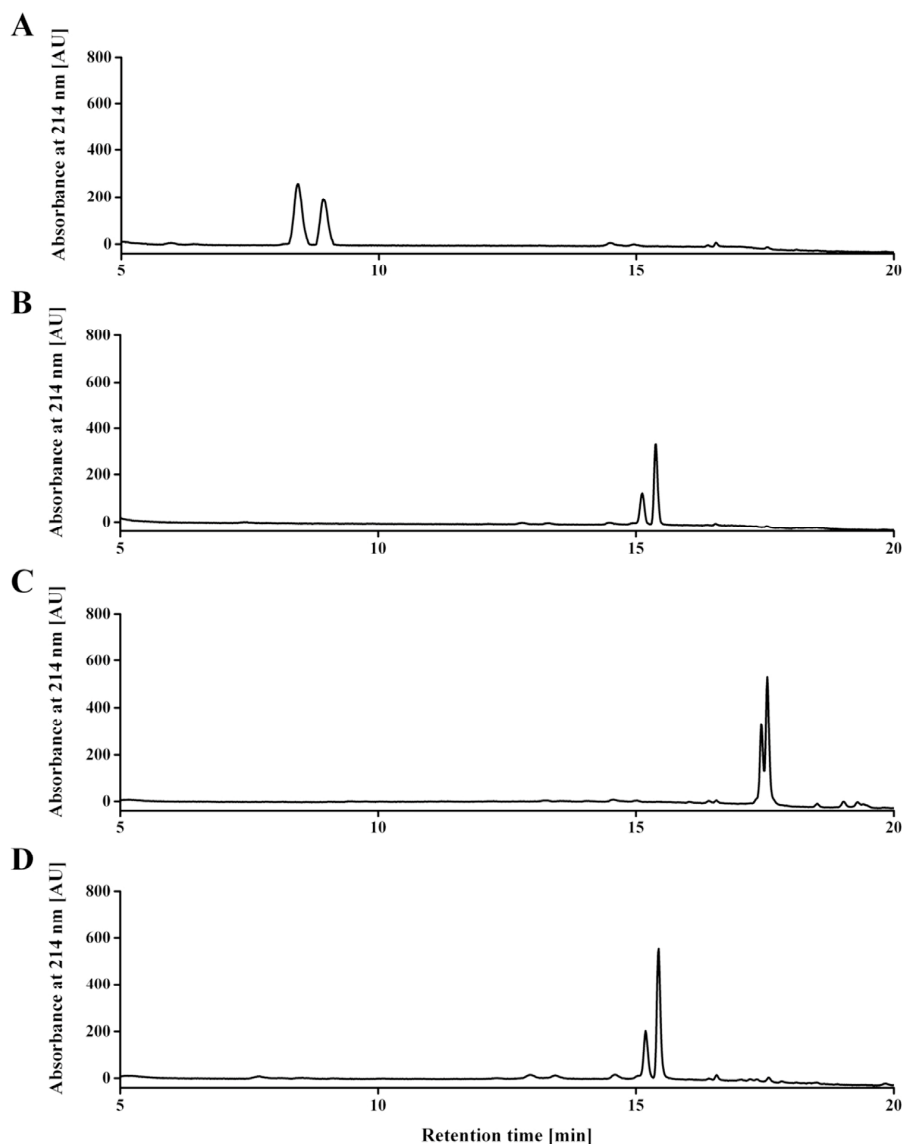


Fig. 2. RP-HPLC analysis of prenyl transferase assay. Authentic standards of **A:** geranyl diphosphate (GDP), **B:** farnesyl diphosphate (FDP), **C:** geranylgeranyl diphosphate (GGDP). **D:** prenyl transferase product synthesized by purified BMD_4442 when incubated with isopentenyl diphosphate (IDP) and geranyl diphosphate (GDP). Authentic standards and reaction products were hydrolyzed as mentioned in the method [Section 2.4](#). The resulting prenyl alcohols show double peaks representing the expected structural isomers. The corresponding retention orders were not determined.

desaturases ([Fig. 4](#))

Comparison of the deduced amino acid sequence of *S. aureus* diapophytoene desaturase (CrtN) revealed the presence of multiple putative diapophytoene desaturases in the genome of *B. megaterium*. As mentioned before, two of them, BMD_0657 as well as BMD_0658, are localized immediately upstream of the putative diapophytoene synthase BMD_0659, while a third putative diapophytoene desaturase BMD_1977 seems to be not only spatially, but probably also functionally separated from the other diapophytoene desaturases. BMD_0657 is a protein of 503 aa and has a calculated molecular weight of 57.8 kDa, which is comparable to CrtN of *S. aureus*. They share a sequence identity of 31% and sequence similarity of 55% as well as a characteristic N-terminal amino acid motif [GxGxxG(x)₁E] as part of the Rossmann fold that allows binding of the dinucleotide cofactor FAD, which is crucial for the putative desaturase activity ([Dym and Eisenberg, 2001](#)). BMD_0658 is a slightly smaller protein, comprising just 485 aa and has a computed

molecular weight of 54.6 kDa. It shares a sequence identity of 32% and sequence similarity of 54% to CrtN of *S. aureus*. The aforementioned FAD binding motif is also present in the N-terminus of BMD_0658. The open reading frame of the third putative diapophytoene desaturase BMD_1977 contains a protein of 503 aa with an estimated molecular weight of 56.1 kDa. Although sharing a sequence identity of 30% and sequence homology of 56% to the crtN of *S. aureus*, the unique localization apart from the proposed C30 carotenoid cluster raised the question whether BMD_1977 could possibly serve a different function. Comparison to the amino acid sequence of *S. aureus* diaponeurosporene oxidase (CrtP) revealed that BMD_1977 shares significantly higher sequence identity of 45% as well as sequence similarity of 63% to CrtP. Hence, these findings emphasize the hypothesis that the open reading frame of BMD_1977 rather encodes a diaponeurosporene oxidase than a diapophytoene desaturase.

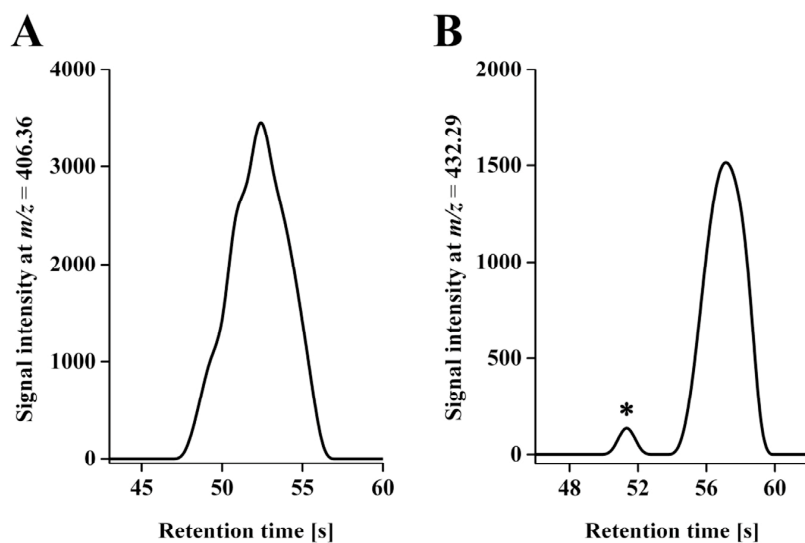
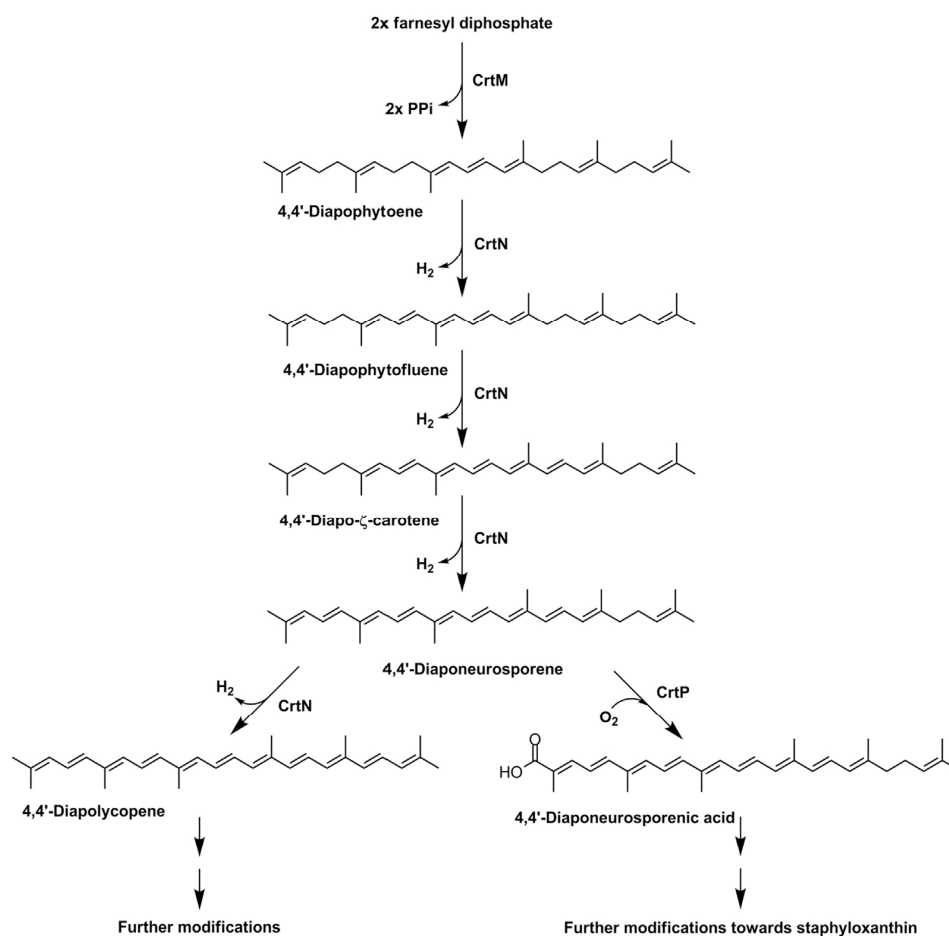


Fig. 3. Representative LC-MS chromatograms for the identified carotenoid metabolites in *B. megaterium* MS941. **A:** Measurement in positive ion mode revealed the presence of the C30 carotenoid 4,4'-diapophytofluene with m/z of 406.36. **B:** Measurement in negative ion mode confirmed the presence of the oxygenated C30 carotenoid derivative 4,4'-diaponeurosporenic acid with m/z of 432.29. * Not characterized metabolite.



Scheme 1. Biosynthetic pathway for 4,4'-diapocarotenoid formation. A diapophytoene synthase (CrtM) catalyzes the initial condensation of two molecules of farnesyl diphosphate to form the C30 carotenoid precursor 4,4'-diapophytoene. 4,4'-diapophytoene subsequently undergoes several stepwise desaturation reactions catalyzed by a single or multiple diapophytoene desaturases (CrtN) yielding the desaturation product 4,4'-diaponeurosporene. 4,4'-diaponeurosporene can be either desaturated to 4,4'-diapolycopene or oxidized by a mixed carotenoid oxygenase (CrtP) to form the staphyloxanthin precursor 4,4'-diaponeurosporenic acid.

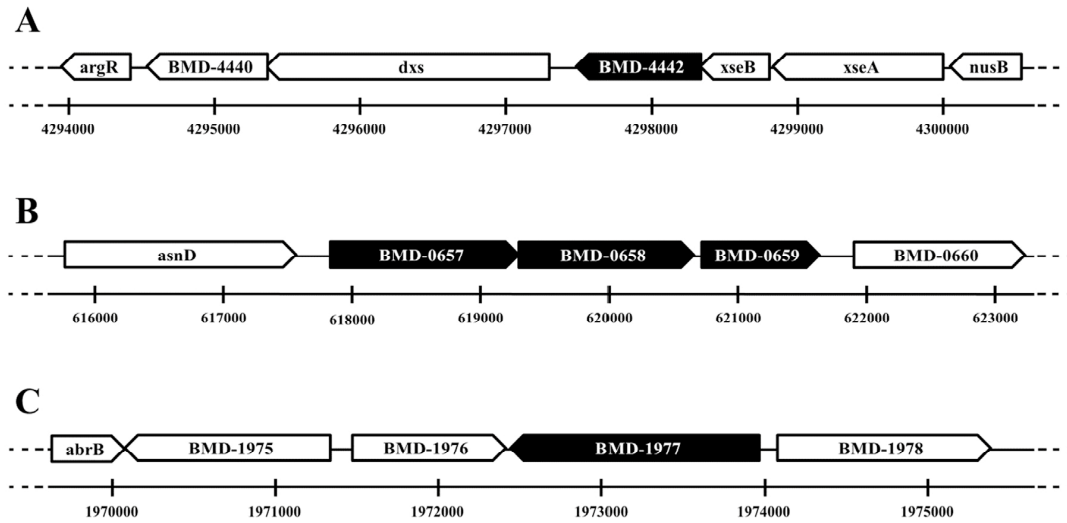


Fig. 4. Genomic context of the putative carotenoid biosynthetic ORFs in *B. megaterium* MS941. **A:** The ORF of the putative farnesyl diphosphate synthase BMD_4442 is located upstream of the 1-deoxy-D-xylulose-5-phosphate synthase gene (*dxs*) which catalyzes the initial step of the mevalonate independent isoprenoid biosynthesis. **B:** The putative diaphophytoene desaturases ORFs (BMD_0657 and BMD_0658) are organized as cluster together with a putative diaphophytoene synthase gene (BMD_0659). **C:** A third putative diaphophytoene desaturase (BMD_1977) seems to be separated from any other isoprenoid/carotenoid associated pathway. The annotation and location of the genes are shown according to MegaBac v9 database.

3.4. Formation of 4,4'-diaphophytoene is catalyzed by BMD_0659, but the stepwise desaturation to 4,4'-diaponeurosporene requires the combined activity of BMD_0657 and BMD_0658

In order to characterize the individual activities of the ORFs encoded by the putative *B. megaterium* C30 carotenoid cluster, the corresponding genes were assembled as synthetic operons into expression vector pSMF.A in different combinations (Fig. 5).

The vector pSMF.A contained the ORF of the newly identified BmFDS to provide adequate amounts of FDP for C30 carotenoid biosynthesis. Resulting constructs were expressed in *E. coli* TOP10 cells and screened for C30 carotenoid biosynthesis by RP-HPLC (Fig. 6).

Combined expression of BmFDS together with the putative diaphophytoene synthase BMD_0659 on the expression vector pSMF.AM lead to the accumulation of 4,4'-diaphophytoene with a retention time of 11.2 min and a characteristic absorption maximum of 287 nm. This

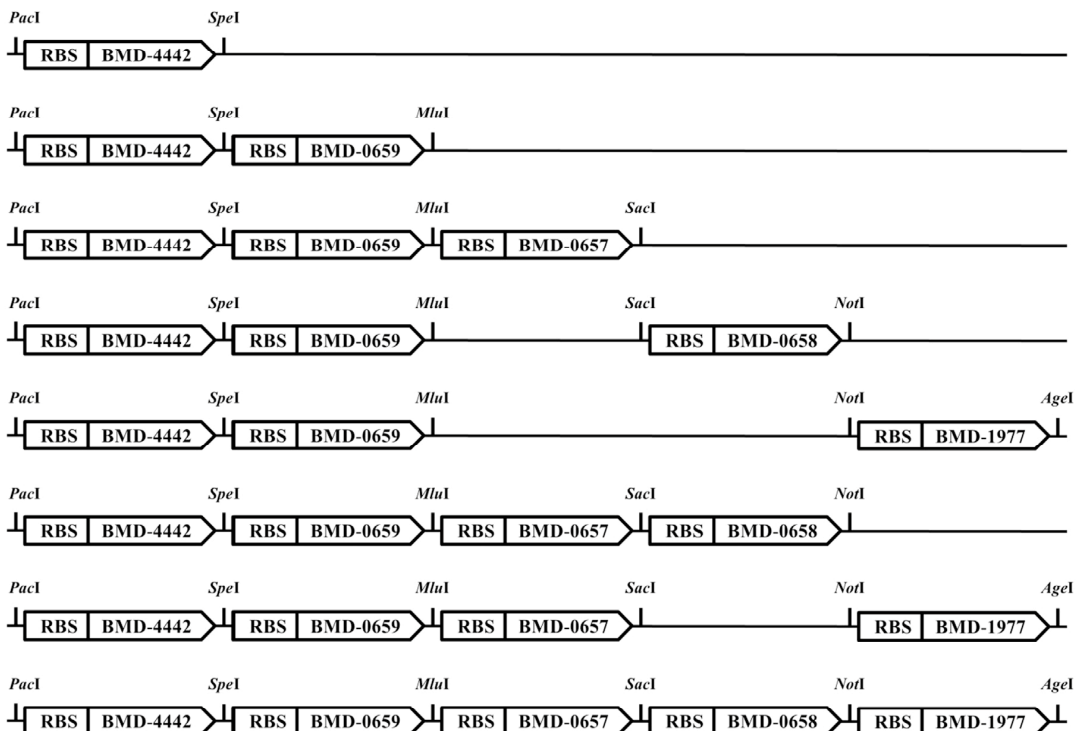


Fig. 5. Synthetic operons constructed for the characterization of the putative carotenoid operon from *B. megaterium* MS941.

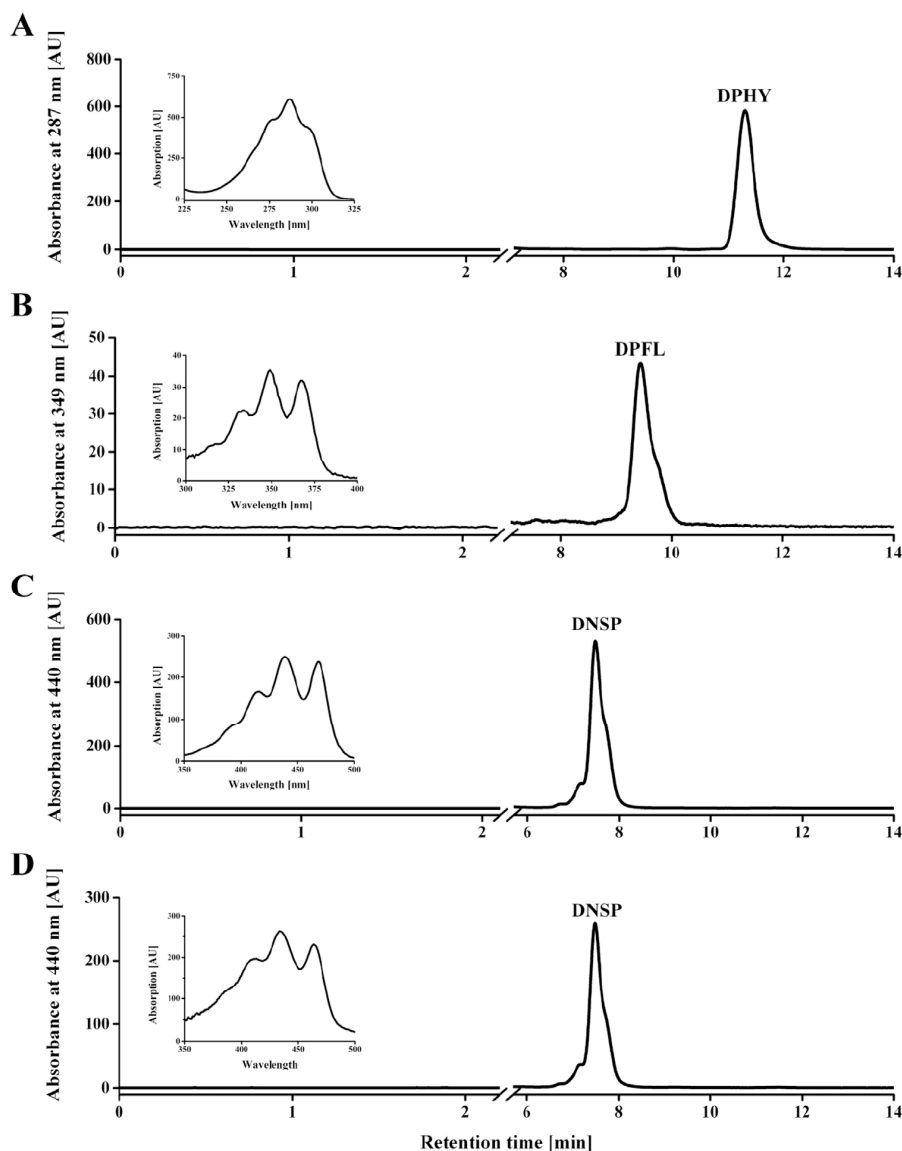


Fig. 6. RP-HPLC profiles of carotenoid extracts recorded at the indicated wavelength. Panels show representative extracts prepared from *E. coli* cells that were transformed with the vectors **A:** pSMF.AM, **B:** pSMF.AMN₁, **C:** pSMF.AMN₁₂, **D:** pSMF.AMN₁₂₃. The name of each carotenoid is indicated above the corresponding peak. DPHY: 4,4'-diapophytoene; DPFL: 4,4'-diapophytofluene; DNSP: 4,4'-diaponeurosporene. Insets show the absorption spectra of the corresponding carotenoid metabolites and were recorded with a photodiode array (PDA) detector.

result confirmed the hypothesis of BMD_0659 being a diapophytoene synthase in the carotenoid cluster. Since BMD_0659 possesses the same activity as the diapophytoene synthase (CrtM) of *S. aureus*, the corresponding enzyme will be referred to as *Bacillus megaterium* diapophytoene synthase (BmCrtM) (Wieland et al., 1994).

Upon additional expression of the first putative diapophytoene desaturase BMD_0657 on the vector pSMF.AMN₁ a new but low abundant metabolite was detected with a shorter retention time of 9.6 min and the characteristic spectral properties of carotenoids, comprising multiple absorption maxima at 333 nm, 349 nm and 368 nm. Comparison to a focused library of absorption spectra for different C30 carotenoids generated by Takaichi showed that the recorded absorption spectrum is identical to that of 4,4'-diapophytofluene (Takaichi, 2000).

BMD_0657 will be hereinafter referred to as *Bacillus megaterium* diapophytoene desaturase (BmCrtN₁). The activity of the heterologously expressed BmCrtN₁ fundamentally differs from the

diapophytoene desaturase activity reported in *S. aureus* and other pigmented *Bacillus* species, where a single diapophytoene desaturase is responsible for the three and four step desaturation reaction of diapophytoene to yield both, 4,4'-diaponeurosporene as well as 4,4'-diapophytofluene, unselectively (Kim and Lee, 2012; Pelz et al., 2005; Steiger et al., 2015). BmCrtN₁, however, catalyzes the single step desaturation reaction of 4,4'-diapophytoene to yield predominantly 4,4'-diapophytofluene.

When the second putative diapophytoene desaturase BMD_0658 was expressed in a synthetic operon with BmFDS and the diapophytoene synthase BMD_0659 on vector pSMF.AMN₂, no desaturation product of 4,4'-diapophytoene was identified (data not shown). The putative diapophytoene desaturase BMD_0658 was obviously not able to accept 4,4'-diapophytoene as substrate. Only the combined expression of both putative diapophytoene desaturases, BMD_0657 and BMD_0658, on the expression vector pSMF.AMN₁₂ yielded a new

Table 4
4,4'-diaponeurosporene yields with *Escherichia coli* based microbial cell factory.

Medium	Specific 4,4'-diaponeurosporene yield [$\mu\text{g/g}$ wet cell weight]	Total 4,4'-diaponeurosporene yield [$\mu\text{g/L}$ culture]
LB (Miller)	181.1 \pm 10.6 (14.6)	1293 \pm 76
M9CA	110.2 \pm 16.2 (8.9)	956 \pm 141
2xYT	134.2 \pm 6.6 (10.8)	1894 \pm 93
TB	146.1 \pm 5.1 (11.8)	4032 \pm 137
Strain	Specific 4,4'-diaponeurosporene yield in TB medium [$\mu\text{g/g}$ wet cell weight]	Total 4,4'-diaponeurosporene yield in TB medium [$\mu\text{g/L}$ culture]
BL21(DE3)	11.3 \pm 2.1	423 \pm 78
C43(DE3)	30.5 \pm 1.6	1084 \pm 58
JM109	75.2 \pm 3.1	2801 \pm 115
TOP10	146.1 \pm 5.1	4032 \pm 137

In brackets is the fold increase in C30 carotenoid yields compared to Chae et al. (2010).

carotenoid metabolite apart from 4,4'-diapophytofluene with a retention time of 7.6 min and multiple absorption maxima at 416 nm, 439 nm and 468 nm. According to the studies of Takaichi these absorption maxima are characteristic for the C30 carotenoid 4,4'-diaponeurosporene (Takaichi, 2000). Consequently, coexpression of BMD_0657 is crucial to provide adequate amounts of 4,4'-diapophytofluene which seems to be the preferred substrate for the desaturation reaction of BMD_0658. Thus, BMD_0658 should be assigned as diapophytofluene desaturase, catalyzing the two step desaturation reaction of 4,4'-diapophytofluene to 4,4'-diaponeurosporene. BMD_0658 will, therefore, be referred to as *Bacillus megaterium* diapophytofluene desaturase (BmCrtN₂). To the best of our knowledge, this uncommon diapophytofluene desaturation activity, in terms of 4,4'-diapophytofluene being the preferred substrate as well as 4,4'-diaponeurosporene being the exclusive desaturation product, is unique among C30 carotenoid accumulating organisms and has not been described before.

In a following experiment, the expression vector pSMF.AMN₁₂₃ was used to study the proposed diaponeurosporene oxygenase activity of BMD_1977. Intriguingly, only the substrate for the oxygenation reaction, 4,4'-diaponeurosporene, but no oxygenated derivative were detected. Thus, despite high sequence identity to the diaponeurosporene oxygenase (crtP) of *S. aureus*, BMD_1977 does evidently not possess an oxygenase activity but might rather be a third putative diapophytoene desaturase. To test this hypothesis, BMD_1977 was coexpressed with BmFDS and diapophytoene synthase BMD_0659 on the vector pSMF.AMN₃. However, no desaturation product of 4,4'-diapophytoene was identified (data not shown). The assumed desaturation activity of BMD_1977 was also not detectable when 4,4'-diapophytofluene was provided as substrate by coexpression of the diapophytoene desaturase BMD_0657 on the vector pSMF.AMN₁₃ (data not shown). Apparently neither 4,4'-diapophytoene nor 4,4'-diapophytofluene can serve as substrate of BMD_1977. The role of BMD_1977 in the C30 carotenoid biosynthesis of *B. megaterium* thus remains unclear.

Nevertheless, there are reasonable indications that other carotenoid oxygenases, apart from BMD_1977 could be involved in the conversion of 4,4'-diaponeurosporene to 4,4'-diaponeurosporenic acid. Several studies have shown that also cytochromes P450 monooxygenases are able to act as carotenoid hydroxylase in plants, fungi and even bacteria (Fiore et al., 2006; Álvarez et al., 2006; Blasco et al., 2004). Over the last decades different cytochromes P450 from *B. megaterium* were identified and their corresponding substrate ranges were investigated intensively (Ho and Fulco, 1976; Brill et al., 2014; Abdumughni et al., 2017; Józwick et al., 2016). In the course of these studies, mainly steroidal as well as terpenoid compounds were identified as substrates (Kiss et al., 2015; Putkaradze et al., 2017). Particularly CYP102A1 (BM3)

was, furthermore, demonstrated to show activity towards the hydrocarbon skeleton of fatty acids with varying chain lengths, which is similar to the structure of acyclic carotenoids (Whitehouse et al., 2012).

3.5. The engineered carotenoid operon can be used for the efficient biosynthesis of the pharmaceutically relevant C30 carotenoid 4,4'-diaponeurosporene

The functional characterization of the engineered carotenoid clusters from *B. megaterium* revealed unique features of diapophytoene desaturation that allow for a selective production of the pharmaceutically relevant C30 carotenoid 4,4'-diaponeurosporene. For this reason, we aimed to establish an *E. coli* based microbial cell factory and assessed its potential for the biotechnological production of 4,4'-diaponeurosporene. We chose *E. coli* due to the fact that side product formation has to be expected in *B. megaterium* MS941 with regard to a still not identified diaponeurosporene oxidase (CrtP). For this purpose, *E. coli* TOP10 cells were transformed with the pSMF2.1 based vector, expressing BmFDS, BmCrtM, BmCrtN₁ and BmCrtN₂. 4,4'-diaponeurosporene levels were determined after 20 h of incubation in shaking flasks.

As shown in Table 4, initial 4,4'-diaponeurosporene production with the *E. coli* TOP10 based microbial cell factory in LB medium resulted in a specific amount of 4,4'-diaponeurosporene of 181.1 $\mu\text{g g}^{-1}$ wet cell weight. This is a 14.6 fold increase in C30 carotenoid production compared to representative studies on 4,4'-diapolyycopene formation in *E. coli* TOP10 where only 12.4 $\mu\text{g g}^{-1}$ wet cell weight were achieved (Chae et al., 2010). Subsequent to these promising results, we tested the influence of different culture media on specific 4,4'-diaponeurosporene yields to further improve C30 carotenoid yields. But all other tested media had a negative impact on specific 4,4'-diaponeurosporene yields, with a minimum amount of 110.2 $\mu\text{g g}^{-1}$ wet cell weight, when *E. coli* cells were grown in modified M9CA medium. However, higher cell densities in complex media positively affected the total 4,4'-diaponeurosporene yields with 1894 $\mu\text{g L}^{-1}$ in 2xYT medium and 4032 $\mu\text{g L}^{-1}$ in TB medium, respectively (Table 4).

Chae and coworkers furthermore showed that C30 carotenoid accumulation levels in metabolically engineered *E. coli* cells are strongly dependent on the *E. coli* strain itself. They were able to increase the specific C30 carotenoid yield to 35.3 $\mu\text{g g}^{-1}$ and 49.4 $\mu\text{g g}^{-1}$ wet cell weight using *E. coli* strains SURE and JM109, respectively. For this reason, several alternative *E. coli* strains, including BL21 (DE3), C43 (DE3) as well as JM109, were transformed with the expression vector pSMF.AMN₁₂ and analyzed for 4,4'-diaponeurosporene yields in TB medium. As shown in Table 4, all tested *E. coli* strains showed significantly lower specific as well as total 4,4'-diaponeurosporene yields compared to the corresponding carotenoid accumulation levels in *E. coli* strain TOP10. Particularly *E. coli* strains BL21 (DE3) and C43 (DE3) seemed to be inappropriate microbial cell factories for the high yield production of 4,4'-diaponeurosporene. However, the C30 carotenoid yields with the engineered carotenoid biosynthesis cluster from *B. megaterium* were considerably higher when expressed in *E. coli* TOP10 cells, exceeding the carotenoid yields of the optimized system from Chae and coworkers by the factor 3.6–5.1.

Kim et al. use a similar constitutive system for the heterologous expression the C30 carotenoid cluster from *S. aureus* in *E. coli* K12 derivatives (Kim et al., 2013). In the corresponding study, total C30 carotenoid yields were reported to vary from approximately 400 $\mu\text{g L}^{-1}$ to 700 $\mu\text{g L}^{-1}$ culture without regard to the product ratios of 4,4'-diaponeurosporene and 4,4'-diapolyycopene. Compared to the study of Kim and coworkers, the total yield of 4,4'-diaponeurosporene was improved at least by the factor 5.7 by using the engineered carotenoid biosynthesis cluster from *B. megaterium* in *E. coli* TOP10.

A more recent study by Furubayashi and coworkers describes the establishment of an arabinose inducible system for the heterologous expression of the C30 carotenoid cluster from *S. aureus* in *E. coli* strain

XL10-Gold (Furubayashi et al., 2014). 4,4'-diaponeurosporene yields of $48.8 \mu\text{g g}^{-1}$ to $112 \mu\text{g g}^{-1}$ dry cell weight were reported, however, with considerable side product formation of 4,4'-diapolycopene. When converted to 4,4'-diaponeurosporene yields per gram wet cell weight, the corresponding carotenoid yields are supposed to be approximately between $12 \mu\text{g g}^{-1}$ and $28 \mu\text{g g}^{-1}$ wet cell weight (Bratbak and Dundas, 1984). Consequently, the highest specific 4,4'-diaponeurosporene yields of this work ($181 \mu\text{g g}^{-1}$), was generated by using the constitutive promoter system and exceeds the best 4,4'-diaponeurosporene yield of the arabinose inducible expression system of Furubayashi and colleagues by the factor 6.4.

These results further emphasize that the engineered whole-cell system, applying the engineered carotenoid biosynthesis operon from *B. megaterium*, is superior to other reported *E. coli* based systems with regard to product yield as well as product selectivity.

To the best of our knowledge our system provides an industrially promising microbial cell factory, which can be employed for high yield C30 carotenoid production.

4. Conclusion

Initial studies with 4,4'-diaponeurosporene convincingly demonstrate its stimulating effects on the immune system of mice (Liu et al., 2016). Administration of 4,4'-diaponeurosporene has been shown to effectively activate dendritic cells, thereby not only enhancing the degree of the immune response but providing resistance to *Salmonella* infections (Jing et al., 2017; Liu et al., 2017). However, efficient biocatalysts for the biosynthesis of 4,4'-diaponeurosporene are not available because natural producers, like *Staphylococcus aureus*, are pathogenic or not able to produce 4,4'-diaponeurosporene selectively. In these bacteria, a major amount of 4,4'-diaponeurosporene is converted by desaturation, oxygenation and even glycosylation to additional products like 4,4'-diapolycopene, 4,4'-diapolycopenal and staphyloxanthin, respectively (Wieland et al., 1994; Steiger et al., 2015; Perez-Fons et al., 2011).

We were able to elucidate a novel biosynthetic pathway for C30 carotenoid production, thereby revealing unique features concerning the desaturation reaction of 4,4'-diapophytoene to 4,4'-diaponeurosporene. Furthermore, we successfully expressed the engineered carotenoid operons from *B. megaterium* to establish *E. coli* as microbial cell factory for a high-yield production of 4,4'-diaponeurosporene. Our system was shown to be superior to previous *E. coli* based systems for the production of C30 carotenoids in terms of product yield and product selectivity. Simple shaking flask experiments resulted in a significant increase in C30 carotenoid accumulation levels compared to studies (Chae et al., 2010; Furubayashi et al., 2014; Kim et al., 2013). Most importantly, we were able to produce 4,4'-diaponeurosporene as exclusive carotenoid without any side products, which will facilitate downstream processing in future large scale fermentation processes.

Acknowledgements

We would like to thank Antje Eiden-Plach and Birgit Heider-Lips for their assistance with the bioreactor.

Funding

This work was supported by the grant AZ 32269/01 of the Deutsche Bundesstiftung Umwelt (DBU).

Competing interests

The authors declare no competing interest.

Author's contribution

P.H. performed all experiments, analyzed and interpreted the data. M.M. participated in the design of the experimental setup, the interpretation of the results and assisted in drafting the manuscript. P.H. wrote the manuscript. S.T. carried out metabolome analyses. F.H. and R.B. supervised the experiments, participated in the interpretation and discussion of the data as well as the writing of the manuscript.

References

- Abdulmughni, A., Józwiak, I.K., Putkaradze, N., Brill, E., Zapp, J., Thunnissen, A.-M.W.H., Hannemann, F., Bernhardt, R., 2017. Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D3 hydroxylase. *J. Biotechnol.* 243, 38–47.
- Alarcon, M., Cori, O., Rojas, C., 1992. Hydrolyses of terpenoid diphosphates. Effects of azide ion on products of hydrolysis. *J. Phys. Org. Chem.* 5, 83–92.
- Álvarez, V., Rodríguez-Sáiz, M., de la Fuente, J.L., Gudiña, E.J., Godio, R.P., Martín, J.F., Barredo, J.L., 2006. The crtS gene of *Xanthophyllomyces dendrorhous* encodes a novel cytochrome-P450 hydroxylase involved in the conversion of β -carotene into astaxanthin and other xanthophylls. *Fungal Genet. Biol.* 43, 261–272.
- Blasco, F., Kauffmann, I., Schmid, R.D., 2004. CYP175A1 from *Thermus thermophilus* HB27, the first beta-carotene hydroxylase of the P450 superfamily. *Appl. Microbiol. Biotechnol.* 64, 671–674.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- β -boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl. Microbiol. Biotechnol.* 93, 1135–1146.
- Bratbak, G., Dundas, I., 1984. Bacterial dry matter content and biomass estimations. *Appl. Environ. Microbiol.* 48, 755–757.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2014. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1701–1717.
- Britton, G., 1995. Structure and properties of carotenoids in relation to function. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 9, 1551–1558.
- Chae, H.S., Kim, K.-H., Kim, S.C., Lee, P.C., 2010. Strain-dependent carotenoid productions in metabolically engineered *Escherichia coli*. *Appl. Biochem. Biotechnol.* 162, 2333–2344.
- Chew, B.P., Park, J.S., 2004. Carotenoid action on the immune response. *J. Nutr.* 134, 257S–261S.
- Cramer, F., Rittersdorf, W., 1967. Die Hydrolyse von Phosphaten und Pyrophosphaten einiger monoterpenalkohole. *Tetrahedron* 23, 3015–3022.
- Duc, L.H., Fraser, P.D., Tam, N.K.M., Cutting, S.M., 2006. Carotenoids present in halo-tolerant *Bacillus* spore formers. *FEMS Microbiol. Lett.* 255, 215–224.
- Dym, O., Eisenberg, D., 2001. Sequence-structure analysis of FAD-containing proteins. *Protein Sci. Publ. Protein Soc.* 10, 1712–1728.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., et al., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J. Bacteriol.* 193, 4199–4213.
- Fiore, A., Dall'Osto, L., Fraser, P.D., Bassi, R., Giuliano, G., 2006. Elucidation of the β -carotene hydroxylation pathway in *Arabidopsis thaliana*. *FEBS Lett.* 580, 4718–4722.
- Frank, H.A., Cogdell, R.J., 1996. Carotenoids in photosynthesis. *Photochem. Photobiol.* 63, 257–264.
- Furubayashi, M., Li, L., Katabami, A., Saito, K., Umeno, D., 2014. Construction of carotenoid biosynthetic pathways using squalene synthase. *FEBS Lett.* 588, 436–442.
- Gao, P., Davies, J., Kao, R.Y.T., 2017. Dehydroxysqualene Desaturase as a novel target for anti-virulence therapy against *Staphylococcus aureus*. *MBio* 8 (e01224-17).
- Garrido-Fernández, J., Maldonado-Barragán, A., Caballero-Guerrero, B., Hornero-Méndez, D., Ruiz-Barba, J.L., 2010. Carotenoid production in *Lactobacillus plantarum*. *Int. J. Food Microbiol.* 140, 34–39.
- George-Nascimento, C., Pont-Lezica, R., Cori, O., 1971. Non enzymic formation of nerolidol from farnesyl pyrophosphate in the presence of bivalent cations. *Biochem. Biophys. Res. Commun.* 45, 119–124.
- Gerster, H., 1993. Anticarcinogenic effect of common carotenoids. *Int. J. Vitam. Nutr. Res. Int. Z. Vitam. - Ernähr. J. Int. Vitaminol. Nutr.* 63, 93–121.
- Gruszecki, W.I., Strzałka, K., 2005. Carotenoids as modulators of lipid membrane physical properties. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* 1740, 108–115.
- Ho, P.P., Fulco, A.J., 1976. Involvement of a single hydroxylase species in the hydroxylation of palmitate at the omega-1, omega-2 and omega-3 positions by a preparation from *Bacillus megaterium*. *Biochim. Biophys. Acta* 431, 249–256.
- Jing, Y., Liu, H., Xu, W., Yang, Q., 2017. Amelioration of the DSS-induced colitis in mice by pretreatment with 4,4'-diaponeurosporene-producing *Bacillus subtilis*. *Exp. Ther. Med.* 14, 6069–6073.
- Józwiak, I.K., Kiss, F.M., Gricman, L., Abdulmughni, A., Brill, E., Zapp, J., Pleiss, J., Bernhardt, R., Thunnissen, A.-M.W.H., 2016. Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J* 283, 4128–4148.
- Khaneja, R., Perez-Fons, L., Fakhry, S., Baccigalupi, L., Steiger, S., To, E., Sandmann, G., Dong, T.C., Ricca, E., Fraser, P.D., et al., 2010. Carotenoids found in *Bacillus*. *J. Appl. Microbiol.* 108, 1889–1902.
- Kim, S.H., Lee, P.C., 2012. Functional expression and extension of staphylococcal

- staphyloxanthin biosynthetic pathway in *Escherichia coli*. *J. Biol. Chem.* 287, 21575–21583.
- Kim, J.R., Kim, S.H., Lee, S.Y., Lee, P.C., 2013. Construction of homologous and heterologous synthetic sucrose utilizing modules and their application for carotenoid production in recombinant *Escherichia coli*. *Bioresour. Technol.* 130, 288–295.
- Kirti, K., Amita, S., Priti, S., Mukesh Kumar, A., Jyoti, S., 2014. Colorful World of Microbes: carotenoids and Their Applications. *Adv. Biol.* 2014, 1–13.
- Kiss, F.M., Khatri, Y., Zapp, J., Bernhardt, R., 2015. Identification of new substrates for the CYP106A1-mediated 11-oxidation and investigation of the reaction mechanism. *FEBS Lett.* 589, 2320–2326.
- Köcher, S., Breitenbach, J., Müller, V., Sandmann, G., 2009. Structure, function and biosynthesis of carotenoids in the moderately halophilic bacterium *Halobacillus halophilus*. *Arch. Microbiol.* 191, 95–104.
- Liu, C.-I., Liu, G.Y., Song, Y., Yin, F., Hensler, M.E., Jeng, W.-Y., Nizet, V., Wang, A.H.-J., Oldfield, E., 2008. A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. *Science* 319, 1391–1394.
- Liu, H., Xu, W., Chang, X., Qin, T., Yin, Y., Yang, Q., 2016. 4,4'-diaponeurosporene, a C30 carotenoid, effectively activates dendritic cells via CD36 and NF- κ B signaling in a ROS independent manner. *Oncotarget* 7, 40978–40991.
- Liu, H., Xu, W., Yu, Q., Yang, Q., 2017. 4,4'-diaponeurosporene-producing *Bacillus subtilis* increased mouse resistance against salmonella typhimurium infection in a CD36-dependent manner. *Front. Immunol.* 8, 483.
- Marshall, J.H., Wilmoth, G.J., 1981. Proposed pathway of triterpenoid carotenoid biosynthesis in *Staphylococcus aureus*: evidence from a study of mutants. *J. Bacteriol.* 147, 914–919.
- McDevitt, T.M., Tchao, R., Harrison, E.H., Morel, D.W., 2005. Carotenoids normally present in serum inhibit proliferation and induce differentiation of a human monocyte/macrophage cell line (U937). *J. Nutr.* 135, 160–164.
- Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K., Harashima, K., 1990. Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. *J. Bacteriol.* 172, 6704–6712.
- Mitchell, C., Iyer, S., Skomurski, J.F., Vary, J.C., 1986. Red pigment in *Bacillus megaterium* spores. *Appl. Environ. Microbiol.* 52, 64–67.
- Miyake, Y., Fukushima, W., Tanaka, K., Sasaki, S., Kiyohara, C., Tsuboi, Y., Yamada, T., Oeda, T., Miki, T., Kawamura, N., et al., 2011. Dietary intake of antioxidant vitamins and risk of Parkinson's disease: a case-control study in Japan. *Eur. J. Neurol.* 18, 106–113.
- Moeller, R., Horneck, G., Facius, R., Stackebrandt, E., 2005. Role of pigmentation in protecting *Bacillus sp. endospores* against environmental UV radiation. *FEMS Microbiol. Ecol.* 51, 231–236.
- Nisar, N., Li, L., Lu, S., Khin, N.C., Pogson, B.J., 2015. Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82.
- Norris, S.R., Barrette, T.R., DellaPenna, D., 1995. Genetic dissection of carotenoid synthesis in *Arabidopsis* defines plastoquinone as an essential component of phytoene desaturation. *Plant Cell* 7, 2139–2149.
- Obulesu, M., Dowlathabad, M.R., Bramhachari, P.V., 2011. Carotenoids and Alzheimer's disease: an insight into therapeutic role of retinoids in animal models. *Neurochem. Int.* 59, 535–541.
- Pelz, A., Wieland, K.-P., Putzbach, K., Hentschel, P., Albert, K., Götz, F., 2005. Structure and biosynthesis of staphyloxanthin from *Staphylococcus aureus*. *J. Biol. Chem.* 280, 32493–32498.
- Perez-Fons, L., Steiger, S., Khaneja, R., Bramley, P.M., Cutting, S.M., Sandmann, G., Fraser, P.D., 2011. Identification and the developmental formation of carotenoid pigments in the yellow/orange *Bacillus* spore-formers. *Biochim. Biophys. Acta* 1811, 177–185.
- Putkaradze, N., Litzenburger, M., Abdulmughni, A., Milhim, M., Brill, E., Hannemann, F., Bernhardt, R., 2017. CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*. *Appl. Microbiol. Biotechnol.* 101, 8379–8393.
- Racine, F.M., Vary, J.C., 1980. Isolation and properties of membranes from *Bacillus megaterium* spores. *J. Bacteriol.* 143, 1208–1214.
- Rao, A.V., Agarwal, S., 2000. Role of antioxidant lycopene in cancer and heart disease. *J. Am. Coll. Nutr.* 19, 563–569.
- Rodríguez-Sáiz, M., de la Fuente, J.L., Barredo, J.L., 2010. *Xanthophyllomyces dendrorhous* for the industrial production of astaxanthin. *Appl. Microbiol. Biotechnol.* 88, 645–658.
- Schmidt, A., Wachtler, B., Temp, U., Kreckling, T., Seguin, A., Gershenzon, J., 2010. A bifunctional geranyl and geranylgeranyl diphosphate synthase is involved in terpene oleoresin formation in *Picea abies*. *PLANT Physiol.* 152, 639–655.
- Stammen, S., Müller, B.K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E., Jahn, D., 2010. High-yield intra- and extracellular protein production using *Bacillus megaterium*. *Appl. Environ. Microbiol.* 76, 4037–4046.
- Steiger, S., Perez-Fons, L., Fraser, P.D., Sandmann, G., 2012. Biosynthesis of a novel C30 carotenoid in *Bacillus firmus* isolates. *J. Appl. Microbiol.* 113, 888–895.
- Steiger, S., Perez-Fons, L., Cutting, S.M., Fraser, P.D., Sandmann, G., 2015. Annotation and functional assignment of the genes for the C30 carotenoid pathways from the genomes of two bacteria: *Bacillus indicus* and *Bacillus firmus*. *Microbiol. Read. Engl.* 161, 194–202.
- Szkołpińska, A., Plochocka, D., 2005. Farnesyl diphosphate synthase; regulation of product specificity. *Acta Biochim. Pol.* 52, 45–55.
- Takaichi, S., Inoue, K., Akaike, M., Kobayashi, M., Oh-oka, H., Madigan, M.T., 1997. The major carotenoid in all known species of heliobacteria is the C30 carotenoid 4,4'-diaponeurosporene, not neurosporene. *Arch. Microbiol.* 168, 277–281.
- Takaichi, S., 2000. Characterization of carotenes in a combination of a C(18) HPLC column with isocratic elution and absorption spectra with a photodiode-array detector. *Photosynth. Res.* 65, 93–99.
- Umeno, D., Arnold, F.H., 2003. A C35 carotenoid biosynthetic pathway. *Appl. Environ. Microbiol.* 69, 3573–3579.
- Wackenroder, H., 1831. Ueber das oleum radices Dauci aetherum, das Carotin, den Carotenzucker und den officinellen succus Dauci; so wie auch über das Mannit, welches in dem Möhrensafte durch eine besondere art der Gahrung gebildet wird. *Geigers Mag. Pharm.* 33, 144–172.
- Whitehouse, C.J.C., Bell, S.G., Wong, L.-L., 2012. P450(BM3) (CYP102A1): connecting the dots. *Chem. Soc. Rev.* 41, 1218–1260.
- Wieland, B., Feil, C., Gloria-Maercker, E., Thumm, G., Lechner, M., Bravo, J.M., Poralla, K., Götz, F., 1994. Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. *J. Bacteriol.* 176, 7719–7726.
- Wittchen, K.D., Meinhardt, F., 1995. Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl. Microbiol. Biotechnol.* 42, 871–877.
- Yabuzaki, J., 2017. Carotenoids database: structures, chemical fingerprints and distribution among organisms. *Database J. Biol. Databases Curation* 2017.
- Yoshida, K., Ueda, S., Maeda, I., 2009. Carotenoid production in *Bacillus subtilis* achieved by metabolic engineering. *Biotechnol. Lett.* 31, 1789–1793.
- Zile, M.H., Cullum, M.E., 1983. The function of vitamin A: current concepts. *Proc. Soc. Exp. Biol. Med. Soc. Exp. Biol. Med. N.Y.* 172, 139–152. <<http://www.bccresearch.com/report/FOD025E.html>>.

2.6 (Milhim et al., 2016)

Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin.

Milhim, M., Putkaradze, N., Abdulmughni, A., Kern, F., **Hartz, P.** and Bernhardt, R.

Journal of Biotechnology. 2016 Nov; 240:68-75

DOI: 10.1016/j.jbiotec.2016.11.002

Reprinted with permission of Journal of Biotechnology. All rights reserved.



Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin



Mohammed Milhim, Natalia Putkaradze, Ammar Abdulmughni, Fredy Kern, Philip Hartz, Rita Bernhardt*

Institute of Biochemistry, Saarland University, 66123 Saarbrücken, Germany

ARTICLE INFO

Article history:

Received 2 September 2016
Received in revised form 28 October 2016
Accepted 1 November 2016
Available online 2 November 2016

Keywords:

Bacillus megaterium
CYP107 family
Bacterial cytochrome P450
Mevastatin
Pravastatin
BmCPR
Fdx2

ABSTRACT

In the current work, we describe the identification and characterization of the first plasmid-encoded P450 (CYP107DY1) from a *Bacillus* species. The recombinant CYP107DY1 exhibits characteristic P450 absolute and reduced CO-bound difference spectra. Reconstitution with different redox systems revealed the autologous one, consisting of BmCPR and Fdx2, as the most effective one. Screening of a library of 18 pharmaceutically relevant compounds displayed activity towards mevastatin to produce pravastatin. Pravastatin is an important therapeutic drug to treat hypercholesterolemia, which was described to be produced by oxyfunctionalization of mevastatin (compactin) by members of CYP105 family. The hydroxylation at C6 of mevastatin was also suggested by docking this compound into a computer model created for CYP107DY1. Moreover, in view of the biotechnological application, CYP107DY1 as well as its redox partners (BmCPR and Fdx2) were successfully utilized to establish an *E. coli* based whole-cell system for an efficient biotransformation of mevastatin. The in vitro and in vivo application of the CYP107DY1 also offers the possibility for the screening of more substrates, which could open up further biotechnological usage of this enzyme.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Selective oxyfunctionalization of nonactivated carbon-hydrogen bonds represents a challenge in the industrial field. In general, the use of chemical methods has many disadvantages such as hazardous conditions, cost-efficiency and the lack of chemo-, stereo-, and regioselectivity. Therefore, in the last years many efforts have been carried out in the search for selective and efficient enzymatic systems that are able to incorporate oxygen into nonactivated carbon-hydrogen bonds. Cytochromes P450 (P450s) [E.C.1.14.-] are heme-iron containing enzymes that catalyze the monooxygenation of various nonactivated hydrocarbons with high regio-, stereo- and enantioselectivity including the biosynthesis of hormones, signal molecules, defense-related chemicals and secondary metabolites in addition to their central role in the metabolism of endogenous (steroids and fatty acids) and exogenous (drugs and toxins) substances (Bernhardt, 2006;

Bernhardt and Urlacher, 2014). They are found in all kingdoms of life including mammals, plants, insects, fungi, archaea, and bacteria (except *E. coli*) as well as in viruses. In eukaryotes, they are mostly integral membrane-bound proteins, whereas prokaryotic P450 systems are more likely soluble and located in the cytoplasm. P450s rely for their activities on redox partners. Based on the composition of the redox partner involved in the transfer of electrons, the P450 systems can be categorized into different classes (Hannemann et al., 2007) of which the most researched ones are the classes I and II. Class I contains the bacterial and the eukaryotic mitochondrial P450s, which obtain electrons from NADPH using two proteins, a flavin adenine dinucleotide (FAD)-containing ferredoxin reductase and an iron-sulfur containing ferredoxin. Class II comprises the eukaryotic microsomal P450s, which obtain electrons from NADPH via a FAD and FMN-containing P450 reductase.

The numbers of the newly identified P450s increased drastically over the past few years (Nelson, 2009), but there is nevertheless a still growing demand to exploit novel P450s as a valuable biocatalyst in the industrial field.

Bacillus megaterium is a nonpathogenic, aerobic, Gram-positive rod-shaped bacterium. Due to its high protein production capacity, plasmid stability and the ability to take up a variety of hydrophobic

* Corresponding author at: Institute of Biochemistry, Campus B2.2, Saarland University, D-66123 Saarbrücken, Germany.

E-mail address: ritabern@mx.uni-saarland.de (R. Bernhardt).

<http://dx.doi.org/10.1016/j.jbiotec.2016.11.002>

0168-1656/© 2016 Elsevier B.V. All rights reserved.

substrates, *B. megaterium* gained throughout the last decades a lot of interest in the industrial field for the production of biotechnologically relevant substances (Bunk et al., 2010). The publication of the complete genome sequence of the *B. megaterium* strains QM B1551 and DSM319 in 2011 (Eppinger et al., 2011) enabled the identification of new proteins. Among them are cytochromes P450 and a NADPH dependent diflavin reductase, which have been shown to be very important for biotechnological and pharmaceutical applications (Brill et al., 2013; Milhim et al., 2016). *B. megaterium* encodes for several P450s. The self-sufficient CYP102A1 (also known as BM3) is the most investigated bacterial P450 so far, which has been used and redesigned to catalyze the oxidation of a variety of biotechnologically interesting substances (Whitehouse et al., 2012). In addition, the biotechnologically valuable CYP106 family, CYP106A1 from *B. megaterium* strain DSM319 (Brill et al., 2013) and CYP106A2 from *B. megaterium* strain ATCC 13368 (Berg et al., 1976, 1979), was characterized to be associated with the biotransformation of a diverse array of substrates such as steroids and terpenoid substances (Brill et al., 2013; Schmitz et al., 2012). Furthermore, CYP109E1 was recently identified from *B. megaterium* strain DSM319 as steroid hydroxylase (Jóźwik et al., 2016).

In this study, we report the identification and characterization of a new plasmid-encoded P450 from the *B. megaterium* QM B1551. The bioinformatic analysis of the new P450 showed that it belongs to the CYP107 family. It was successfully cloned and expressed in *E. coli*. Screening of a potential substrates showed that CYP107DY1 possesses hydroxylation activity towards mevastatin.

2. Materials and methods

2.1. Strains, expression vectors, enzymes, and chemicals

E. coli TOP10 from Invitrogen (Karlsruhe, Germany) was used for cloning experiments. *E. coli* C43 (DE3) and the expression vector pET17b, both from Novagen (Darmstadt, Germany), were used for recombinant gene expression. Substrates were obtained from TCI (Eschborn, Germany). Pravastatin lactone was from Santa Cruz Biotechnology (Heidelberg, Germany). All other chemicals were purchased from Sigma–Aldrich (Schnellendorf, Germany).

2.2. Cloning of the gene encoding CYP107DY1

For protein purification purposes, the DNA fragment encoding the full length CYP107DY1 (Supplementary Fig. S1) was synthesized (Genart, Regensburg, Germany) and cloned into the expression vector pET17b with the *Nde*I/*Kpn*I restriction sites. For purification with IMAC, the 3' end of the gene was extended with a sequence coding for six histidines. Plasmid was verified by sequencing.

2.3. Heterologous gene expression and purification of CYP107DY1, reductases and ferredoxins

For heterologous gene expression, *E. coli* C43 (DE3) cells were co-transformed with the expression vector pET17b congaing the sequence for CYP107DY1 and the chaperone GroEL/GroES-encoding plasmid pGro12, which has a kanamycin resistance gene (Brixius-Anderko et al., 2015; Nishihara et al., 1998). Cultures were grown at 37 °C to an optical density of 0.6 in 200 ml TB medium containing the suitable antibiotics. The expression of the protein was induced by adding 1 mM IPTG, the synthesis of heme was enhanced by addition of 1 mM heme precursor δ -ALA. The cells were grown at 28 °C and 180 rpm for 24 h.

For purification, cell pellets were sonicated in 50 ml lysis buffer (50 mM potassium phosphate pH 7.4, 20% glycerol, 0.1 mM DTE, 500 mM sodium acetate, and 0.1 mM PMSF). After centrifugation

at 30,000g for 30 min at 4 °C, the supernatant was applied on a Ni–NTA agarose column equilibrated with lysis buffer. The column was washed with 100 ml equilibration buffer supplemented with 40 mM imidazole followed by 20 ml elution buffer supplemented with 200 mM imidazole (50 mM potassium phosphate pH 7.4, 20% glycerol, 0.1 mM DTE, and 0.1 mM PMSF). The eluted protein was dialyzed against elution buffer without imidazole, concentrated and stored at –80 °C.

The *B. megaterium* redox system BmCPR and Fdx2 were purified as reported previously (Brill et al., 2013; Milhim et al., 2016). The purification of the redox system from the fission yeast *Schizosaccharomyces pombe* Arh1 and Etp1^{fd} was carried out as described before (Bureik et al., 2002; Ewen et al., 2008). Recombinant bovine AdR and the Adx₄₋₁₀₈ (truncated form of Adx comprising amino acids 4–108) were purified as mentioned elsewhere (Sagara et al., 1993; Uhlmann et al., 1992).

The concentration of recombinant P450 was estimated using the CO-difference spectral assay as described previously with $\epsilon_{450-490} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ (Omura and Sato, 1964). The concentration of BmCPR was quantified by measuring the flavin absorbance at 456 nm with $\epsilon_{456} = 21 \text{ mM}^{-1} \text{ cm}^{-1}$ for the oxidized enzyme (Milhim et al., 2016). The concentrations of the AdR and Arh1 were measured using the extinction coefficient $\epsilon_{450} = 11.3 \text{ mM}^{-1} \text{ cm}^{-1}$ (Ewen et al., 2008; Hiwatashi et al., 1976). The concentrations of Fdx2 and Etp1^{fd} were measured using the extinction coefficient $\epsilon_{390} = 6.671 \text{ mM}^{-1} \text{ cm}^{-1}$ and $\epsilon_{414} = 9.8 \text{ mM}^{-1} \text{ cm}^{-1}$, respectively (Brill et al., 2013; Schiffler et al., 2004).

2.4. Investigation of electron transfer partners

The functional interaction of the electron transfer partners for a particular P450 can be determined by recording the NADPH reduced CO-complex peak at 450 nm when P450 was coupled with the different ferredoxins/ferredoxin reductases in the absence of substrate. For this, CYP107DY1 was mixed with ferredoxins (Fdx2, Etp1^{fd} or Adx₄₋₁₀₈) and ferredoxin reductases (BmCPR, Arh1 or AdR) with ratios of 1:40:5 μM [CYP107DY1:ferredoxin:ferredoxin reductase] in 50 mM HEPES buffer pH 7.4 and NADPH was added to a final concentration of 1 mM. The spectrum of NADPH-reduced samples was recorded after bubbling the sample with carbon monoxide (CO) gas. The reduction efficiency of the redox partners was then evaluated by comparing the peak at 450 nm of the CO-complexed CYP107DY1 reduced with the different redox systems and the peak at 450 nm of the CO-complexed CYP107DY1 reduced with sodium-dithionite.

2.5. In vitro conversion and HPLC analysis

The in vitro conversion of the substrates was carried out with a reconstituted system at 30 °C in conversion buffer (50 mM HEPES, pH 7.4, 20% glycerol). The reconstituted system contained 0.5 μM CYP107DY1, 2.5 μM BmCPR, 20 μM Fdx2, 1 mM MgCl_2 , 5 mM glucose-6-phosphate, 1 U glucose-6-phosphate dehydrogenase for NADPH regeneration and 100 μM substrate. The reaction was started by adding NADPH (200 μM) and stopped after 15 min by the addition of 1 vol ethyl acetate and extracted twice. The organic phase was evaporated under vacuum. Residuum was dissolved in 20% acetonitrile/water mixture and subjected to HPLC analysis. HPLC analysis was performed using a Jasco system. A reversed-phase ec MN Nucleodur C18 (4.0 \times 125 mm) column (Macherey–Nagel) was used for all experiments at an oven temperature of 40 °C. Mevastatin and its metabolite pravastatin were eluted from the column using a gradient of acetonitrile from 20 to 100% in water over 20 min. The detection wavelength of mevastatin and its metabolite pravastatin was 236 nm.

2.6. Circular dichroism (CD) spectroscopy

Circular dichroism (CD) spectra were recorded at 30 °C using a JASCO J-715 spectropolarimeter over the wavelength range 190–260 nm for the far-UV region and 300–500 nm for the near-UV/Vis region at a protein concentration of 2 µM (~0.1 mg/ml) and 20 (~1 mg/ml), respectively, dissolved in 10 mM potassium phosphate buffer pH 7.4 with the following parameters: path-length of 0.1 cm for the far-UV region and 1 cm for the near-UV/Vis region measurement, data pitch of 0.1 nm, band width of 5 nm, accumulation 3 times. Spectra were recorded in triplicate and averaged.

2.7. Homology modeling of CYP107DY1 and molecular docking with mevastatin

Using the homology modeling program Modeller 9.14 (University of California San Francisco, USA), a model of CYP107DY1 was calculated using CYP107RB1 (Vdh) (PDB accession code: 3A4G) as template. The coordinates of the heme-porphyrin atoms from the template structure were added subsequently to the obtained homology model. No further structural refinement of the model was performed.

The three dimensional structure of the mevastatin molecule was obtained from PubChem (Kim et al., 2016). The docking simulations of the homology model of CYP107DY1 with the mevastatin molecule were carried out using Autodock 4.0 (Morris et al., 2009). The Windows version 1.5.6 of Autodock Tools was used to compute Kollman charges for the enzyme and Gasteiger-Marsili charges for the ligand (Sanner, 1999). 200 docking runs were carried out applying the Lamarckian genetic algorithm using default parameter settings.

2.8. In vivo whole-cell biotransformation

For the establishment of a CYP107DY1-dependent whole-cell system, a tricistronic pET17b-based vector, encoding the CYP107DY1-BmCPR-Fdx2 genes, was constructed. CYP107DY1 coding sequence was amplified and cloned via the restriction sites *NdeI/HindIII*. The resulting vector (pET17b.CYP107DY1) served then as a backbone for the cloning of BmCPR, which was amplified via PCR and cloned with the restriction sites *BamHI/NotI*. Fdx2 coding region was PCR amplified and cloned downstream CYP107DY1-BmCPR using the restriction sites *NotI/XhoI*. All resulting vectors were verified by sequencing.

E. coli C43 (DE3) cells were co-transformed with the suitable expression vector based on pET17b, coding for CYP107DY1-BmCPR-Fdx2, and the chaperone GroEL/GroES-coding plasmid pGro12, which has a kanamycin resistance gene (Brixius-Anderko et al., 2015; Nishihara et al., 1998). Transformed cells were grown overnight in 50 ml LB broth medium supplemented with 100 µg/ml ampicillin and 50 µg/ml kanamycin at 37 °C and shaking at 180 rpm. For the expression of proteins, 50 ml TB medium containing 100 µg/ml ampicillin and 50 µg/ml kanamycin were inoculated (1:100) with the transformed cells and cultivated at 37 °C with rotary shaking at 140 rpm. Protein expression was induced at OD₆₀₀ = 0.6–0.8 with 1 mM IPTG, 4 mg/ml arabinose (for induction of GroES/GroEL expression) and 1 mM δ-ALA as heme precursor. The temperature was then reduced to 28 °C. After incubation for 24 h, cells were harvested by centrifugation (4000g) for 20 min at 4 °C and washed once with 1 vol of conversion buffer (50 mM potassium phosphate buffer (pH 7.4) supplemented with 2% glycerol). After a second centrifugation, the cell pellets were resuspended in conversion buffer to an end cell-suspension concentration of 60g wet cell weight (wcw)/L buffer. The substrate mevastatin was added to a final concentration of 100 µM and the culture was incubated for the indicated time at 30 °C and 140 rpm

for 20 h. To enable higher substrate conversion, EDTA (20 mM) or polymyxin B (32 µg/ml) was added to increase permeability and substrate uptake of the *E. coli* cells (Janocha and Bernhardt, 2013; Kern et al., 2016). Substrate was extracted twice with the same volume of ethyl acetate and the organic phase was evaporated using a rotary evaporator. After that, the residues were dissolved in the high performance liquid chromatography (HPLC) mobile phase (20% ACN) and subjected to HPLC analysis using the same method mentioned in Section 2.5.

3. Results and discussion

3.1. Bioinformatic analysis

The publication of the complete genome sequence of the *B. megaterium* strains QM B1551 and DSM319 in 2011 (Eppinger et al., 2011) enabled the characterization of new proteins. Recently, we were able to identify and characterize different P450s and related proteins from the strain DSM319 (Brill et al., 2013; Gerber et al., 2015; Milhim et al., 2016). In contrast to the DSM319 strain, strain QM B1551 harbors seven indigenous plasmids (Eppinger et al., 2011). Sequence analysis of the indigenous plasmids showed that plasmid number 5 (pBM500) has an open reading frame (BMQ_pBM50008) containing 1233 base pairs, which encodes for a protein comprising 410 amino acids with a predicted molecular weight of about 46.742 kDa. The analysis of the BMQ_pBM50008 domains using the Pfam Database (Finn et al., 2016) for highly conserved motifs of the P450 family proved the presence of the heme-binding motif (F-x-x-G-x-x-C-x-G), the (A/G-G-x-E/D-T-T/S) motif in the I-helix and the (E-x-x-R) motif in the K-helix (Supplementary Fig. S1), suggesting its identification as a cytochrome P450.

Multiple sequence alignment showed that the protein sequence of BMQ_pBM50008 belongs to the CYP107 family. After submission of the protein sequence of BMQ_pBM50008 to the P450 nomenclature committee (Prof. Dr. David Nelson), it was found to match best to CYP107DA1 with 49% sequence identity and, therefore, was assigned as a new subfamily with the name CYP107DY1. CYP107 is the largest family among bacterial P450s comprising more than 2500 subfamily members (<https://cyped.biocatnet.de/sFam/107>). Besides the CYP105 family, members of the CYP107 family are shown to be the most studied bacterial P450s that participate in the degradation and biotransformation of a broad spectrum of xenobiotics as well as in secondary metabolite biosynthesis, and, therefore, are considered to be important for industrial biotechnology; for example CYP107BR1 from *Pseudomonas autotrophica* for the activation of Vitamin D₃ (Sakaki et al., 2011), CYP107E from *Micromonospora griseorubida* for mycinamicin biosynthesis (Inouye et al., 1994), and P450_{terf} (CYP107L) from *Streptomyces platensis* for the hydroxylation of terfenadine (Lombard et al., 2011).

CYP107DY1 is the first P450 found to be encoded on a bacillus sp. plasmid. There are only few examples of plasmid-encoded P450s described so far in the literature, such as the CYP107A2 (LkmF) and CYP107AP1 (LkmK) from *Streptomyces rochei*, which participate in the biosynthesis of the macrolide antibiotic lankomycin (Arakawa et al., 2006) and CYP102H1 from *Nocardia farcinica* that catalyzes the hydroxylation of linoleic acid (Chung et al., 2012).

The plasmid-encoded nature of the identified P450 is very interesting since the absence of CYP107 members in the genome of *B. megaterium* indicates that the presence of the corresponding gene of the CYP107DY1 on the plasmid may be due to horizontal gene transfer displacement rather than intragenic transfer from the chromosome to the plasmid. This suggests that the P450s can play an important role in adaptation and evolution in prokaryotes.

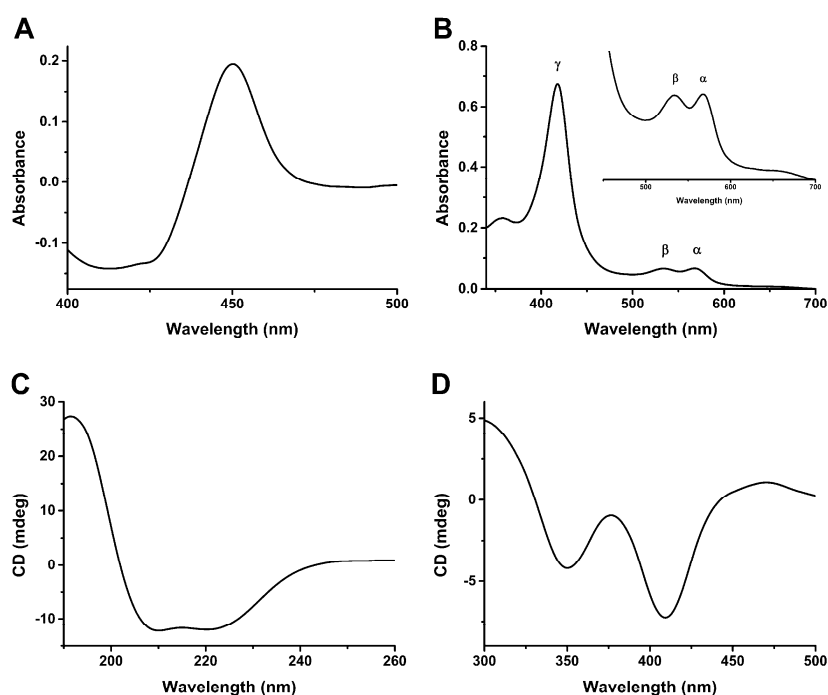


Fig. 1. Spectral characteristics of CYP107DY1. (A) The CO difference spectrum of CYP107DY1. (B) The UV-vis spectral characteristics of the purified CYP107DY1. The inset shows the magnification of the spectrum in the α and β band region. Circular dichroism (CD) spectra in the far-UV (C) and in the near-UV/visible (D) region at a CYP107DY1 concentration of 2 μ M and 20 μ M, respectively, resuspended in 10 mM potassium phosphate buffer pH 7.4. The CD spectrum was recorded using the following parameters: path length 1 mm (for far-UV) and 5 mm (near-UV/visible); time constant 2 s; band pass 5 nm; number of scans 3.

3.2. Expression, purification and spectrophotometric characterization

The DNA fragment encoding the full length CYP107DY1 was cloned into the expression vector pET17b and expressed in *E. coli* C43 (DE3). The expression levels were calculated by measuring reduced CO difference spectrum in the cell lysate (Omura and Sato, 1964). The carbon monoxide bound form gave a typical peak maximum at 450 nm (Fig. 1A). The full-length CYP107DY1 was purified in a soluble form with an expression level of 20 nmol/L. It was previously shown that the co-expression of some P450s with chaperones leads to an improvement of protein folding and thus increases of the expression level (Arase et al., 2006; Brixius-Anderko et al., 2015; Nishihara et al., 1998). The co-expression with the chaperones GroES and GroEL also improved the expression of CYP107DY1 more than 10 times, yielding 210 nmol/L.

Besides the CO-difference spectrum, the UV-vis absorption spectrum provides a simple technique for the characterization of P450 enzymes. The oxidized form of substrate free CYP107DY1 exhibited a major (γ) Soret peak at 417 and the smaller α and β bands at 567 and 535 nm, respectively (Fig. 1B), indicating a low spin state of the heme iron in the P450. In addition, the CD spectra of the oxidized CYP107DY1 were measured in the far-UV- and near UV-vis region. The far-UV CD spectrum showed a negative dichroic double band with minima at 208 and 222 nm (Fig. 1C), an indication of predominantly α -helical secondary structure (Poulos et al., 1986, 1987; Ravichandran et al., 1993). In the near UV-vis region, CYP107DY1 displayed two large negative signals at 350 nm and at 408 nm (Fig. 1D). These are in correspondence with the characteristic peaks for other bacterial P450s (Lepesheva et al., 2001; Munro et al., 1994).

Taken together, the spectrophotometric properties of the purified CYP107DY1 indicate that the enzyme is produced in the active form with proper heme incorporation.

3.3. Searching for a suitable redox partner

The availability of a redox partner/s is essential for studying the functionality of CYP107DY1. By searching the ORFs around the CYP107DY1 coding region as well as the ORFs on the seven indigenous plasmids of the QM B1551 strain, no redox partner (reductase and/or ferredoxin) was identified. Therefore, we tested herein different autologous and heterologous redox systems. As autologous electron transfer partners, the diflavin reductase BmCPR and the ferredoxin Fdx2 of *B. megaterium* DSM319 were selected. The BmCPR-Fdx2 system has been shown to support efficiently the activity of CYP106A1 (Milhim et al., 2016). As heterologous redox partners, the soluble *Schizosaccharomyces pombe* redox system adrenodoxin reductase homologue 1 (Arh1) and its ferredoxin (Etp1^{fd}) as well as the bovine adrenodoxin reductase (AdR) and adrenodoxin (Adx₄₋₁₀₈) were used. The redox systems Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈ are described to transfer electrons to different classes of P450s (Brixius-Anderko et al., 2015; Ly et al., 2012).

Based on the measurement of the reduced CO-bound spectrum of CYP107DY1, the redox partners were tested and compared. Using the heterologous redox partners Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈ very low peaks (<5%) compared with the dithionite reduced CO difference peak at 450 nm were recovered within 5 min. In contrast, the autologous redox partners BmCPR-Fdx2 were very efficient and able to recover ~99.5% of the peak (Fig. 2). Therefore, the BmCPR-Fdx2 redox system was selected for further investigations with CYP107DY1.

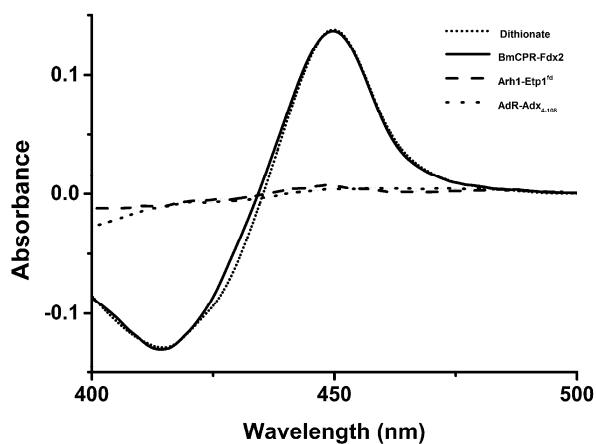


Fig. 2. Determination of CYP107DY1 reduction using autologous and heterologous electron transfer partners. The dithionite reduced CO-difference spectrum (short dot line) of CYP107DY1 was compared with the NADPH reduced BmCPR-Fdx2 (solid line), Arh1-Etp1^{fd} (dash line) and AdR-Adx₄₋₁₀₈ (dot line) CO-complexed spectrum. The NADPH (1 mM) reduced CO-difference spectra were recorded in a 200 μ l mixture of CYP/ferredoxin/reductase with a 1:40:5 molar ratio in 50 mM HEPES buffer pH 7.4 containing 20% glycerol (see Materials and methods).

3.4. In vitro conversion of mevastatin

The identification of suitable substrates and the characterization of the catalytic activity of a new enzyme are significant challenges. Therefore, at first a phylogenetic based approach was used in this study to predict a functional homolog of the new P450. The deduced amino acid sequence of CYP107DY1 was aligned with some related bacterial P450s with known function or substrate/s. The aligned sequences were then used for the construction of a neighbor-joining phylogenetic tree, which shows that CYP107DY1 is clustered with CYP267B1 from *Sorangium cellulosum* So ce56 (Fig. 3). CYP267B1 belongs to the CYP107 family (<http://drnelson.uthsc.edu/Bacteria.html>) and was previously characterized as a versatile drug metabolizer (Kern et al., 2016). Therefore, using an in vitro reconstituted assay we screened the activity of CYP107DY1 towards various pharmaceutical substrates (Supplementary Table 1). Among the 18 tested drugs, CYP107DY1 was found to metabolize mevastatin producing one product at a retention time of 9 min (Fig. 4B) compared with the negative control (Fig. 4A). The conversion ratio of 100 μ M mevastatin within 15 min was ~30%. The product retention time was identical to that of a pravastatin authentic standard (Fig. 4C). The results presented in Fig. 5 showed that CYP107DY1 activity can be supported by all tested redox systems, with an obvious higher efficiency using BmCPR-Fdx2. The conversion ratio of the 100 μ M mevastatin using Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈ as redox partners was ~5% and ~3%, respectively, compared with when using of the BmCPR-Fdx2.

The regioselective oxyfunctionalization of the precursor mevastatin is crucial for the production of pravastatin, the widely used therapeutic agent for hypercholesterolemia (Lamon-Fava, 2013). This was described previously using CYP105A3 (P450sca2) from *Streptomyces carbophilus* (Matsuoka et al., 1989) and mutant CYP105AS1 from *Amycolatopsis orientalis* (McLean et al., 2015). However, CYP107DY1 is the first reported P450 of the CYP107 family, which can highly selectively hydroxylate such statin drugs.

3.5. Homology modeling and docking of CYP107DY1 with mevastatin

For the prediction of the three dimensional structure of CYP107DY1, homology modeling was performed. In order to select

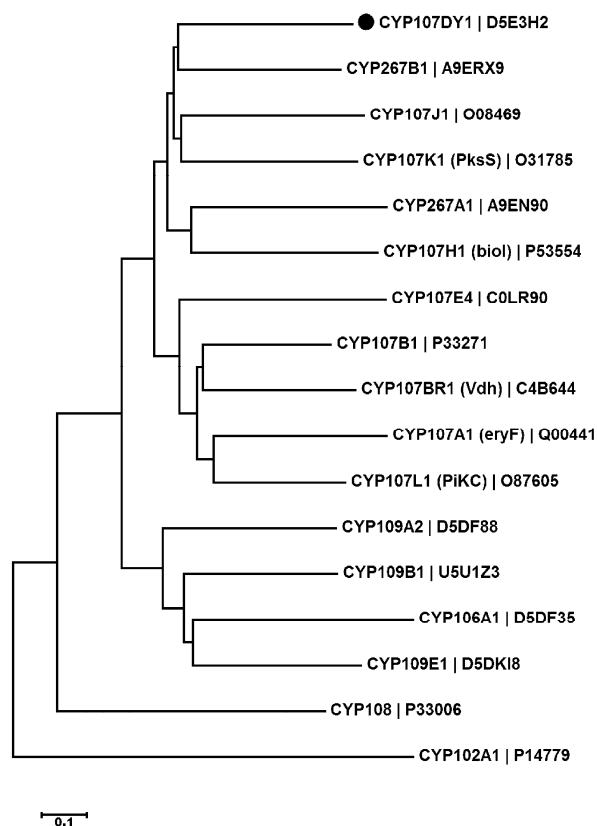


Fig. 3. Evolutionary relationships of CYP107DY1. The alignment was done with 10-gap setting and 0.1-gap extension, with slow alignment input in ClustalW2 server. The tree was constructed by neighbor joining algorithm with bootstrap analysis of 1000 replicates. The scale bar represents 0.1 amino acid substitution per amino acid. The CYP107DY1 is indicated with the closed circle (●). The number next to the gene name represents the UniProtKB accession number.

a suitable template, the amino acid sequence of the CYP107DY1 was submitted to the RCSB protein data bank BLAST server (<http://www.rcsb.org/pdb/search/advSearch.do?search=new>). The CYP107BR1 (Vdh) from *Pseudonocardia autotrophica* (UniProt acc. no.: C4B644) showed a sequence identity of 40% to CYP107DY1. The corresponding crystal structure of CYP107BR1 (Vdh) (PDB acc. code: 3A4G) was, therefore, chosen as a structural template. The CYP107DY1 homology model (Supplementary Fig. S2) has a typical three-dimensional P450 structure consisting of a C-terminus relatively rich in α -helices and an N-terminus relatively rich in β -sheets. Overall, it comprises 13 α -helices designated A, B, B₁ and C-L and 12 β -sheets grouped into 5 regions. The prosthetic heme group is embedded in the active site and surrounded by the I-helix from the proximal side and the L-helix from the distal side. The conserved cysteine residue (Cys360) is located in the loop region preceding the L-helix. The key structural features of the CYP107DY1 suggest that its overall topology correlates with the general folding properties of P450s (Supplementary Fig. S2).

Three-dimensional structures of P450s are very helpful to understand the enzyme-substrate interaction. Therefore, mevastatin was docked into CYP107DY1. In the docking simulation the target protein was kept as rigid body. The flexible bonds of the ligands were assigned automatically and verified by manual inspection. A cubic grid box with a size of 52 $\text{Å} \times 52 \text{Å} \times 52 \text{Å}$ was fixed above the heme moiety to cover the whole active site of the enzyme. As illustrated in Fig. 6, the substrate molecule is located over the

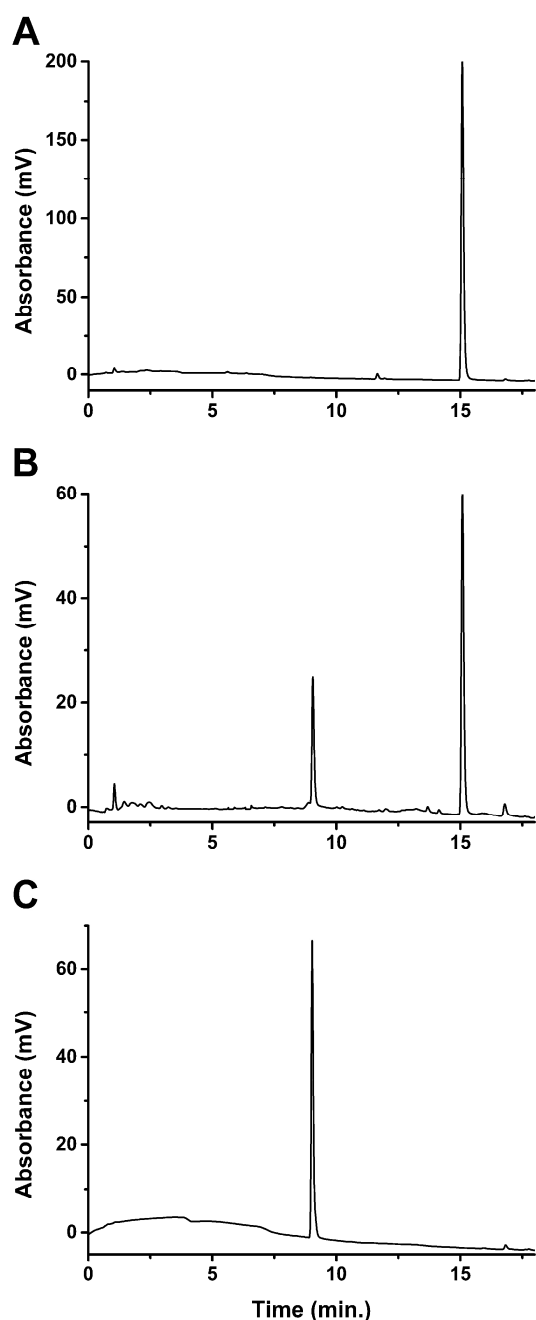


Fig. 4. HPLC analysis of the conversion of mevastatin with CYP107DY1. In vitro conversion of 100 μ M mevastatin in the (A) absence of CYP107DY1 or (B) presence of CYP107DY1. The reaction was carried out in 50 mM HEPES buffer pH 7.4 with 20% glycerol. The reconstituted system contained 0.5 μ M CYP107DY1, 2.5 μ M BmCPR, 20 μ M Fdx2 and the NADPH regeneration system. The conversion was carried out for 15 min at 30 °C. (C) Pravastatin authentic standard.

heme. In addition, the side chains of seven amino acid residues (including Leu99, Leu245, Thr249, Ala250, Glu253, Thr254 and Gly300) were found to define a cavity with mevastatin. The keto group of the 2-methylbutanoate side chain and the hydroxyl group on the lactone ring of the mevastatin molecule formed two hydrogen bonds with Thr249 and Glu253, respectively. These interactions enable the orientation of the molecule in a position that allows

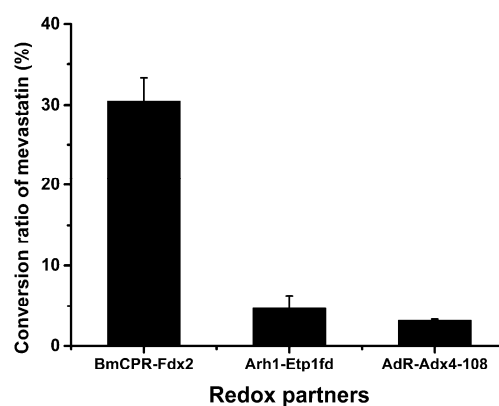


Fig. 5. Effect of the different redox partners on the in vitro conversion of mevastatin by CYP107DY1. In vitro conversion of 100 μ M mevastatin by CYP107DY1 using the redox system BmCPR-Fdx2 of, Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈. The reaction was carried out in 50 mM HEPES buffer pH 7.4 with 20% glycerol. The reconstituted system contained 0.5 μ M CYP107DY1, 2.5 μ M reductase (BmCPR, AdR or Arh1), and 20 μ M ferredoxin (Fdx2, Etp1^{fd} or Adx₄₋₁₀₈) and the NADPH regeneration system. The reaction was carried out for 15 min at 30 °C. The data is represented as mean \pm SD of three separate measurements.

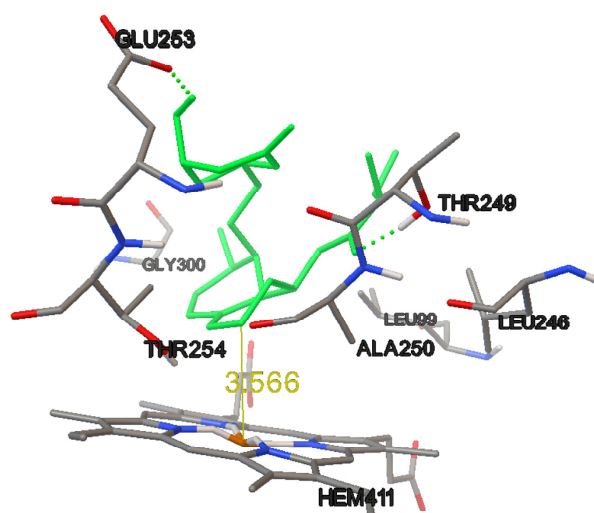
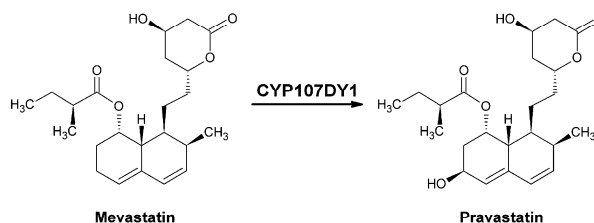


Fig. 6. Docking conformation of mevastatin. The ligand mevastatin is shown in green. The distance of the C-6 atom from the heme iron is shown in yellow in Å. Some of the amino acids that form the active site of the CYP107DY1 in the presence of mevastatin are shown and named. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



Scheme 1. Conversion of mevastatin to pravastatin by CYP107DY1.

atom C-6 to face the heme iron (distance \sim 3.5 Å) (Fig. 6), favorable for a hydroxylation at this position to produce the pravastatin (Scheme 1), supporting our experimental data shown in Fig. 4.

Table 1
CYP107DY1-based ^(a) *Escherichia coli* Whole-cell biotransformation yield of mevastatin.

Additives ^(b)	Pravastatin yield (mg/L)
Without additives	13.2 ± 2.5
Polymyxin B	17.9 ± 2.1
EDTA	28.5 ± 3.1

^a *E. coli* C43 (DE3) cells were co-transformed with the expression vector based on pET17b, coding for CYP107DY1-BmCPR-Fdx2, and the chaperone GroEL/GroES-coding plasmid pGro12.

^b Polymyxin B and EDTA were added to a final concentration of 32 µg/ml and 20 mM, respectively.

3.6. Biotransformation of mevastatin using a whole-cell based system

The in vitro activity of CYP107DY1 towards mevastatin was very encouraging to address the question whether it is applicable for the in vivo production of pravastatin. Therefore, and as a proof of concept, a CYP107DY1 based *E. coli* whole-cell biotransformation system, utilizing BmCPR and Fdx2 as redox partners, was established. We chose *E. coli* as host, because, in contrast to *B. megaterium*, it has the advantage that no endogenous P450 can interfere with the desired reaction. For this purpose, a tricistronic pET17b-based vector, expressing CYP107DY1, BmCPR and Fdx2, was constructed. The whole-cell biotransformation was performed with resting cells in potassium phosphate buffer instead of the terrific broth complex medium, since indole, which results from the tryptophan metabolism by *E. coli*, may have an inhibitory effect on the activity of P450s (Brixius-Anderko et al., 2016; Ringle et al., 2013).

As shown in Table 1, the whole-cell biotransformation system yielded about 13.2 mg/L pravastatin. This amount is comparable to the previously reported WT CYP105A3 (P450sca2) *E. coli*-based whole-cell system, which yielded 12.8 mg/L pravastatin (Ba et al., 2013a, 2013b). It was shown previously that the hindered transport of the substrate into the *E. coli* cells is most likely responsible for the low biotransformation ratio (Janocha and Bernhardt, 2013). Therefore, we used the peptide antibiotic polymyxin B and the chelating agent EDTA, which are reported to exhibit permeabilization activity towards the outer membrane of the Gram-negative bacteria. Concerning this, *E. coli*-based whole cell systems were also designed for other P450s in our laboratory, in which optimal concentrations of polymyxin B (32 µg/ml) and EDTA (20 mM) were employed to achieve the highest conversion of substrates (Janocha and Bernhardt, 2013; Kern et al., 2016; Litzenburger et al., 2015). The use of polymyxin B lead to an 1.3 fold increase in the pravastatin production, yielding 17.9 mg/L, while the highest yield was achieved using EDTA with an 2.2 fold increase yielding 28.5 mg/L (Table 1), which suggests the potential of the constructed whole-cell system in the production of pravastatin.

Novel P450s are increasingly emerging because of the availability of many genome sequences. However, the multi-component nature of the P450s and other factors (Bernhardt and Urlacher, 2014) limit their biotechnological use. The characterization of CYP107DY1 and its successful employment in the in vivo pravastatin production is the first step in the industrialization of this interesting P450. The application of this novel pravastatin-producing P450 in an industrially optimized host as recently shown for the mutant CYP105AS1 (McLean et al., 2015) will certainly pave the way for a successful use in a biotechnological process.

4. Conclusion

Regio- and stereoselective oxidation of non-activated carbon atoms using P450s is of great interest for synthetic biology. The

robustness of the bacterial P450s is attractive for the industrial production of different valuable products (Bernhardt and Urlacher, 2014). However, despite the high potential and the recent development with respect to exploiting the P450s in biotechnological applications, there are various challenges limiting their use in industrial processes, for instance their dependency on electron transfer proteins. In the recent study, we have identified, cloned and expressed a new plasmid-encoded cytochrome P450 from *B. megaterium* QM B1551, which was categorized into a new sub-family (CYP107DY1). In addition, we were able to find a redox system (BmCPR-Fdx2) that can efficiently support the activity of the CYP107DY1, which enabled the identification of mevastatin as a substrate. The successful construction of *E. coli*-based whole cell system utilizing CYP107DY1 is of great importance not only for efficient production of pravastatin but also for future characterization of new substrates, functionalization of drugs and the production of new metabolites. Further improvements, either at the cellular level by expressing CYP107DY1 in other microorganisms (McLean et al., 2015), or at the molecular level by rational design to improve the activity of this P450 (Ba et al., 2013b) will contribute to the set-up of efficient industrial processes using this novel P450.

Author contributions

M.M. carried out all experiments, analyzed and interpreted the data and drafted the manuscript. N.P. and F.K. participated in the establishment of the experiments. A.A. participated in the substrate docking. P.H. and R.B. participated in the interpretation and discussion of experimental results and writing of the manuscript.

Conflict of interest

The authors declare that they have no competing interest.

Acknowledgments

The authors would like to thank Birgit Heider-Lips for the purification of AdR and Adx, Dr. Elisa Brill for cloning of Fdx2 and Tanja Sagadin for the purification of Arh1 and Etp1^{td}. The authors also gratefully acknowledge Prof. Dr. David Nelson (The University of Tennessee Health Science Center) for helping in the CYP107DY1 classification.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jbiotec.2016.11.002>.

References

- Arakawa, K., Kodama, K., Tatsuno, S., Ide, S., Kinashi, H., 2006. Analysis of the loading and hydroxylation steps in lankamycin biosynthesis in *Streptomyces rochei*. *Antimicrob. Agents Chemother.* 50, 1946–1952.
- Arase, M., Waterman, M.R., Kagawa, N., 2006. Purification and characterization of bovine steroid 21-hydroxylase (P450c21) efficiently expressed in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 344, 400–405.
- Ba, L., Li, P., Zhang, H., Duan, Y., Lin, Z., 2013a. Engineering of a hybrid biotransformation system for cytochrome P450sca-2 in *Escherichia coli*. *Biotechnol. J.* 8, 785–793.
- Ba, L., Li, P., Zhang, H., Duan, Y., Lin, Z., 2013b. Semi-rational engineering of cytochrome P450sca-2 in a hybrid system for enhanced catalytic activity: insights into the important role of electron transfer. *Biotechnol. Bioeng.* 110, 2815–2825.
- Berg, A., Gustafsson, J.A., Ingelman-Sundberg, M., 1976. Characterization of a cytochrome P450-dependent steroid hydroxylase system present in *Bacillus megaterium*. *J. Biol. Chem.* 251, 2831–2838.
- Berg, A., Ingelman-Sundberg, M., Gustafsson, J.A., 1979. Isolation and characterization of cytochrome P450meg. *Acta Biol. Med. Ger.* 38, 333–344.

- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98, 6185–6203.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124, 128–145.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2013. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1703–1717.
- Brixius-Anderko, S., Schiffer, L., Hannemann, F., Janocha, B., Bernhardt, R., 2015. A CYP21A2 based whole-cell system in *Escherichia coli* for the biotechnological production of premedrol. *Microb. Cell Fact.* 14, <http://dx.doi.org/10.1186/s12934-015-0333-2>.
- Brixius-Anderko, S., Hannemann, F., Ringle, M., Khatri, Y., Bernhardt, R., 2016. An indole deficient *Escherichia coli* strain improves screening of cytochromes P450 for biotechnological applications. *Biotechnol. Appl. Biochem.*, <http://dx.doi.org/10.1002/bab.1488>.
- Bunk, B., Schulz, A., Stammen, S., Münch, R., Warren, M.J., Rohde, M., Jahn, D., Biedendieck, R., 2010. A short story about a big magic bug. *Bioeng. Bugs* 1, 85–91.
- Bureik, M., Schiffler, B., Hiraoka, Y., Vogel, F., Bernhardt, R., 2002. Functional expression of human mitochondrial CYP11B2 in fission yeast and identification of a new internal electron transfer protein, etp1. *Biochemistry (Mosc.)* 41, 2311–2321.
- Chung, Y.-H., Song, J.-W., Choi, K.-Y., Yoon, J.W., Yang, K.-M., Park, J.-B., 2012. Cloning, expression, and characterization of P450 monooxygenase CYP102H1 from *Nocardia farcinica*. *J. Korean Soc. Appl. Biol. Chem.* 55, 259–264.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin, M., Elbourne, L.D.H., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukyanchuk, V., Ball, B., Ul-Qamar, R., Seibel, J., Bremer, E., Jahn, D., Ravel, J., Vary, P.S., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J. Bacteriol.* 193, 4199–4213.
- Ewen, K.M., Schiffler, B., Uhlmann-Schiffler, H., Bernhardt, R., Hannemann, F., 2008. The endogenous adrenodoxin reductase-like flavoprotein arh1 supports heterologous cytochrome P450-dependent substrate conversions in *Schizosaccharomyces pombe*. *FEMS Yeast Res.* 8, 432–441.
- Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Fact.* 14, 107, <http://dx.doi.org/10.1186/s12934-015-0300-y>.
- Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R., 2007. Cytochrome P450 systems—biological variations of electron transport chains. *Biochim. Biophys. Acta BBA Gen. Subj.* 1770 (P450), 330–344.
- Hiwatashi, A., Ichikawa, Y., Maruya, N., Yamano, T., Aki, K., 1976. Properties of crystalline reduced nicotinamide adenine-dinucleotide phosphate-adrenodoxin reductase from bovine adrenocortical mitochondria. I. Physicochemical properties of holo-NADPH-adrenodoxin and apo-NADPH-adrenodoxin reductase and interaction between non-heme iron proteins and reductase. *Biochemistry (Mosc.)* 15, 3082–3090.
- Inouye, M., Takada, Y., Muto, N., Beppu, T., Horinouchi, S., 1994. Characterization and expression of a P450-like mycinamicin biosynthesis gene using a novel *Micromonospora- Escherichia coli* shuttle cosmid vector. *Mol. Gen. Genet.* MGG 245, 456–464.
- Jóźwik, I.K., Kiss, F.M., Gricman, L., Abdulmughni, A., Brill, E., Zapp, J., Pleiss, J., Bernhardt, R., Thunnissen, A.-M.W.H., 2016. Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J.*, <http://dx.doi.org/10.1111/febs.13911>.
- Janocha, S., Bernhardt, R., 2013. Design and characterization of an efficient CYP105A1-based whole-cell biocatalyst for the conversion of resin acid diterpenoids in permeabilized *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 97, 7639–7649.
- Kern, F., Khatri, Y., Litzenburger, M., Bernhardt, R., 2016. CYP267A1 and CYP267B1 from *Sorangium cellulosum* So ce56 are highly versatile drug metabolizers. *Drug Metab. Dispos.* 44, 495–504.
- Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., Wang, J., Yu, B., Zhang, J., Bryant, S.H., 2016. PubChem substance and compound databases. *Nucleic Acids Res.* 44, D1202–1213.
- Lamon-Fava, S., 2013. Statins and lipid metabolism: an update. *Curr. Opin. Lipidol.* 24, 221–226.
- Lepesheva, G.I., Podust, L.M., Bellamine, A., Waterman, M.R., 2001. Folding requirements are different between sterol 14 α -demethylase (CYP51) from *Mycobacterium tuberculosis* and human or fungal orthologs. *J. Biol. Chem.* 276, 28413–28420.
- Litzenburger, M., Kern, F., Khatri, Y., Bernhardt, R., 2015. Conversions of tricyclic antidepressants and antipsychotics with selected P450s from *Sorangium cellulosum* So ce56. *Drug Metab. Dispos.* 43, 392–399.
- Lombard, M., Salard, I., Sari, M.-A., Mansuy, D., Buisson, D., 2011. A new cytochrome P450 belonging to the 107L subfamily is responsible for the efficient hydroxylation of the drug terfenadine by *Streptomyces platensis*. *Arch. Biochem. Biophys.* 508, 54–63.
- Ly, T.T.B., Khatri, Y., Zapp, J., Hutter, M.C., Bernhardt, R., 2012. CYP264B1 from *Sorangium cellulosum* So ce56: A fascinating norisoprenoid and sesquiterpene hydroxylase. *Appl. Microbiol. Biotechnol.* 95, 123–133.
- Matsuoka, T., Miyakoshi, S., Tanzawa, K., Nakahara, K., Hosobuchi, M., Serizawa, N., 1989. Purification and characterization of cytochrome P450sca from *Streptomyces carbophilus*. ML-236B (compactin) induces a cytochrome P450sca in *Streptomyces carbophilus* that hydroxylates ML-236B to pravastatin sodium (CS-514), a tissue-selective inhibitor of 3-hydroxy-3-methylglutaryl-coenzyme-A reductase. *Eur. J. Biochem. FEBS* 184, 707–713.
- McLean, K.J., Hans, M., Meijrink, B., van Scheppingen, W.B., Vollebregt, A., Tee, K.L., van der Laan, J.-M., Leys, D., Munro, A.W., van den Berg, M.A., 2015. Single-step fermentative production of the cholesterol-lowering drug pravastatin via reprogramming of *Penicillium chrysogenum*. *Proc. Natl. Acad. Sci. U. S. A.* 112, 2847–2852.
- Milhim, M., Gerber, A., Neunzig, J., Hannemann, F., Bernhardt, R., 2016. A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase. *J. Biotechnol.* 231, 83–94.
- Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791.
- Munro, A.W., Lindsay, J.G., Coggins, J.R., Kelly, S.M., Price, N.C., 1994. Structural and enzymological analysis of the interaction of isolated domains of cytochrome P450 BM3. *FEBS Lett.* 343, 70–74.
- Nelson, D.R., 2009. The cytochrome P450 homepage. *Hum. Genomics* 4, 59–65.
- Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., Yura, T., 1998. Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* 64, 1694–1699.
- Omura, T., Sato, R., 1964. The carbon monoxide-binding pigment of liver microsomes. II. Solubilization purification, and properties. *J. Biol. Chem.* 239, 2379–2385.
- Poulos, T.L., Finzel, B.C., Howard, A.J., 1986. Crystal structure of substrate-free *Pseudomonas putida* cytochrome P450. *Biochemistry (Mosc.)* 25, 5314–5322.
- Poulos, T.L., Finzel, B.C., Howard, A.J., 1987. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* 195, 687–700.
- Ravichandran, K.G., Boddupalli, S.S., Hasermann, C.A., Peterson, J.A., Deisenhofer, J., 1993. Crystal structure of hemoprotein domain of P450 BM3, a prototype for microsomal P450's. *Science* 261, 731–736.
- Ringle, M., Khatri, Y., Zapp, J., Hannemann, F., Bernhardt, R., 2013. Application of a new versatile electron transfer system for cytochrome P450-based *Escherichia coli* whole-cell bioconversions. *Appl. Microbiol. Biotechnol.* 97, 7741–7754.
- Sagara, Y., Wada, A., Takata, Y., Waterman, M.R., Sekimizu, K., Horiuchi, T., 1993. Direct expression of adrenodoxin reductase in *Escherichia coli* and the functional characterization. *Biol. Pharm. Bull.* 16, 627–630.
- Sakaki, T., Sugimoto, H., Hayashi, K., Yasuda, K., Munetsuna, E., Kamakura, M., Ikushiro, S., Shiro, Y., 2011. Bioconversion of vitamin D to its active form by bacterial or mammalian cytochrome P450: *Biochim. Biophys. Acta BBA Proteins Proteomics* 1814, 249–256.
- Sanner, M.F., 1999. Python: a programming language for software integration and development. *J. Mol. Graph. Model.* 17, 57–61.
- Schiffler, B., Bureik, M., Reinle, W., Müller, E.-C., Hannemann, F., Bernhardt, R., 2004. The adrenodoxin-like ferredoxin of *Schizosaccharomyces pombe* mitochondria. *J. Inorg. Biochem.* 98, 1229–1237.
- Schmitz, D., Zapp, J., Bernhardt, R., 2012. Hydroxylation of the triterpenoid dipterocarpol with CYP106A2 from *Bacillus megaterium*. *FEBS J.* 279, 1663–1674.
- Uhlmann, H., Beckert, V., Schwarz, D., Bernhardt, R., 1992. Expression of bovine adrenodoxin in *Escherichia coli* and site-directed mutagenesis of (2-Fe-2s) cluster ligands. *Biochem. Biophys. Res. Commun.* 188, 1131–1138.
- Whitehouse, C.J.C., Bell, S.G., Wong, L.-L., 2012. P450 BM3 (CYP102A1): connecting the dots. *Chem. Soc. Rev.* 41, 1218–1260.

2.6 (Milhim et al., 2016)

Supplemental information

Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin.

Milhim, M., Putkaradze, N., Abdilmughni, A., Kern, F., **Hartz, P.** and Bernhardt, R.

Journal of Biotechnology. 2016 Nov; 240:68-75

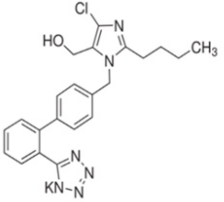
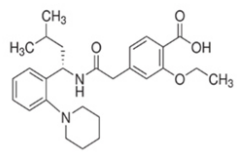
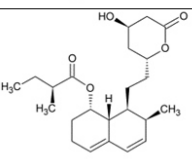
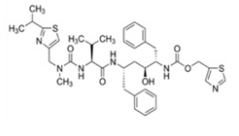
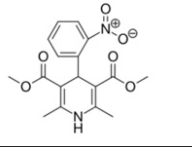
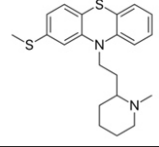
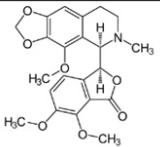
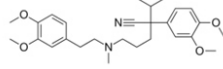
DOI: 10.1016/j.jbiotec.2016.11.002

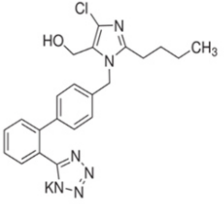
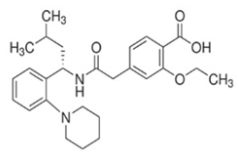
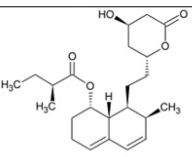
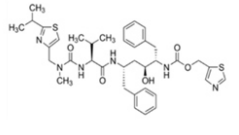
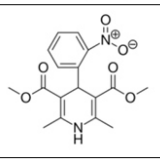
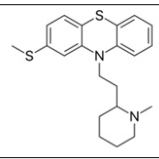
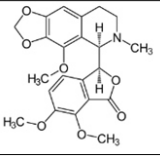
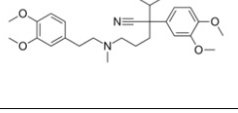
Reprinted with permission of Journal of Biotechnology. All rights reserved.

Supplementary Table 1. Protein sequence identity values of CYP107DY1 with different cytochrome P450s.

Closest match with different CYP107					
Protein Abb.	Natural function or substrate/s	No. of aa	Species	Identity% ^(a)	UniProtKB accession numbers. ^(b)
CYP107DY1	---	410	<i>B. megaterium</i> QM B1551	100	D5E3H2
CYP267B1	Drug metabolism	405	<i>Sorangium cellulosum</i> So ce5	43.5	A9ERX9
CYP107B1	7-ethoxycoumarin	405	<i>Saccharopolyspora erythraea</i>	41.9	P33271
CYP107K1 (PksS)	Polyketide biosynthesis (bacillaene biosynthesis)	405	<i>Bacillus subtilis</i> (Strain 168)	41.8	O31785
CYP107H1 (bioI)	Biotin biosynthesis	395	<i>Bacillus subtilis</i> (Strain 168)	40.8	P53554
CYP107J1	Testosterone enanthate	410	<i>Bacillus subtilis</i> (Strain 168)	40.7	O08469
CYP107BR1 (Vdh)	Vitamin D3	403	<i>Pseudonocardia autotrophica</i>	40	C4B644
CYP107L1 (PiKC)	Pikromycin biosynthesis	416	<i>Streptomyces venezuelae</i>	39.2	O87605
CYP107A1 (eryF)	Erythromycin biosynthesis	404	<i>Saccharopolyspora erythraea</i>	36.2	Q00441
CYP107E4	Diclofenac	396	<i>Actinoplanes sp.ATCC 53771</i>	36	C0LR90
CYP109B1	Various substrates (e.g. (+)-valencene)	396	<i>Bacillus subtilis</i>	35.4	U5U1Z3
CYP267A1	Drugs metabolism	429	<i>Sorangium cellulosum</i> So ce5	35.3	A9EN90
CYP109A2	Steroids	403	<i>B. megaterium</i> DSM319	34.6	D5DF88
CYP109E1	Steroids	404	<i>B. megaterium</i> DSM319	32.7	D5DKI8
CYP106A1	Steroids and terpenoids	410	<i>B. megaterium</i> DSM319	29.1	D5DF35
CYP108	α -terpinol	428	<i>Pseudomonas sp.</i>	28	P33006
CYP102A1	Fatty acids	1,049	<i>B. megaterium</i> DSM319	18.4	P14779

(a) Based on Clustal Omega Multiple Sequence Alignment (MSA) (<http://www.ebi.ac.uk/Tools/msa/>)(b) <http://www.uniprot.org/>**Supplementary Table 2.** List of tested substrates with CYP107DY1

Losartan		Repaglinide	
Mevastatin		Ritonavir	
Nifedipine		Thioridazine	
Noscapine		Verapamil	

Losartan		Repaglinide	
Mevastatin		Ritonavir	
Nifedipine		Thioridazine	
Noscapine		Verapamil	

```

1  ATGAAAAGGTTACAGTTGATGATTTAGCTCTCCAGAAAATATGCACGATGTCATCGGATTTTATAAAAACTCACTGAACATCAAGAACCTCTTATTCGTTTG
1  M K K V T V D D F S S P E N M H D V I G F Y K K L T E H Q E P L I R L

106  GATGATTATTACGGGTTGGGACCGGCATGGGTCGCATTACGTCATGACGATGTTGTTACGACTAAAGAACCCCGTTTTCTCAAAGATGTACGGAAGTTCACA
36  D D Y Y G L G P A W V A L R H D D V V T I L K N P R F L K D V R K F T

211  CCATTGCAAGATAAAAAGGATTTCTATAGATGATAGCACATCTGCGAGCAAACCTGTTTGAATGGATGATGAATATGCCGAATATGCTTACGGTCGATCCACCCGAT
71  P L Q D K K D S I D D S T S A S K L F E W M M N M P N M L T V D P P D

316  CACACTCGTTTGCAGGTTGGCCTCTAAAGCCTTTACGCCACGTATGATCGAGAATCTTCGACCTCGTATACAGCAGATTACCAATGAGCTATTGGATTAGTA
106  H T R L R R L A S K A F T P R M I E N L R P R I Q Q I T N E L L D S V

421  GAAGGAAAAGGAATATGGATCTTGTTCGGATTTTCTTTTCTCTGCCCATTATTGTTCATTTAGAGATGCTAGGGATTCCACCTTTAGATCAGAAAATGCTG
141  E G K R N M D L V A D F S F P L P I I V I S E M L G I P P L D Q K R F

526  CGCGACTGGACAGATAAACTCATCAAGCAGCTATGGATCCTAGCCAAGGGGCTGTAGTTATGGAACACTCAAGGAGTTTATTGATTACATCAAAAAATGCTG
176  R D W T D K L I K A A M D P S Q G A V V M E T L K E F I D Y I K K M L

631  GTCGAAAAGCGCAACCATCCAGACGATGATGTGATGAGTGTCTTGTGCAAGCACATGAGCAAGAAGATAAGTTGAGCGAGAACGAGCTTCTTTCCACGATTGG
211  V E K R N H P D D D V M S A L L Q A H E Q E D K L S E N E L L S T I W

736  CTACTCATTACAGCCGACATGAGACGACGGCCCATCTAATCAGCAACGGGCTACTGGCGCTATTGAAGCATCCGAAACAAATGCGCCTGCTTCGGGATAATCCT
246  L L I T A G H E T T A H L I S N G V L A L L K H P E Q M R L L R D N P

841  TCTTTACTCCCCTCTGCCGTTGAAGAGCTGTACGCTATGCCGACCGGTCATGATGGTGGGCGTTTTGCGGGTGAAGATATCATGATGCAATGAAAAATGATT
281  S L L P S A V E E L L R Y A G P V M I G G R F A G E D I I M H G K M I

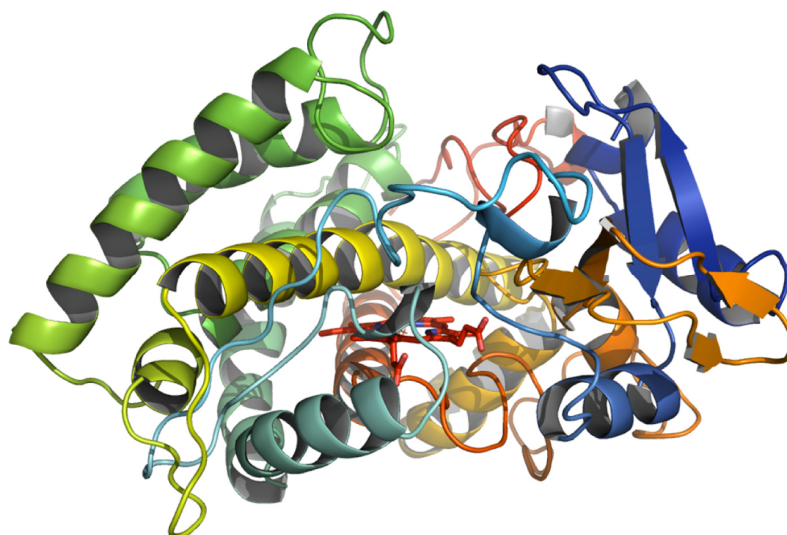
946  CCCAAGGTGAAATGGTGTCTGCTGTTCTGCGTGGTCCGCAATATTGATTCACAGAAATCTCTTATCCTGAGGATGGATATTACACGCGAGGAGAAATGAGCAT
316  P K G E M V L F S L V A A N I D S Q K F S Y P E G L D I T R E E N E H

1051  CTCACCTTCGGAAGGATCCATCATTTGTTGGGAGCGCCTTTGGCGCGCATGGAAGCACATATCGCTTTCGGCACATTGCTTCAACGGTTTCTGATTTACGA
351  L T F G K G I H H C L G A P L A R M E A H I A F G T L L Q R F P D L R

1156  TTGGCAATCGAATCGGAGCAACTGGTTTATAACAACAGCACATTGCGTTCTCTTAAAAGCTTGCCAGTTATTTCTAA
386  L A I E S E Q L V Y N N S T L R S L K S L P V I F *

```

Supplementary Fig. S1. The open reading frame sequence of the CYP107DY1. The upper and the lower lines represent the nucleotide and deduced amino acid sequences, respectively. The one-letter code for each amino acid is aligned with the first nucleotide of each codon. Several conserved motifs used for the identification of cytochrome P450s are underlined. I-helix (A/G-G-x-E/D-T-T/S), K-helix (E-x-x-R), and the heme pocket (F-x-x-G-x-x-x-C-x-G).



Supplementary Fig. S2. Homology model of CYP107DY1. A model of CYP107DY1 was calculated using CYP107RB1 (Vdh) (PDB accession code: 3A4G) as template. The coordinates of the heme-porphyrin atoms from the template structure were added subsequently to the obtained homology model. Program Modeller 9.14 (University of California San Francisco, USA) was used.

3. General Discussion

3.1 Identification and characterization of novel promoter systems in *B. megaterium*

Genome-wide differential gene expression analyses of the *B. megaterium* strain MS941 revealed growth phase dependent transcription patterns for many open reading frames (ORFs) (see publication 2.1, Figure 1). Throughout the evaluation process of the generated microarray data it became obvious that most ORFs showed significantly changed transcription levels during the late exponential and the early stationary phases of bacterial growth. Contrarily, the early and mid-exponential growth phases were rather associated with some of the highest signal intensities. For this reason, the corresponding putative promoter elements were categorized into the group of “fold change promoters”, representing the genes with the most significantly changed transcription levels (see publication 2.1, Figure 2A) and the group of “signal intensity promoters” which showed the highest transcription levels (see publication 2.1, Figure 2B). Strikingly, some of the ORFs with the highest transcription levels were also reported to be found among the most abundant proteins of *B. megaterium* (Wang et al., 2005), thus further emphasizing the quality of the generated microarray data. Furthermore, novel sugar-inducible promoter elements were identified based on the architecture of the xylose promoter-repressor system (Rygus et al., 1991) (see publication 2.1, Figure 5A). A β -galactosidase-deficient *B. megaterium* strain was generated (see publication 2.1, Figure 3) and the putative promoter elements were successfully characterized via β -galactosidase (LacZ) screening (see publication 2.1, Figure 4 and 5B) with particular focus on the comparison of the different promoter strengths with that of the extensively used and already optimized xylose inducible promoter system (pXyl*) of *B. megaterium* (Rygus and Hillen, 1991; Stammen et al., 2010). As a result, a diverse set of 19 promoters, comprising constitutive, growth phase dependent as well as inducible promoters, was characterized with promoter strengths ranging from 15% to 145% compared to that of the reference promoter (summarized in Table 3.1).

Table 3.1. Overview of the main characteristics of the novel promoter library.

Promoter	Class	Strength (%) ^a
p0123	early exponential	14.9
p4756	early exponential	75.0
p0706	early exponential	91.5
p0551	mid-exponential	< 1.0
p0450	mid-exponential	15.1
p2097	mid-exponential	16.8
p2618	late exponential	14.8
p4632	late exponential	93.9
p2386	late exponential	108.3
p3537	late exponential	145.6
p2948	early stationary	< 1.0
p0462	early stationary	< 1.0
p1081	early stationary	15.0
p3115	constitutive	20.7
p1214	constitutive	34.5
pSuc	not inducible	< 1.0
pGal	inducible	23.9
pBga	inducible	34.9
pAra	inducible	75.4
pXyl*	inducible	100.0

^a compared to the LacZ activity of the optimized xylose inducible promoter system (pXyl*)

The best promoter p3537 was additionally shown to be responsive to iron which ultimately resulted in a nearly 45% increase in LacZ expression levels under optimized cultivation conditions compared with the optimized xylose inducible promoter (see publication 2.1, Supplemental Figure 2B). Furthermore, promoter p4632 was characterized as the first heat inducible promoter in *B. megaterium*, showing 2.6 fold increased LacZ expression levels at elevated cultivation temperatures that almost matched those of the reference promoter (see publication 2.1, Supplemental Figure 2C). The LacZ expression levels correlated remarkably with the corresponding LacZ activities as shown in SDS-PAGE analyses (see publication 2.1, Supplemental Figure 1). To the best of our knowledge, this is the first and most comprehensive study for the identification of novel promoter elements for *B. megaterium*, especially with regard to growth phase dependent promoters. For this reason there are no comparable data available in *B. megaterium*. However, based on the increasing number of recent publications that deal with the identification and characterization of growth phase dependent promoters in general, they seem to have considerable importance for biotechnological applications in different microorganisms, not least because of the fact that they are independent of the addition of expensive or toxic inducing agents that possibly impair their usage in industrially scaled processes (Hofsten, 1961; Nocadello and Swennen, 2012; Wilson et al., 1981). Similar promoter screening approaches for *Bacillus subtilis* recently provided the basis for the identification of the highly active stationary phase promoters Pylb and the constitutive promoters PsodA as well as PyszA. Their application was demonstrated to significantly improve the recombinant expression of pullulanase, a widely used enzyme with application in the starch processing industries (Hii et al., 2012; Liu et al., 2018; Meng et al., 2018; Yu et al., 2015). Unfortunately, most of these studies are merely focused on the identification of the strongest promoters for the recombinant expression of single proteins. However, excessive and fixed protein expression under the control of the same promoter does not meet the requirements of modern rational strain design, since it is nowadays often associated with increasing reaction network complexity and involves multi enzyme cascades that necessarily rely on well defined and harmonized protein expression levels (Jones et al., 2015; Pitera et al., 2007; Xu et al., 2014). This aspect highlights the great potential of the expanded promoter toolbox with its diverse promoter activities as identified and evaluated here for *B. megaterium*. Particularly, the newly identified class of growth phase dependent promoters is virtually predestined to be exploited for the rational strain design of *B. megaterium* allowing the simultaneous use of multiple promoters for both the temporal and the quantitative control of individual protein expression levels within the complex crosstalk of metabolic networks. However, the simultaneous application of multiple sugar-inducible promoters is rather limited for a similar fine tuning of protein expression levels in multi enzyme cascades due to their mutual susceptibility to carbon catabolite repression (CCR) (Giacalone et al., 2006; Miyada et al., 1984; Rygus and Hillen, 1992). Unsurprisingly, a corresponding effect was also observed for the non optimized (pXyl) and optimized xylose inducible promoter (pXyl*) as well as for the novel arabinose inducible promoter systems (pAra) in *B. megaterium* (listed in Table 3.2).

Table 3.2. Effect of different sugars on the carbon catabolite repression (%) of selected inducible promoter systems in *B. megaterium*.

Sugar ^a	pAra ^b	pXyl ^b	pXyl* ^b
(L)-Arabinose	0.0 ^c	48.9 ± 3.4	14.1 ± 12.1
(D)-Xylose	40.3 ± 11.6	0.0 ^c	0.0 ^c
(D)-Glucose	95.8 ± 0.5	91.8 ± 0.5	30.0 ± 12.4
(D)-Galactose	79.7 ± 4.8	5.2 ± 3.8	13.3 ± 4.2
(D)-Fructose	86.3 ± 1.2	73.1 ± 0.4	0.0
(D)-Sucrose	81.6 ± 2.6	14.1 ± 2.2	0.0

^a the final concentration of the carbon catabolite repressor was 3 % (w/v)

^b the final concentration of the inducing sugar was 0.5 % (w/v)

^c no carbon catabolite repressor was supplemented

In general, the carbon catabolite repression (CCR) profiles of the sugar-inducible promoter systems of *B. megaterium* differentiated significantly. While the optimized xylose inducible promoter system (pXyl*) was obviously insensitive to CCR except for glucose, the non optimized xylose promoter (pXyl) showed significant CCR for xylose, glucose and sucrose. However, the arabinose inducible promoter (pAra) seemed to be remarkably CCR susceptible to most of the tested sugars. Consequently the simultaneous application of the xylose and arabinose inducible promoter system in the same whole-cell biocatalyst will be associated with a significant decrease in recombinant protein expression levels and is therefore not recommended. Alternatively, the evaluation and engineering of a CCR insensitive *B. megaterium* strain would be a reasonable approach to avoid mutual repression of sugar-inducible promoter systems as described for other *Bacillus* species (Ludwig and Stülke, 2001; van der Voort et al., 2008). Intriguingly, glycerol did not negatively affect any of the tested promoters, conversely leading to higher β -galactosidase (LacZ) activities of up to 66.8% (see Table 3.3). This beneficial effect on recombinant protein expression in *B. megaterium* makes glycerol an ideal and cheap carbon source for industrial fermentation processes not only for the sugar-inducible promoter systems but also for the class of growth phase dependent promoters.

Table 3.3. Stimulatory effect of glycerol supplementation on the β -galactosidase activities (%) of selected inducible promoter systems in *B. megaterium*.

	pAra ^b	pXyl ^b	pXyl* ^b
- Glycerol	100	100	100
+ Glycerol ^a	114.9 ± 4.4	120.1 ± 4.0	166.8 ± 4.9

^a the final concentration of glycerol was 3 % (w/v)

^b the final concentration of the inducing sugar was 0.5 % (w/v)

The rational adaptation of protein expression levels using different promoters with versatile promoter strengths is considered to be particularly attractive for the optimization of cytochrome P450 driven reactions, such as epoxidation, dealkylation, C-C bond cleavage, but mainly hydroxylation of valuable compounds (Bernhardt and Urlacher, 2014). Cytochromes P450 belong to the group of heme containing monooxygenases and necessarily rely on auxiliary redox proteins which support the electron transport from an external coenzyme (NADPH) to the substrate (Bernhardt, 2006). *In vitro* studies with well defined ratios of CYP260A1 from *Sorangium cellulosum* to its surrogate redox partners adrenodoxin (Adx) and adrenodoxin reductase (AdR) demonstrated a significant influence not only on the activity but also on the product pattern of CYP260A1 (Khatri et al., 2017). Similarly, the engineering of the multi component redox chain of bovine CYP11B1 in an *E. coli* based whole-cell system was shown to accelerate as well as improve production of the common glucocorticoid cortisol, remarkably (Schiffer et al., 2015). Since the repertoire of useful promoter elements in *B. megaterium* has been limited to the xylose inducible system so far, the novel promoter toolbox will offer the innovative possibility to deliberately manipulate and harmonize ratios of cytochromes P450 and corresponding redox proteins *in vivo*, thereby not only modulating the product pattern but possibly also the efficiency of P450 mediated steroid conversions in *B. megaterium*.

3.2 Biotransformation of steroids in *B. megaterium* using new promoters

The promising potential of *B. megaterium* for the biotransformation of various steroids and steroid derived compounds has been shown in several studies before. The endogenous cytochromes P450 of *B. megaterium* as well as engineered variants thereof were successfully applied for the development of efficient whole-cell systems for the conversion of a diverse set of natural steroids such as cholesterol (Putkaradze et al., 2019), testosterone (Jóźwik et al., 2016; van Vugt-Lussenburg et al., 2006), progesterone (Lee et al., 2015; Zehentgruber et al., 2010) or dehydroepiandrosterone (Schmitz et al., 2014) as well as synthetic steroids like prednisone, dexamethasone (Putkaradze et al., 2017a) or cyproterone acetate (Kiss et al., 2015). Recently, *B. megaterium* was also shown to be a suitable host for the recombinant expression of membrane associated mammalian cytochromes P450 (Ehrhardt et al., 2016; Gerber et al., 2015). Particularly, the CYP11A1 mediated side-chain cleavage of cholesterol to pregnenolone, the initial reaction of the steroid hormone biosynthesis, was demonstrated to be extraordinarily efficient in *B. megaterium* compared with other microbial cell factories (Duport et al., 1998). Another important step towards the biosynthesis of steroid hormones comprises the conversion of pregnenolone to progesterone (see publication 2.1, Figure 6A), a key intermediate for the production of corticosteroids (see Figure 1.3). For this reason, we aimed to apply the novel promoter toolbox in *B. megaterium* for the heterologous expression of the cholesterol oxidase 2 (BCO2) from *Brevibacterium sterolicum*, which was previously shown to catalyze the desired conversion of Δ^5 -3 β -hydroxysteroids, including pregnenolone, to the corresponding steroids with Δ^4 -3-keto configuration, such as progesterone (Coulombe et al., 2001; Croteau and Vrieling, 1996). As shown in publication 2.1, Figure 6B-E, all tested promoters of a focused promoter library were successfully applied for the BCO2 mediated conversion of pregnenolone to progesterone in a genetically engineered *B. megaterium* strain with innate steroid hydroxylase as well as 20 α -hydroxysteroid dehydrogenase deficiencies (see publication 2.2, Table 1). Strikingly, the observed progesterone formation rates and yields accurately reflected the individual promoter strengths that were previously determined via β -galactosidase (LacZ) assays. Consequently, the initial pregnenolone conversion with the arabinose inducible promoter system (pAra) seemed to be decelerated compared to the reference promoter (pXyl*). The final progesterone yield of 3.5 mM/L/d, however, was identical to that of pXyl*. Similarly, the early exponential growth-phase dependent promoters p0706 and p4756 showed slower progesterone formation rates that ultimately resulted in slightly decreased progesterone yields of 3.3 mM/L/d and 2.9 mM/L/d, respectively. Since the late exponential growth-phase dependent promoter p3537 was screened as the strongest promoter according to the results of the LacZ assay, it was expected to be the most potent promoter for the BCO2 mediated production of progesterone in *B. megaterium*. Indeed, the final progesterone yield of 3.6 mM/L/d was found to be slightly improved compared to the optimized xylose inducible promoter system (pXyl*). In this context, it should be emphasized that, in contrast to the heavily engineered and optimized xylose inducible promoter system (pXyl*), the regulatory DNA sequences of all novel promoters are wildtype sequences. Hence, their

potential for the future optimization of recombinant protein production in *B. megaterium* seems to be more than promising, particularly considering the fact that pregnenolone conversions with the comparable wildtype xylose inducible promoter system (pXyl) barely resulted in higher progesterone yields than 1.4 mM/L/d (see publication 2.1, Figure 6F). The BCO2 mediated formation of progesterone in *B. megaterium* was further optimized using high cell density conversion approaches. The application of p3537 as the best promoter resulted in significantly accelerated progesterone formation rates that allowed a nearly total biotransformation of 5 mM pregnenolone yielding approximately 1.55 g/L progesterone within the first 2 h of conversion with resting cells (data not shown). To the best of our knowledge, the cholesterol oxidase mediated conversion of Δ^5 -3 β -hydroxysteroids has not been carried out with recombinant whole-cell biocatalysts, yet. Conventional biocatalysis of the corresponding Δ^4 -3-keto steroids was either performed with extensively purified cholesterol oxidase enzymes (Alexander and Fisher, 1995; Labaree et al., 1997) or natural microbial strains which both resulted in low product titers of barely more than 0.2 g/L or complex product mixtures (Ahmad et al., 1991). Consequently, the established *B. megaterium* based whole-cell biocatalyst offers an efficient and innovative platform with nearly 8 fold increased Δ^4 -3-keto steroid formation rates compared to other systems.

Since localized at the crossroad of corticosteroid biosynthesis (see Figure 1.3), progesterone can either be directed towards the biosynthesis of mineralocorticoids by the CYP21A1 mediated formation of 21 α -hydroxyprogesterone (deoxycorticosterone; DOC) or transformed by CYP17A1 to 17 α -hydroxyprogesterone (17 α OH-P) which is considered an important intermediate for the biosynthesis of pharmaceutically valuable glucocorticoids, like cortisol (Gilep et al., 2011). While DOC was demonstrated not being metabolized in *B. megaterium* (see publication 2.1, Supplemental Figure 3), 17 α OH-P was evidently further converted to several unwanted side products (see publication 2.2, Figure 3). The major side product was characterized as 17 α ,20 α -dihydroxyprogesterone (17 α ,20 α DiOH-P) whereas its minor side product was identified as 17 α ,20 β -dihydroxyprogesterone (17 α ,20 β DiOH-P). Bioinformatic analyses were applied to identify putative open reading frames (ORFs) with hydroxysteroid dehydrogenase (HSD) activity. Upon overexpression, a corresponding 20 α -HSD activity was confirmed for 4 of the screened ORFs (see publication 2.2, Figure 4B). The stepwise deletion of all 20 α -HSD genes ultimately enabled the development of the *B. megaterium* strain GHH8 with completely abolished 20 α -HSD activity (see publication 2.2, Figure 4D). Among the other ORFs with potential HSD activity, the 3-oxoacyl-(acyl-carrier-protein) reductase (FabG) was identified accordingly to be associated with the residual 20 β -HSD activity (see publication 2.2, Figure 5A). Unfortunately, FabG is an essential component during the bacterial fatty acid synthesis (FAS) and therefore not suitable for the generation of a deletion mutant with reduced 20 β -HSD activity (Zhang and Cronan, 1998). Compounding this problem is the fact that further transformation of 17 α OH-P, such as the CYP21A1 catalyzed hydroxylation to 17 α ,21-dihydroxyprogesterone (11-deoxycortisol, RSS) or the CYP11B1 mediated hydroxylation of RSS to cortisol seems to increase the

substrate affinity of $17\alpha\text{OH-P}$ derived steroids to FabG (unpublished data), thereby facilitating the formation of pharmaceutically unfavorable 20β -hydroxysteroids in *B. megaterium*. Intriguingly, the supplementation of acetone during steroid conversions seemed to have an inhibitory effect on the 20β -HSD activity of FabG, though accumulation of 20β -hydroxysteroids was not completely prevented (see publication 2.3, Supplemental Figure 1). As substrate for the recombinantly expressed alcohol dehydrogenase of *Lactobacillus brevis* (LbADH), acetone was originally intended to drive a corresponding NADP⁺ regeneration system in *B. megaterium* (see publication 2.3, Figure 12) which was determined as crucial factor for the efficient oxidation of the pharmaceutically valuable cortisol to cortisone (see publication 2.3, Figure 11). However, the increased pool of NADP⁺ also contributed significantly to the reduced 20β -HSD activity of FabG which, itself, is dependent on NADPH.

More important than the reduced side product formation, the productivity of the developed *B. megaterium* based whole-cell system for the 11β -hydroxysteroid dehydrogenase (11β -HSD) mediated conversion of cortisol to cortisone was improved considerably by applying the novel arabinose inducible promoter system (pAra) and the growth phase dependent promoter p0706 for the recombinant expression of a modified type 1 guinea pig 11β -HSD (Lawson et al., 2009). Compared to the optimized xylose inducible promoter (pXyl*) which already showed a high volumetric cortisone production of approximately 6 g/L/d under optimized cultivation conditions, cortisone formation rates were further increased 1.3 fold with pAra and 2.2 fold with p0706 resulting in extraordinary space-time yields of nearly 8 g/L/d and 14 g/L/d cortisone, respectively (see publication 2.3, Figure 11). To the best of our knowledge, the previously highest reported yield of 9 g/L/d for a similar microbial glucocorticoid conversion of cortisone to cortisol was achieved with an *E. coli* based whole-cell biocatalyst (Zhang et al., 2014). Consequently, the application of the novel promoter systems in *B. megaterium* lead to the accumulation of 1.6 fold higher glucocorticoid levels compared to the formerly best microbial cell factory of Zhang and coworkers. Besides the improved cholesterol oxidase (BCO2) mediated conversion of pregnenolone to progesterone, the significantly enhanced activity of the 11β -HSD is yet another striking example for the potential of the novel promoter toolbox as key component for the development and optimization of *B. megaterium* as efficient and versatile platform for protein production and steroid transformations.

The ability of *B. megaterium* for the biotransformation of pharmaceutically important steroids can be summarized as follows. Although the novel promoters were successfully applied for the heterologous expression of a diverse set of enzymes from different enzyme classes (including the BCO2 from *B. sterolicum*, the guinea pig 11β -HSD as well as several mammalian steroidogenic cytochromes P450 (unpublished data)) the residual 20β -HSD activity of FabG has mainly impaired the application of *B. megaterium* as efficient microbial cell factory for the production of cortisol and cortisone, so far. Consequently, the elimination or minimization of this 20β -HSD activity remains the major challenge

for future engineering approaches to develop *B. megaterium* into an ideal microbial platform for steroid biotransformations.

3.3 Identification of a novel short chain dehydrogenase for the production of the sesquiterpene (+)-nootkatone in *B. megaterium*

In the course of the studies for the identification of putative hydroxysteroid dehydrogenases (HSDs) in *B. megaterium* we found a promising open reading frame (BMD_2094) which was not only able to act as 17 β -HSD on the sex hormone testosterone to produce androstenedione (unpublished data) but more importantly to catalyze the efficient oxidation of (trans)-nootkatol to the industrially valuable flavor and fragrance compound (+)-nootkatone (see publication 2.4, Figure 1). Commercial (+)-nootkatone production involves a two-step biotransformation process comprising the regioselective hydroxylation of (+)-valencene to (trans)-nootkatol which is then further converted to (+)-nootkatone (Fraatz et al., 2009) (see publication 2.4, Scheme 1). While the first catalytic step of (+)-nootkatone formation has been elucidated as a cytochrome P450 mediated reaction (Cankar et al., 2011; Gavira et al., 2013; Girhard et al., 2009; Schulz et al., 2015; Sowden et al., 2005), only scarce information is known about suitable enzymes for the final oxidation step which is believed to involve a dehydrogenase (Wriessnegger et al., 2014). To the best of our knowledge, the novel short chain dehydrogenase (SDR) BMD_2094 is the first bacterial enzyme with activity towards the conversion of (trans)-nootkatol to (+)-nootkatone. As demonstrated in publication 2.4, Figure 2, the recombinant expression of BMD_2094 under the control of the optimized xylose inducible promoter (pXyl*) resulted in a very efficient *B. megaterium* based whole-cell system with (+)-nootkatone formation rates of 44 mg/L/h. While other microbial cell factories were reported to yield approximately 0.35 mg/L/h (+)-nootkatone in *Pichia pastoris* (Wriessnegger et al., 2014) and 8.06 mg/L/h (+)-nootkatone in *Yarrowia lipolytica* (Palmerín-Carreño et al., 2016), the corresponding space-time yields for (+)-nootkatone production in *B. megaterium* were shown to be over 120 fold and 5 fold higher than those of *P. pastoris* and *Y. lipolytica*, respectively. This extraordinary productivity strikingly illustrates the promising potential of the novel dehydrogenase BMD_2094 as well as the potential of *B. megaterium* for the development of efficient microbial cell factories for the improved production of the industrially valuable sesquiterpene (+)-nootkatone.

3.4 Biosynthesis of C30 carotenoids in *B. megaterium*

Metabolic profiling of *B. megaterium* MS941 by LC-MS analyses revealed the presence of two carotenoid species with low abundance. According to the authentic standard library of our collaboration partner Metabolic Discoveries in Potsdam, the identified carotenoids showed identical m/z ratios and retention times to those of the C30 carotenoids 4,4'-diapophytofluene (DPFL) and its oxygenated derivative 4,4'-diaponeurosporenic acid (DNSA), respectively (see publication 2.5, Figure 3). These findings agreed with scientific reports from the early 1980s that describe the accumulation of two major pigments during spore formation of *B. megaterium* (Racine and Vary, 1980). Although the pigments were not further characterized, they intriguingly showed a carotenoid-like absorption spectrum with multiple maxima in the range of visible light (Mitchell et al., 1986). Particularly the absorption spectrum of the yellow pigment observed by Mitchell and coworkers was strikingly similar to the later characterized spectrum of 4,4'-diaponeurosporene (DNSP), one of several pigments in the staphyloxanthin biosynthetic pathway of the pathogen *Staphylococcus aureus* (Kim et al., 2016; Marshall and Wilmoth, 1981; Wieland et al., 1994). Meanwhile, various spore forming bacteria including the group of pigmented bacilli were shown to accumulate a diverse set of C30 carotenoids (Köcher et al., 2009; Perez-Fons et al., 2011; Steiger et al., 2012; Takaichi et al., 1997). While most of these carotenoid structures have successfully been elucidated along with their underlying biosynthetic pathways and involved enzymes, there was no information or experimental data about a cognate biosynthetic route for *B. megaterium*. Consequently, the published genome sequence of the parental *B. megaterium* strain DSM319 was used for the identification of a putative biosynthetic pathway that evidently led to the accumulation of DPFL and DNSA in vegetative cells of the *B. megaterium* strain MS941. In the course of these bioinformatic analyses, several open reading frames (ORFs) were identified with significant sequence identities to the deduced amino acid sequences of the corresponding enzymes in the extensively studied staphyloxanthin biosynthetic pathway of *S. aureus* (Pelz et al., 2005) (see publication 2.5, Table 3). Intriguingly, some of them seemed to be coordinated as a biosynthetic operon (see publication 2.5, Figure 4). A functional characterization of the individual enzyme activities in the non carotenogenic model organism *Escherichia coli* partially revealed unique features of the putative C30 carotenoid biosynthesis operon of *B. megaterium* MS941.

In general, the number of carbon atoms in the carotenoid backbone as well as its structure is strongly dependent on the chain lengths of the prenyl diphosphates that are used for the initial condensation reaction during carotenoid biosynthesis (Norris et al., 1995). Compared to other pigmented bacilli, the biosynthesis of prenyl diphosphates in *B. megaterium* was demonstrated to be limited to the formation of farnesyl diphosphate (FDP) (see publication 2.5, Figure 2), thus providing the basis for the subsequent head-to-head condensation of two molecules of FDP (C15) to form the symmetric C30 carotenoid backbone structure of 4,4'-diapocarotenoids. As a result, the activity of the putative carotenoid synthase (BmCrtM) of *B. megaterium* exclusively yielded the symmetric carotenoid

precursor 4,4'-diapophytoene (DPHY) (see publication 2.5, Figure 6A). Contrarily, this condensation reaction was shown in other pigmented bacilli to preferentially take place between one molecule of geranyl diphosphate (C10) and one molecule of geranylgeranyl diphosphate (C20) yielding the asymmetric C30 carotenoid backbone structure of apo-8'-carotenoids (Perez-Fons et al., 2011). Furthermore, striking conceptual differences were observed between the desaturation process of DPHY in *B. megaterium*, which involved the coordinated action of multiple carotenoid desaturases (BmCrtN₁₋₂) and the desaturation process of the corresponding apocarotenoids in *S. aureus* as well as other pigmented bacilli, where this desaturation reaction is catalyzed by a single carotenoid desaturase (Pelz et al., 2005; Steiger et al., 2015) (summarized in Figure 3.1). The carotenoid desaturase BmCrtN₁ of *B. megaterium* was characterized to perform the insertion of a single double bond into the backbone of DPHY which resulted in the formation of 4,4'-diapophytofluene (DPFL) (see publication 2.5, Figure 6B). In the course of these studies, the activity of BmCrtN₁ was shown to be essential to provide adequate amounts of DPFL for the subsequent desaturation steps catalyzed by the carotenoid desaturase BmCrtN₂ (see publication 2.5, Figure 6C). This second carotenoid desaturase of *B. megaterium* showed extraordinary properties in terms of substrate specificity (with DPFL being the preferred substrate), desaturation activity (by catalyzing a defined two-step desaturation reaction) and product specificity (with 4,4'-diaponeurosporene (DNSP) being the exclusive desaturation product).

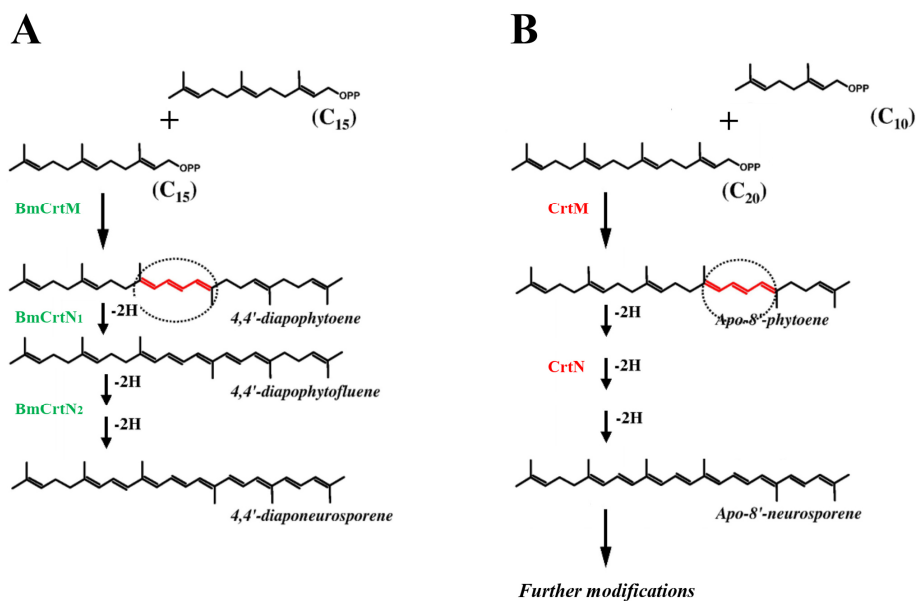


Figure 3.1. Biosynthetic routes of C30 carotenoid formation. A: Formation of symmetric 4,4'-diapocarotenoids in *B. megaterium*. B: Formation of asymmetric apo-8'-carotenoids in other pigmented bacilli. Adopted and modified from (Perez-Fons et al., 2011).

To the best of our knowledge these unique features, particularly the product specificity, have neither been described for the carotenoid biosynthesis in *S. aureus* nor in other pigmented bacilli, where DNSP is further converted to the completely desaturated carotenoid structure of 4,4'-diapolycopene (DLPN) (Yoshida et al., 2009) (see publication 2.5, Scheme 1). A third putative carotenoid desaturase BmCrtN₃ was found to be located far apart from the original DNSP biosynthesis operon (see publication 2.5, Figure 4C). However, due to a considerably higher sequence identity to the carotenoid oxygenases (CrtP) of *S. aureus*, this ORF was rather supposed to act as the missing carotenoid oxygenase towards the biosynthesis of 4,4'-diaponeurosporenic acid (DNSA), one of the two carotenoid species that were detected during the initial metabolite profiling of *B. megaterium* MS941. Despite various attempts to characterize BmCrtP as carotenoid oxygenase or carotenoid desaturase, BmCrtP ultimately seemed not to be involved in the process of C30 carotenoid biosynthesis (see publication 2.5, Figure 6D).

Nevertheless, there is a reasonable possibility that other, yet to identify, carotenoid oxygenases apart from CrtN₃ are responsible for the terminal oxygenation of DNSP to form DNSA in *B. megaterium*. The most promising candidates certainly belong to the group of cytochrome P450 monooxygenases, since multiple studies have demonstrated their potential to act as carotenoid hydroxylases in various organisms including plants, fungi and bacteria (Alvarez et al., 2006; Blasco et al., 2004; Fiore et al., 2006). *B. megaterium* also possesses several cytochromes P450, whose substrate ranges were subject to intensive studies that were predominantly focused on steroidal compounds. Just recently, CYP109E1 from *B. megaterium* was identified as efficient hydroxylase for the conversion of different carotenoid derived aroma compounds (Putkaradze et al., 2017b). Furthermore, CYP102A1, probably better known as BM3, was demonstrated to have activity towards the linear hydrocarbon skeleton of fatty acids, which resembles the linear backbone structure of acyclic carotenoids like DNSP (Whitehouse et al., 2012). Based on these encouraging findings, both CYP109E1 and CYP102A1 could become priority targets with regard to a future identification as well as characterization of the still obscure carotenoid oxygenase activity in *B. megaterium*.

DNSP itself has recently drawn considerable attention as promising pharmaceutical compound due to its stimulatory effects on the murine immune system (Liu et al., 2016). Several studies convincingly demonstrated an anti-inflammatory effect on chronic gastrointestinal diseases as well as a reduced susceptibility to *Salmonella* infections (Jing et al., 2019, 2017; Liu et al., 2017). However, there are no suitable microbial cell factories for the efficient and selective biosynthesis of DNSP, since natural producer strains, like *S. aureus*, are either pathogenic or associated with considerable side product formation (Archer, 1998; Chae et al., 2010). A major amount of DNSP in these strains is converted by a broad set of carotenoid modifying enzymes, including desaturases, oxygenases or glycosylases to the corresponding DNSP derivatives like 4,4'-diapolycopene, 4,4'-diapolycopenal or staphyloxanthin, respectively (Perez-Fons et al., 2011; Steiger et al., 2015; Wieland et al., 1994). Until now, the effects

of the aforementioned DNSP modifications have not been tested, hence it is not clear whether they have a comparable positive effect on the immune system or whether they are associated with adverse effects. At least staphyloxanthin was hypothesized to promote immune evasion, thereby significantly contributing to the pathogenic phenotype of *S. aureus* (Clauditz et al., 2006). Consequently, the selective production of DNSP in a microbial cell factory with GRAS (generally recognized as safe) status is considered a crucial prerequisite for the development of biotechnological processes that are compliant with the strict regulations of the pharmaceutical industry. In order to address the selective production of DNSP, the unique features of the newly identified C30 carotenoid biosynthetic cluster of *B. megaterium* were exploited to establish and assess the potential of an *Escherichia coli* based whole-cell system (see publication 2.5, Table 4). The engineered carotenoid gene cluster additionally contained the farnesyl diphosphate synthase (BmFDS) of *B. megaterium* to provide adequate amounts of FDP for the DNSP biosynthesis. This approach not only resulted in unrivaled product selectivity but also in significantly higher DNSP yields, thus clearly demonstrating that the application of the engineered DNSP gene cluster from *B. megaterium* provides not only a competitive but also a superior alternative for the efficient production of DNSP in comparison to other *E. coli* based microbial cell factories (summarized in Table 3.4). Despite significant longer production times of up to 78 h, most of them showed substantially lower specific C30 carotenoid yields ranging between 12 $\mu\text{g/g}$ wet cell weight and 49.4 $\mu\text{g/g}$ wet cell weight (Chae et al., 2010; Furubayashi et al., 2014; Xue et al., 2015). Only the heterologous expression of the C30 carotenoid gene cluster from *S. aureus* in genetically engineered *E. coli* K12 strains, optimized for the consumption of sucrose, was reported to produce higher specific C30 carotenoid yields of approximately 375 $\mu\text{g/g}$ wet cell weight (Kim et al., 2013). However the utilization of high sucrose concentrations seemed to prevent high cell density fermentation, thus limiting the total C30 carotenoid yield to 0.7 mg/L compared to approximately 4 mg/L obtained with the engineered carotenoid operon derived from *B. megaterium*. These results further emphasize the potential of the developed *E. coli* whole-cell system for a high yield DNSP production, most importantly without any side products, which will significantly facilitate downstream processing in future industrially scaled processes.

Table 3.4. C30 carotenoid yields obtained with different microbial cell factories.

Source	Expression Strain	C30 yield (µg/g wcw)	C30 yield (mg/L)	Time	Reference
<i>B. megaterium</i> DSM319	<i>E. coli</i> TOP10	146.1 ^{a, b}	4.03 ^b	20 h	publication 2.5
<i>B. megaterium</i> DSM319	<i>B. megaterium</i> MS941	262.7 ^{a, b}	7.91 ^b	20 h	unpublished
<i>S. aureus</i>	<i>B. subtilis</i> 168	68 ^{a, b}	n.d.	24 h	(Xue et al., 2015)
<i>S. aureus</i>	<i>E. coli</i> XL1-Blue	28.0 ^{a, b}	n.d.	60 h	(Furubayashi et al., 2014)
<i>H. sapiens</i> ; <i>S. aureus</i>	<i>E. coli</i> XL1-Blue	12.0 ^{a, b}	n.d.	60 h	(Furubayashi et al., 2014)
<i>S. aureus</i>	<i>E. coli</i> SURE	49.4 ^c	n.c.	48 h	(Chae et al., 2010)
<i>S. aureus</i>	<i>E. coli</i> XL1-Blue	35.3 ^c	n.c.	48 h	(Chae et al., 2010)
<i>S. aureus</i>	<i>E. coli</i> TOP10	12.4 ^c	n.c.	48 h	(Chae et al., 2010)
<i>S. aureus</i>	<i>E. coli</i> K12 derivative	375.0 ^{a, c}	0.7 ^c	78 h	(Kim et al., 2013)

n.d. not determined

n.c. not comprehensible

^a wet cell weight was estimated according to (Bratbak and Dundas, 1984)

^b carotenoid yields refer to 4,4'-diaponeurosporene

^c carotenoid yields refer to the major product 4,4'-diapolycopene

Encouraged by the outstanding performance of the developed *E. coli* system, which allowed the production of DNSP with unchallenged selectivity and volumetric yields, we wanted to test the potential of *B. megaterium* for the production of C30 carotenoids. The homologous expression of the engineered carotenoid cluster from *B. megaterium* was initially driven by the optimized xylose inducible system and resulted in a total C30 carotenoid yield of 3.84 mg/L (data not shown). This yield was found to be comparable with the 4.03 mg/L DNSP obtained with the corresponding *E. coli* based system (see Table 3.4). Although considerable side product formation was expected due to the not identified and characterized carotenoid oxygenase (BmCrtP), HPLC analyses of the C30 carotenoid extracts from *B. megaterium* surprisingly revealed DNSP as the exclusive carotenoid product. The retention time as well as the absorption spectrum of the isolated C30 carotenoid was identical to those, previously reported for the DNSP extracts from *E. coli* (see Figure 3.2). Obviously, the expression of the unknown carotenoid oxygenase in *B. megaterium* was not induced under the tested growth conditions. One possible reason for the absence of the carotenoid oxygenase activity may be related to the observation that carotenoid modification processes, particularly oxygenations,

seem to be strongly associated with spore formation as in other pigmented *Bacillus* species (Duc et al., 2006). Since long term cultivation of *B. megaterium* cells in complex media was not reported to induce sporogenesis, side product formation during DNSP biosynthesis is obviously negligible and can be ignored.

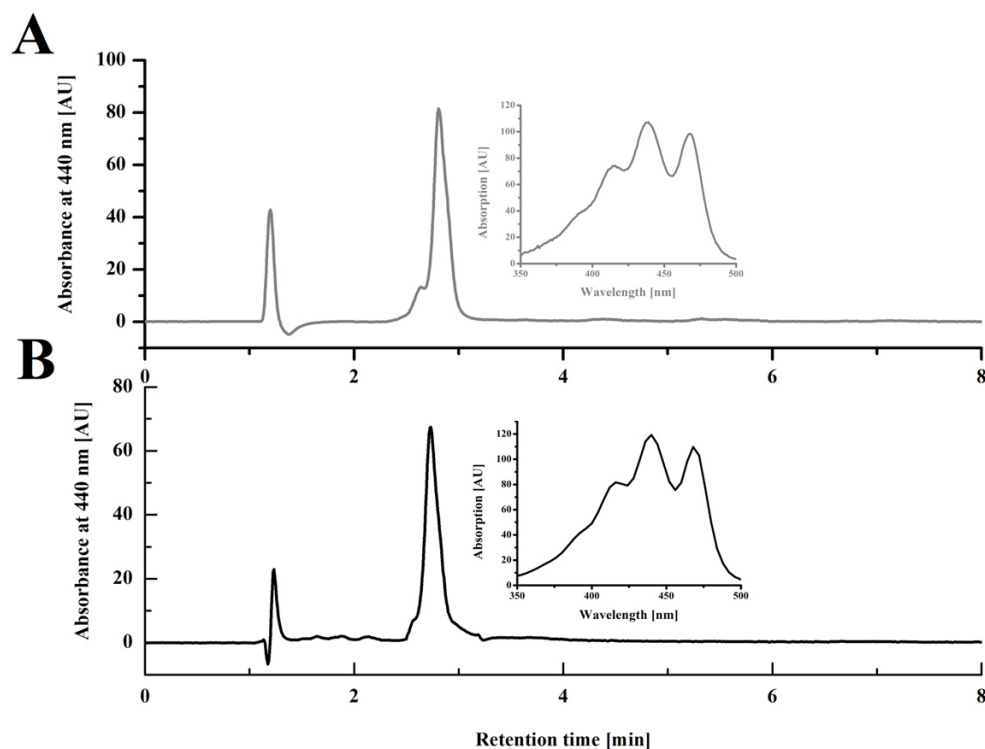


Figure 3.2. RP-HPLC profiles of C30 carotenoids recorded at the indicated wavelength. **A:** authentic DNSP extract of *E. coli*. **B:** C30 carotenoid extract of *B. megaterium* cells. Insets show the absorption spectra of the corresponding C30 carotenoids and were recorded with a photodiode array (PDA) detector.

The unexpected product selectivity of the *B. megaterium* based whole-cell system was a key argument for additional optimization efforts towards a more efficient DNSP biosynthesis. For this reason, a focused promoter library was tested to increase the DNSP yield in *B. megaterium* (unpublished data). This promoter library comprised the best novel promoters for protein expression in *B. megaterium* (see publication 2.1). According to the obtained results, the utilization of the new arabinose inducible promoter system led to the accumulation of 7.91 mg/L DNSP, which is more than a 2 fold increase in carotenoid accumulation levels compared to the optimized xylose inducible system. The total DNSP yield was even nearly 2 fold higher than that of the formerly best *E. coli* based whole-cell system (see Table 3.4). Consequently, the *B. megaterium* based system developed within this work was proven to be the most promising microbial cell factory for the selective biotechnological production of the pharmaceutically important C30 carotenoid DNSP, so far.

3.5 Biosynthesis of C40 carotenoids in *B. megaterium*

Although the C30 carotenoid 4,4'-diaponeurosporene (DNSP) has the potential to become a pharmaceutically valuable compound for the treatment of severe infectious as well as inflammatory diseases, the commercial interest and demand for the production of C40 carotenoids is still considerably higher. The most demanded C40 carotenoids include commonly known representatives like lycopene, β -carotene and oxygenated derivatives thereof such as lutein, zeaxanthin, canthaxanthin or astaxanthin, the carotenoid with the most antioxidant potential (Ernst, 2002; Torregrosa-Crespo et al., 2018). They find versatile application as colorants in the food and feed industry, as additives in cosmetics and as dietary supplements with beneficial effect for the human health, a property that was eponymous for the neologism “nutraceutical” (Anunciato and da Rocha Filho, 2012). Due to the growing demand for natural C40 carotenoids, they are projected to collectively contribute to an annual global market value of 2.0 billion US dollar by 2022 and are consequently products of high interest for the biotechnological industry (BBC Research).

As demonstrated in publication 2.5, Figure 2, *B. megaterium* was evidently incapable of initializing C40 carotenoid formation, since the corresponding biosynthetic pathway necessarily relies on geranylgeranyl diphosphate (GGDP) as substrate for the initial condensation reaction, but biosynthesis of the required prenyl diphosphates is only limited to the formation of farnesyl diphosphate (FDP). In order to overcome this limitation, the innate carotenoid biosynthetic pathway of *B. megaterium* was extended by the recombinant expression of components of the extensively studied C40 carotenoid pathway of the carotenogenic bacterium and plant pathogen *Pantoea ananatis* (formerly *Erwinia uredovora*) (Coutinho and Venter, 2009). The biosynthetic pathway for the production of the C40 carotenoid β -carotene comprises the open reading frames (ORFs) for a GGDP synthase (PaCrtE), a phytoene synthase (PaCrtB), a phytoene desaturase (PaCrtI) as well as a lycopene cyclase (PaCrtY) (Misawa et al., 1990). These ORFs were assembled as a synthetic operon under the control of the optimized xylose inducible promoter system and subsequently expressed in *B. megaterium* MS941. As shown in Figure 3.3, HPLC analyses of the corresponding carotenoid extracts demonstrated the presence of a carotenoid with an identical retention time to that of a commercial standard of the C40 carotenoid β -carotene.

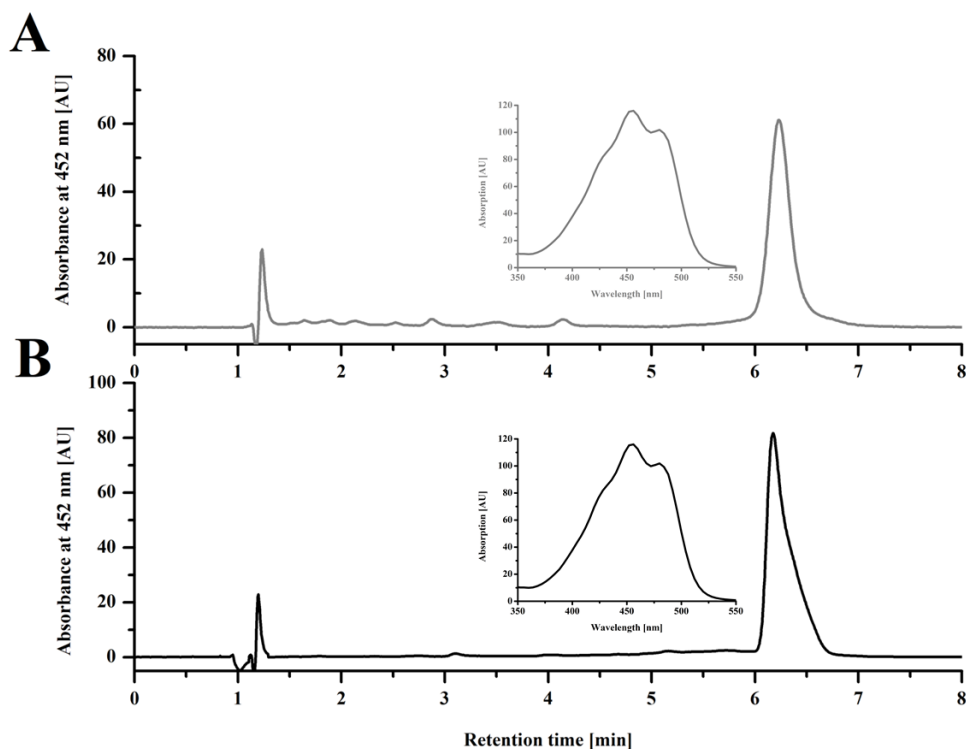


Figure 3.3. RP-HPLC profiles of C40 carotenoids recorded at the indicated wavelength. **A:** commercial β -carotene standard. **B:** C40 carotenoid extract of *B. megaterium* cells. Insets show the absorption spectra of the corresponding C40 carotenoids and were recorded with a photodiode array (PDA) detector.

Moreover, the absorption spectrum of this carotenoid was equivalent to that of β -carotene, showing the characteristic absorption maxima at specific wavelengths of 453 nm and 479 nm, respectively (Takaichi and Shimada, 1992). Most importantly, the C40 carotenoid extract did not show the formation of undesired carotenoid side product possibly originating from a promiscuous activity of the endogenous C30 carotenoid modifying enzymes in *B. megaterium*. Although the selective production of β -carotene in *B. megaterium* seemed quite promising, the total β -carotene yield of 0.28 mg/L was quite low compared to the DNSP yield obtained after homologous expression of the endogenous carotenoid operon. Similar to the optimization approach for the production of DNSP, the β -carotene yields were significantly increased in *B. megaterium* by utilizing the arabinose inducible promoter system described in publication 2.1. Strikingly, this increase turned out to result in a nearly threefold improvement to a maximum of 0.79 mg/L β -carotene.

In summary, the production of the highly demanded C40 carotenoid β -carotene was established in *B. megaterium* for the first time by exploiting the carotenogenic pathway of *P. ananatis*. Despite the presence of an endogenous pathway for the biosynthesis of C30 carotenoids, the β -carotene production in *B. megaterium* was demonstrated to be very selective. Despite the fact, that the β -carotene yields were tripled using the novel arabinose inducible promoter system, the overall performance of the

established *B. megaterium* whole-cell systems was evaluated to be low, particularly with regard to the already optimized *E. coli* systems for the production of β -carotene (Kim et al., 2009). For this reason, C40 carotenoid biosynthesis should be further optimized to be competitive in *B. megaterium*. Popular approaches to obtain a more efficient carotenoid biosynthesis have been reviewed in the literature before and predominantly focus on extensive metabolic pathway engineering to enhance the pool of the isoprenoid precursors IDP and DMADP (Nguyen et al., 2012; Yang and Guo, 2014).

4. Future Perspectives

4.1 Application of the novel promoters and rational strain design of *B. megaterium* for the efficient biotransformation of steroids

B. megaterium was shown to inherently possess beneficial characteristics for the biotransformation of hydrophobic compounds such as steroids and steroid derived drugs. These characteristics include the extraordinary resistance to high concentrations of toxic solvents and the presence of polyhydroxyalkanoate (PHA) bodies. While the resistance to solvents mainly contributes to the enrichment of the cultivation medium with hydrophobic substrates, the PHA bodies were demonstrated to support the accumulation of hydrophobic compounds inside the bacterial cell, thereby significantly facilitating their conversion (Gerber et al., 2015).

Furthermore, *B. megaterium* was established as suitable host for the challenging expression of various oxidoreductases, including several mammalian cytochromes P450, the guinea pig 11 β -hydroxysteroid dehydrogenase (11 β -HSD) as well as the cholesterol oxidase (BCO2) from *B. sterolicum*, which are all involved in the biosynthesis of the pharmaceutically important glucocorticoids cortisol and cortisone. Since the activity of cytochromes P450 is heavily dependent on the availability of auxiliary redox partners, the novel promoter library with its various promoter strengths should be exploited to deliberately manipulate and harmonize the expression levels of cytochromes P450 and corresponding redox partners in order to maximize the efficiency of cytochrome P450 mediated steroid conversions in *B. megaterium*. Moreover, the regulatory DNA sequences of all novel promoters are wildtype sequences, yet some of them were able to compete with the heavily engineered and optimized xylose inducible promoter system in terms of their promoter strength. For this reason, the key regulatory consensus sequences of these promising promoters, like the -10 or -35 regions, should be redesigned according to the example of the optimized xylose inducible promoter system (Stammen et al., 2010). This approach should help to evolve the already powerful novel promoters to even more efficient promoter elements for the recombinant protein production in *B. megaterium*.

Although the abundant side product formation of 20 α -hydroxysteroids was successfully abolished by extensive genetic engineering efforts, the residual 20 β -HSD activity of FabG remains a major challenge for the development of *B. megaterium* towards a selective production of steroidal drugs. As an essential component of the bacterial fatty acid synthesis, the complete elimination of the aberrant 20 β -HSD activity of FabG cannot be approached by gene deletion. Alternatively, the weakest promoters of the novel promoter library could possibly be used to replace the native promoter of the FabG gene, thereby generating a knockdown phenotype with reduced 20 β -HSD activity and consequently reduced side product formation.

4.2 Metabolic engineering of *B. megaterium* towards an improved carotenoid biosynthesis

While the total 4,4'-diaponeurosporene (DNSP) yields in the novel *B. megaterium* based microbial cell factory were shown to be higher than those of any other non-optimized biocatalysts, the corresponding yields of the C40 carotenoid β -carotene were rather low, particularly compared to the β -carotene titers obtained with *E. coli* systems. Similar to these metabolically engineered *E. coli* strains, *B. megaterium* also offers great potential for the optimization of the carotenoid production. As mentioned before, widely applied strategies to boost the carotenoid production in microorganisms aim at increasing the intracellular pool of the common isoprenoid precursors IDP and DMADP. In this context, the recombinant expression of enzymes that catalyze the rate limiting steps of the mevalonate or the non mevalonate pathway was demonstrated to be beneficial for the redistribution of metabolic fluxes towards the biosynthesis of isoprenoids, thus positively contributing to the overall system performance (Nguyen et al., 2012; Yang and Guo, 2014). Accordingly, the overexpression of the DXP synthase (IspB) as well as the DXP reductase (IspC), both involved in the first biosynthetic steps of the non mevalonate pathway, was shown to enhance carotenoid production in *B. subtilis* nearly 8 and 9 fold, respectively (Xue et al., 2015). An equivalent approach would definitely be a reasonable option to optimize carotenoid production in *B. megaterium* significantly.

The most promising potential for a maximization of the carotenoid production in *B. megaterium*, however, is the presence of polyhydroxyalkanoate (PHA) inclusions in the cytoplasmic environment of the bacterial cells. In general, PHAs are built from different monomers, such as saturated as well as unsaturated, branched or substituted 3-hydroxyacids of various lengths, to form large and diverse organic polyesters (Steinbüchel, 1995; Sudesh et al., 2000). PHAs are predominantly produced during excessive nutrient availability and serve as an important source of energy as well as compounds for carbon storage (Pötter and Steinbüchel, 2005). Some microorganisms, including *B. megaterium*, are capable of accumulating up to 70 % of their biomass as PHA bodies (Rodríguez-Contreras et al., 2013). In *B. megaterium*, PHA formation seems to be limited to the monomer 3-hydroxybutyrate, which is derived from the central metabolite acetyl-CoA. The biosynthesis of PHA is initialized by the action of an acetyl-CoA acetyltransferase (PhaA or Acat) which catalyzes the condensation of two molecules of acetyl-CoA to form acetoacetyl-CoA (Grage et al., 2009). This biosynthetic reaction is identical to the initial step of the mevalonate pathway (see Figure 4.1), providing the basis to exploit the immense reservoir of acetoacetyl-CoA in *B. megaterium* for the biosynthesis of IDP and DMADP.

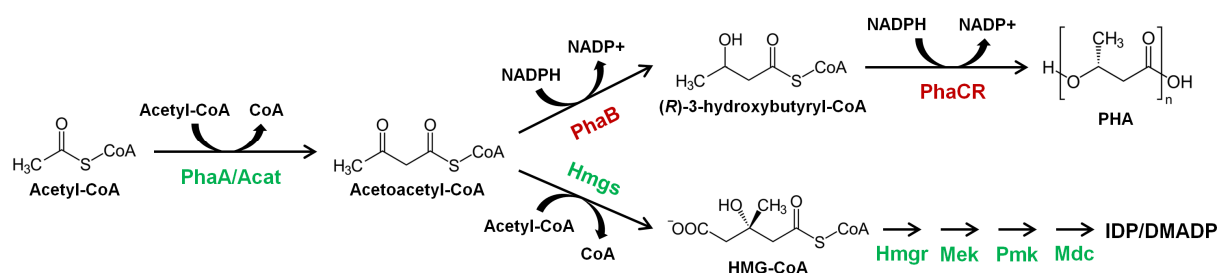
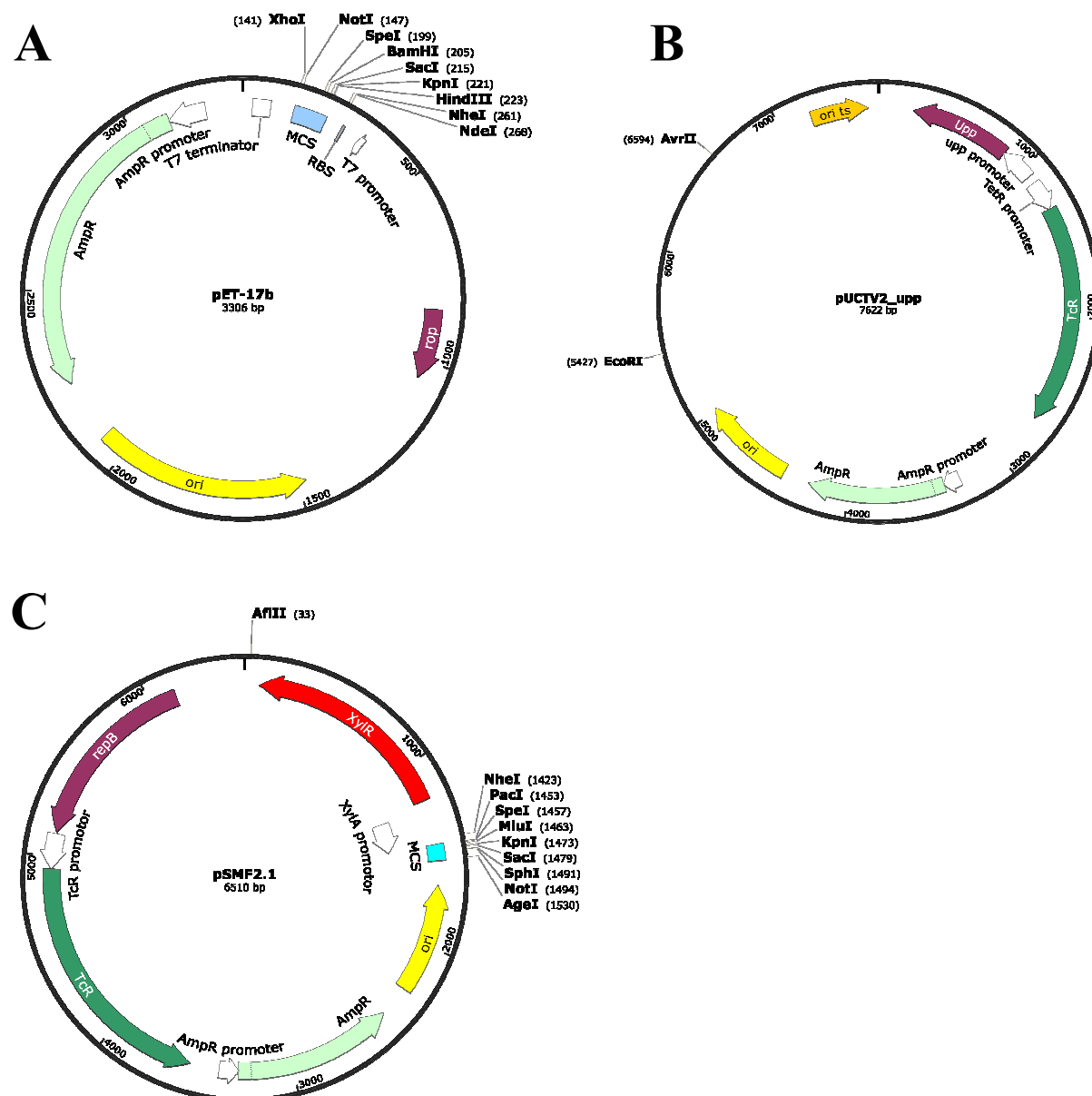


Figure 4.1. Metabolic crossroads of the polyhydroxyalkanoate (PHA) pathway and the mevalonate pathway. Potential knockout targets to block the accumulation of PHAs in *B. megaterium* are shown in red. Enzymes of the mevalonate pathway are shown in green. Their recombinant expression in *B. megaterium* should enable the efficient redirection of acetoacetyl-CoA towards the biosynthesis of IDP and DMADP.

In order to ensure and direct a maximum metabolic flux of acetoacetyl-CoA towards the biosynthetic route of the mevalonate pathway, the subsequent reactions of the PHA biosynthesis necessarily have to be eliminated. These are the NADPH dependent reduction of acetoacetyl-CoA to (*R*)-3-hydroxybutyryl-CoA, catalyzed by the acetoacetyl-CoA reductase (PhaB) and the following iterative condensation of (*R*)-3-hydroxybutyryl-CoA to form the high-molecular weight polymers of PHAs, which is driven by the PHA synthase complex consisting of PhaC and PhaR (McCool and Cannon, 2001; Rehm, 2003). For this reason, a genomic deletion of the corresponding ORFs would be a viable approach for the rational design and development of *B. megaterium* as microbial cell factory for a high yield production of isoprenoid derived compounds with both pharmaceutical and industrial importance.

Appendix

Vector maps



Supplemental Figure. Basic cloning vector maps. Common restriction enzymes and restriction sites are shown in bold next to the vector maps. A: Expression vector pET17b for recombinant protein production in *E. coli*. B: Vector pUCTV2_upp for gene deletions in *B. megaterium*. C: Expression vector pSMF2.1 for recombinant protein expression in *B. megaterium*. AmpR: β -lactamase, MCS: multiple cloning site, ori (ts): (temperature sensitive) origin of replication, RBS: ribosomal binding site, repB: replication protein, rop: regulatory protein of replication, TcR: tetracycline efflux protein, Upp: uracil phosphoribosyltransferase, XylR: xylose repressor, XylA: xylose isomerase.

References

- Abdulmughni, A., Jóźwik, I.K., Putkaradze, N., Brill, E., Zapp, J., Thunnissen, A.-M.W.H., Hannemann, F., Bernhardt, R., 2017. Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D3 hydroxylase. *Journal of Biotechnology* 243, 38–47.
- Abe, H., Fujita, Y., Takaoka, Y., Kurita, E., Yano, S., Tanaka, N., Nakayama, K., 2009. Ethanol-tolerant *Saccharomyces cerevisiae* strains isolated under selective conditions by over-expression of a proofreading-deficient DNA polymerase δ . *Journal of Bioscience and Bioengineering* 108, 199–204.
- Ahmad, S., Roy, P.K., Khan, A.W., Basu, S.K., Johri, B.N., 1991. Microbial transformation of sterols to C19-steroids by *Rhodococcus equi*. *World Journal of Microbiology and Biotechnology* 7, 557–561.
- Alexander, D.L., Fisher, J.F., 1995. A convenient synthesis of 7 α -hydroxycholest-4-en-3-one by the hydroxypropyl- β -cyclodextrin-facilitated cholesterol oxidase oxidation of 3 β ,7 α -cholest-5-ene-3,7-diol. *Steroids* 60, 290–294.
- Alvarez, V., Rodríguez-Sáiz, M., de la Fuente, J.L., Gudiña, E.J., Godio, R.P., Martín, J.F., Barredo, J.L., 2006. The crtS gene of *Xanthophyllomyces dendrorhous* encodes a novel cytochrome-P450 hydroxylase involved in the conversion of β -carotene into astaxanthin and other xanthophylls. *Fungal Genetics and Biology* 43, 261–272.
- An, G.H., Cho, M.H., Johnson, E.A., 1999. Monocyclic carotenoid biosynthetic pathway in the yeast *Phaffia rhodozyma* (*Xanthophyllomyces dendrorhous*). *Journal of Bioscience and Bioengineering* 88, 189–193.
- Anunciato, T.P., da Rocha Filho, P.A., 2012. Carotenoids and polyphenols in nutricosmetics, nutraceuticals, and cosmeceuticals. *Journal of Cosmetic Dermatology* 11, 51–54.
- Archer, G.L., 1998. *Staphylococcus aureus*: a well-armed pathogen. *Clinical Infectious Diseases* 26, 1179–1181.
- Auldridge, M.E., McCarty, D.R., Klee, H.J., 2006. Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Current Opinion in Plant Biology* 9, 315–321.
- Bailey, J.E., 1991. Toward a science of metabolic engineering. *Science* 252, 1668–1675.
- Banerjee, A., Sharkey, T.D., 2014. Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. *Natural Product Reports* 31, 1043–1055.
- Baneyx, F., 1999. Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology* 10, 411–421.
- Bäumchen, C., Roth, A.H.F.J., Biedendieck, R., Malten, M., Follmann, M., Sahn, H., Bringer-Meyer, S., Jahn, D., 2007. D-Mannitol production by resting state whole cell biotransformation of D-fructose by heterologous mannitol and formate dehydrogenase gene expression in *Bacillus megaterium*. *Biotechnology Journal* 2, 1408–1416.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *Journal of Biotechnology* 124, 128–145.

- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Applied Microbiology and Biotechnology* 98, 6185–6203.
- Besserer, A., Puech-Pagès, V., Kiefer, P., Gomez-Roldan, V., Jauneau, A., Roy, S., Portais, J.-C., Roux, C., Bécard, G., Séjalon-Delmas, N., 2006. Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. *PLOS Biology* 4, e226.
- Biedendieck, R., Malten, M., Barg, H., Bunk, B., Martens, J.-H., Deery, E., Leech, H., Warren, M.J., Jahn, D., 2010. Metabolic engineering of cobalamin (vitamin B12) production in *Bacillus megaterium*. *Microbial Biotechnology* 3, 24–37.
- Blasco, F., Kauffmann, I., Schmid, R.D., 2004. CYP175A1 from *Thermus thermophilus* HB27, the first β -carotene hydroxylase of the P450 superfamily. *Applied Microbiology and Biotechnology* 64, 671–674.
- Bleif, S., Hannemann, F., Lisurek, M., von Kries, J.P., Zapp, J., Dietzen, M., Antes, I., Bernhardt, R., 2011. Identification of CYP106A2 as a regioselective allylic bacterial diterpene hydroxylase. *Chembiochem* 12, 576–582.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- β -boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Applied Microbiology and Biotechnology* 93, 1135–1146.
- Boucher, Y., Doolittle, W.F., 2000. The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Molecular Microbiology* 37, 703–716.
- Bratbak, G., Dundas, I., 1984. Bacterial dry matter content and biomass estimations. *Applied and Environmental Microbiology* 48, 755–757.
- Breithaupt, D.E., Bamedi, A., 2001. Carotenoid esters in vegetables and fruits: a screening with emphasis on β -cryptoxanthin esters. *Journal of Agricultural and Food Chemistry* 49, 2064–2070.
- Brey, R.N., Banner, C.D., Wolf, J.B., 1986. Cloning of multiple genes involved with cobalamin (Vitamin B12) biosynthesis in *Bacillus megaterium*. *Journal of Bacteriology* 167, 623–630.
- Briand, L., Marcion, G., Kriznik, A., Heydel, J.M., Artur, Y., Garrido, C., Seigneuric, R., Neiers, F., 2016. A self-inducible heterologous protein expression system in *Escherichia coli*. *Scientific Reports* 6, 33037.
- Britton, G., 1995. Structure and properties of carotenoids in relation to function. *FASEB Journal* 9, 1551–1558.
- Britton, G., Liaaen-Jensen, S., Phander, H., 2008. Carotenoid: natural functions. Birkhäuser Verlag, Basel.
- Brockman, I.M., Prather, K.L.J., 2015a. Dynamic metabolic engineering: new strategies for developing responsive cell factories. *Biotechnology Journal* 10, 1360–1369.
- Brockman, I.M., Prather, K.L.J., 2015b. Dynamic knockdown of *Escherichia coli* central metabolism for redirecting fluxes of primary metabolites. *Metabolic Engineering* 28, 104–113.
- Brunner, M., Bujard, H., 1987. Promoter recognition and promoter strength in the *Escherichia coli* system. *EMBO Journal* 6, 3139–3144.

- Burger, S., Tatge, H., Hofmann, F., Genth, H., Just, I., Gerhard, R., 2003. Expression of recombinant *Clostridium difficile* toxin A using the *Bacillus megaterium* system. *Biochemical and Biophysical Research Communications* 307, 584–588.
- Burke, C.C., Wildung, M.R., Croteau, R., 1999. Geranyl diphosphate synthase: cloning, expression, and characterization of this prenyltransferase as a heterodimer. *Proceedings of the National Academy of Sciences of the U.S.A.* 96, 13062–13067.
- Cankar, K., van Houwelingen, A., Bosch, D., Sonke, T., Bouwmeester, H., Beekwilder, J., 2011. A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. *FEBS Letters* 585, 178–182.
- Cao, W., Ma, W., Wang, X., Zhang, B., Cao, X., Chen, K., Li, Y., Ouyang, P., 2016. Enhanced pinocembrin production in *Escherichia coli* by regulating cinnamic acid metabolism. *Scientific Reports* 6, 32640.
- Carbonell, P., Jervis, A.J., Robinson, C.J., Yan, C., Dunstan, M., Swainston, N., Vinaixa, M., Hollywood, K.A., Currin, A., Rattray, N.J.W., Taylor, S., Spiess, R., Sung, R., Williams, A.R., Fellows, D., Stanford, N.J., Mulherin, P., Le Feuvre, R., Barran, P., Goodacre, R., Turner, N.J., Goble, C., Chen, G.G., Kell, D.B., Micklefield, J., Breitling, R., Takano, E., Faulon, J.-L., Scrutton, N.S., 2018. An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Communications Biology* 1, 66.
- Chae, H.S., Kim, K.-H., Kim, S.C., Lee, P.C., 2010. Strain-dependent carotenoid productions in metabolically engineered *Escherichia coli*. *Applied Biochemistry and Biotechnology* 162, 2333–2344.
- Chen, C., Zou, J., Zhang, S., Zaitlin, D., Zhu, L., 2009. Strigolactones are a new-defined class of plant hormones which inhibit shoot branching and mediate the interaction of plant-AM fungi and plant-parasitic weeds. *Science China Life Sciences* 52, 693–700.
- Chen, Y., Ho, J.M.L., Shis, D.L., Gupta, C., Long, J., Wagner, D.S., Ott, W., Josić, K., Bennett, M.R., 2018. Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nature Communications* 9, 64.
- Chew, B.P., Park, J.S., 2004. Carotenoid action on the immune response. *Journal of Nutrition* 134, 257S-261S.
- Chubukov, V., Mukhopadhyay, A., Petzold, C.J., Keasling, J.D., Martín, H.G., 2016. Synthetic and systems biology for microbial production of commodity chemicals. *Npj Systems Biology and Applications* 2, 16009.
- Clauditz, A., Resch, A., Wieland, K.-P., Peschel, A., Götz, F., 2006. Staphyloxanthin plays a role in the fitness of *Staphylococcus aureus* and its ability to cope with oxidative stress. *Infection and Immunity* 74, 4950–4953.
- Cobb, R.E., Sun, N., Zhao, H., 2013. Directed evolution as a powerful synthetic biology tool. *Methods* 60, 81–90.
- Coulombe, R., Yue, K.Q., Ghisla, S., Vrielink, A., 2001. Oxygen access to the active site of cholesterol oxidase through a narrow channel is gated by an Arg-Glu pair. *Journal of Biological Chemistry* 276, 30435–30441.
- Coutinho, T.A., Venter, S.N., 2009. *Pantoea ananatis*: an unconventional plant pathogen. *Molecular Plant Pathology* 10, 325–335.

- Croteau, N., Vrielink, A., 1996. Crystallization and preliminary X-ray analysis of cholesterol oxidase from *Brevibacterium sterolicum* containing covalently bound FAD. *Journal of Structural Biology* 116, 317–319.
- Dattananda, C.S., Rajkumari, K., Gowrishankar, J., 1991. Multiple mechanisms contribute to osmotic inducibility of proU operon expression in *Escherichia coli*: demonstration of two osmoresponsive promoters and of a negative regulatory element within the first structural gene. *Journal of Bacteriology* 173, 7481–7490.
- De Bary, A., 1884. *Vergleichende Morphologie und Biologie der Pilze, Mycetozen und Bakterien*. Wilhelm Engelmann, Leipzig.
- Di Gennaro, P., Ferrara, S., Bestetti, G., Sello, G., Solera, D., Galli, E., Renzi, F., Bertoni, G., 2008. Novel auto-inducing expression systems for the development of whole-cell biocatalysts. *Applied Microbiology and Biotechnology* 79, 617.
- Donnelly, M.I., Stevens, P.W., Stols, L., Su, S.X., Tollaksen, S., Giometti, C., Joachimiak, A., 2001. Expression of a highly toxic protein, Bax, in *Escherichia coli* by attachment of a leader peptide derived from the GroES cochaperone. *Protein Expression and Purification* 22, 422–429.
- Duc, L.H., Fraser, P.D., Tam, N.K.M., Cutting, S.M., 2006. Carotenoids present in halotolerant *Bacillus* spore formers. *FEMS Microbiology Letters* 255, 215–224.
- Duport, C., Spagnoli, R., Degryse, E., Pompon, D., 1998. Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. *Nature Biotechnology* 16, 186–189.
- Egorova, O.V., Gulevskaya, S.A., Puntus, I.F., Filonov, A.E., Donova, M.V., 2002. Production of androstenedione using mutants of *Mycobacterium* sp. *Journal of Chemical Technology & Biotechnology* 77, 141–147.
- Ehrhardt, M., Gerber, A., Hannemann, F., Bernhardt, R., 2016. Expression of human CYP27A1 in *B. megaterium* for the efficient hydroxylation of cholesterol, vitamin D3 and 7-dehydrocholesterol. *Journal of Biotechnology* 218, 34–40.
- El-Kadi, I.A., Eman Mostafa, M., 2004. Hydroxylation of progesterone by some *Trichoderma* species. *Folia Microbiologica* 49, 285–290.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin, M., Elbourne, L.D.H., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukyanchuk, V., Ball, B., Ul-Qamar, R., Seibel, J., Bremer, E., Jahn, D., Ravel, J., Vary, P.S., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *Journal of Bacteriology* 193, 4199–4213.
- Ernst, H., 2002. Recent advances in industrial carotenoid synthesis. *Pure and Applied Chemistry*. 74, 1369–1382.
- Farnberger, J.E., Lorenz, E., Richter, N., Wendisch, V.F., Kroutil, W., 2017. In vivo plug-and-play: a modular multi-enzyme single-cell catalyst for the asymmetric amination of ketoacids and ketones. *Microbial Cell Factories*. 16, 132.
- Feklistov, A., 2013. RNA polymerase: in search of promoters. *Annals of the New York Academy of Sciences* 1293, 25–32.
- Fernández-Cabezón, L., Galán, B., García, J.L., 2018. New insights on steroid biotechnology. *Frontiers in Microbiology* 9, 958.

- Fernández-Cabezón, L., Galán, B., García, J.L., 2017. Engineering *Mycobacterium smegmatis* for testosterone production. *Microbial Biotechnology* 10, 151–161.
- Fiore, A., Dall’Osto, L., Fraser, P.D., Bassi, R., Giuliano, G., 2006. Elucidation of the β -carotene hydroxylation pathway in *Arabidopsis thaliana*. *FEBS Letters* 580, 4718–4722.
- Fleischmann, P., Zorn, H., 2008. Enzymic pathways for formation of carotenoid cleavage products. In: Britton, G., Liaaen-Jensen, S., Pfander, H. (Eds.), *Carotenoids*. Birkhäuser Basel, Basel, pp. 341–366.
- Fraatz, M.A., Berger, R.G., Zorn, H., 2009. Nootkatone—a biotechnological challenge. *Applied Microbiology and Biotechnology* 83, 35–41.
- Francis, D.M., Page, R., 2010. Strategies to optimize protein expression in *E. coli*. *Current Protocols in Protein Science* 61, 5.24.1-5.24.29.
- Frank, H.A., Cogdell, R.J., 1996. Carotenoids in photosynthesis. *Photochemistry and Photobiology* 63, 257–264.
- Funder, J.W., Krozowski, Z., Myles, K., Sato, A., Sheppard, K.E., Young, M., 1997. Mineralocorticoid receptors, salt, and hypertension. *Recent Progress in Hormone Research*. 52, 247–262.
- Furubayashi, M., Li, L., Katabami, A., Saito, K., Umeno, D., 2014. Construction of carotenoid biosynthetic pathways using squalene synthase. *FEBS Letters* 588, 436–442.
- Gamer, M., Fröde, D., Biedendieck, R., Stammen, S., Jahn, D., 2009. A T7 RNA polymerase-dependent gene expression system for *Bacillus megaterium*. *Applied Microbiology and Biotechnology* 82, 1195.
- Gardner, T.S., Cantor, C.R., Collins, J.J., 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339–342.
- Gavira, C., Höfer, R., Lesot, A., Lambert, F., Zucca, J., Werck-Reichhart, D., 2013. Challenges and pitfalls of P450-dependent (+)-valencene bioconversion by *Saccharomyces cerevisiae*. *Metabolic Engineering* 18, 25–35.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microbial Cell Factories* 14, 107.
- Gerster, H., 1993. Anticarcinogenic effect of common carotenoids. *International Journal for Vitamin and Nutrition Research* 63, 93–121.
- Ghayee, H.K., Auchus, R.J., 2007. Basic concepts and recent developments in human steroid hormone biosynthesis. *Reviews in Endocrine and Metabolic Disorders* 8, 289–300.
- Ghosh, T., Bose, D., Zhang, X., 2010. Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiology Reviews* 34, 611–627.
- Giacalone, M.J., Gentile, A.M., Lovitt, B.T., Berkley, N.L., Gunderson, C.W., Surber, M.W., 2006. Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *BioTechniques* 40, 355–364.

- Gilep, A.A., Sushko, T.A., Usanov, S.A., 2011. At the crossroads of steroid hormone biosynthesis: the role, substrate specificity and evolutionary development of CYP17. *Biochimica et Biophysica Acta* 1814, 200–209.
- Girhard, M., Machida, K., Itoh, M., Schmid, R.D., Arisawa, A., Urlacher, V.B., 2009. Regioselective biooxidation of (+)-valencene by recombinant *E. coli* expressing CYP109B1 from *Bacillus subtilis* in a two-liquid-phase system. *Microbial Cell Factories*. 8, 36.
- Gosset, G., 2008. Microbial production of industrial chemicals. Introduction. *Journal of Molecular Microbiology and Biotechnology* 15, 5–7.
- Grage, K., Jahns, A.C., Parlane, N., Palanisamy, R., Rasiah, I.A., Atwood, J.A., Rehm, B.H.A., 2009. Bacterial polyhydroxyalkanoate granules: biogenesis, structure, and potential use as nano-/microbeads in biotechnological and biomedical applications. *Biomacromolecules* 10, 660–669.
- Greener, A., Callahan, M., Jerpseth, B., 1997. An efficient random mutagenesis technique using an *E. coli* mutator strain. *Molecular Biotechnology* 7, 189–195.
- Greenfield, L., Boone, T., Wilcox, G., 1978. DNA sequence of the araBAD promoter in *Escherichia coli* B/r. *Proceedings of the National Academy of Sciences of the U.S.A.* 75, 4724–4728.
- Guan, C., Cui, W., Cheng, J., Zhou, L., Guo, J., Hu, X., Xiao, G., Zhou, Z., 2015. Construction and development of an auto-regulatory gene expression system in *Bacillus subtilis*. *Microbial Cell Factories* 14, 150.
- Guggisberg, A.M., Amthor, R.E., Odom, A.R., 2014. Isoprenoid biosynthesis in *Plasmodium falciparum*. *Eukaryotic Cell* 13, 1348–1359.
- Guiziou, S., Sauveplane, V., Chang, H.-J., Clerté, C., Declerck, N., Jules, M., Bonnet, J., 2016. A part toolbox to tune genetic expression in *Bacillus subtilis*. *Nucleic Acids Research*. 44, 7495–7508.
- Gupta, A., Brockman Reizman, I.M., Reisch, C.R., Prather, K.L.J., 2017. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature Biotechnology* 35, 273–279.
- Hakki, T., Zearo, S., Drăgan, C.-A., Bureik, M., Bernhardt, R., 2008. Coexpression of redox partners increases the hydrocortisone (cortisol) production efficiency in CYP11B1 expressing fission yeast *Schizosaccharomyces pombe*. *Journal of Biotechnology* 133, 351–359.
- Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R., 2007. Cytochrome P450 systems—biological variations of electron transport chains. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1770, 330–344.
- Hara, M., Yuan, H., Yang, Q., Hoshino, T., Yokoyama, A., Miyake, J., 1999. Stabilization of liposomal membranes by thermozeaxanthins: carotenoid-glucoside esters. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1461, 147–154.
- Hebeda, R.E., Styrlund, C.R., Teague, W.M., 1988. Benefits of *Bacillus megaterium* amylase in dextrose production. *Starch* 40, 33–36.
- Hemmati, H., Basu, C., 2015. Transcriptional analyses of an ethanol inducible promoter in *Escherichia coli* and tobacco for production of cellulase and green fluorescent protein. *Biotechnology & Biotechnological Equipment* 29, 1043–1052.
- Hii, S.L., Tan, J.S., Ling, T.C., Ariff, A.B., 2012. Pullulanase: role in starch hydrolysis and potential industrial applications. *Enzyme Research* 2012, 921362.

- Hofsten, B. v., 1961. The inhibitory effect of galactosides on the growth of *Escherichia coli*. *Biochimica et Biophysica Acta* 48, 164–171.
- Hold, C., Billerbeck, S., Panke, S., 2016. Forward design of a complex enzyme cascade reaction. *Nature Communications* 7, 12971.
- Hopkins, J.D., 1974. A new class of promoter mutations in the lactose operon of *Escherichia coli*. *Journal of Molecular Biology* 87, 715–724.
- Hoshino, T., Fujii, R., Nakahara, T., 1994. Overproduction of carotenoids in *Thermus thermophilus*. *Journal of Fermentation and Bioengineering* 77, 423–424.
- Hwang, H.J., Lee, S.Y., Lee, P.C., 2018. Engineering and application of synthetic nar promoter for fine-tuning the expression of metabolic pathway genes in *Escherichia coli*. *Biotechnology for Biofuels* 11, 103.
- Iniesta, A.A., Cervantes, M., Murillo, F.J., 2008. Conversion of the lycopene monocyclus of *Myxococcus xanthus* into a bicyclus. *Applied Microbiology and Biotechnology* 79, 793–802.
- Jain, S., 2014. Biosynthesis of archaeal membrane ether lipids. *Frontiers in Microbiology* 5.
- Jankowicz-Cieslak, J., Mba, C., Till, B.J., 2017. Mutagenesis for crop breeding and functional genomics. In: Jankowicz-Cieslak, J., Tai, T.H., Kumlehn, J., Till, B.J. (Eds.), *Biotechnologies for Plant Mutation Breeding*. Springer International Publishing, Cham, pp. 3–18.
- Jayaraman, P., Yeoh, J.W., Zhang, J., Poh, C.L., 2018. Programming the dynamic control of bacterial gene expression with a chimeric ligand- and light-based promoter system. *ACS Synthetic Biology* 7, 2627–2639.
- Jing, Y., Liu, H., Xu, W., Yang, Q., 2019. 4,4'-Diaponeurosporene-Producing *Bacillus subtilis* promotes the development of the mucosal immune system of the piglet gut. *Anatomical Record (Hoboken)*.
- Jing, Y., Liu, H., Xu, W., Yang, Q., 2017. Amelioration of the DSS-induced colitis in mice by pretreatment with 4,4'-diaponeurosporene-producing *Bacillus subtilis*. *Experimental and Therapeutic Medicine* 14, 6069–6073.
- Jones, J.A., Vernacchio, V.R., Lachance, D.M., Lebovich, M., Fu, L., Shirke, A.N., Schultz, V.L., Cress, B., Linhardt, R.J., Koffas, M.A.G., 2015. ePathOptimize: a combinatorial approach for transcriptional balancing of metabolic pathways. *Scientific Reports* 5, 11301.
- Joseph, B.C., Pichaimuthu, S., Srimeenakshi, S., 2015. An overview of the parameters for recombinant protein expression in *Escherichia coli*. *Journal of Cell Science & Therapy* 06.
- Jóźwik, I.K., Kiss, F.M., Gricman, Ł., Abdulmughni, A., Brill, E., Zapp, J., Pleiss, J., Bernhardt, R., Thunnissen, A.-M.W.H., 2016. Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *The FEBS Journal* 283, 4128–4148.
- Kaur, Jashandeep, Kumar, A., Kaur, Jagdeep, 2018. Strategies for optimization of heterologous protein expression in *E. coli*: roadblocks and reinforcements. *International Journal of Biological Macromolecules*. 106, 803–822.
- Khatri, Y., Schifrin, A., Bernhardt, R., 2017. Investigating the effect of available redox protein ratios for the conversion of a steroid by a myxobacterial CYP260A1. *FEBS Letters* 591, 1126–1140.

- Khoo, H.-E., Prasad, K.N., Kong, K.-W., Jiang, Y., Ismail, A., 2011. Carotenoids and their isomers: color pigments in fruits and vegetables. *Molecules* 16, 1710–1738.
- Kim, J.H., Kim, S.-W., Nguyen, D.Q.A., Li, H., Kim, S.B., Seo, Y.-G., Yang, J.-K., Chung, I.-Y., Kim, D.H., Kim, C.-J., 2009. Production of β -carotene by recombinant *Escherichia coli* with engineered whole mevalonate pathway in batch and fed-batch cultures. *Biotechnology and Bioprocess Engineering* 14, 559–564.
- Kim, J.R., Kim, S.H., Lee, S.Y., Lee, P.C., 2013. Construction of homologous and heterologous synthetic sucrose utilizing modules and their application for carotenoid production in recombinant *Escherichia coli*. *Bioresource Technology* 130, 288–295.
- Kim, S.H., Kim, M.S., Lee, B.Y., Lee, P.C., 2016. Generation of structurally novel short carotenoids and study of their biological activity. *Scientific Reports* 6.
- Kirti, K., Amita, S., Priti, S., Mukesh Kumar, A., Jyoti, S., 2014. Colorful world of microbes: carotenoids and their applications. *Advances in Biology* 2014, 1–13.
- Kiss, F.M., Khatri, Y., Zapp, J., Bernhardt, R., 2015. Identification of new substrates for the CYP106A1-mediated 11-oxidation and investigation of the reaction mechanism. *FEBS Letters* 589, 2320–2326.
- Kiss, F.M., Lundemo, M.T., Zapp, J., Woodley, J.M., Bernhardt, R., 2015. Process development for the production of 15 β -hydroxycyproterone acetate using *Bacillus megaterium* expressing CYP106A2 as whole-cell biocatalyst. *Microbial Cell Factories* 14, 28.
- Köcher, S., Breitenbach, J., Müller, V., Sandmann, G., 2009. Structure, function and biosynthesis of carotenoids in the moderately halophilic bacterium *Halobacillus halophilus*. *Archives of Microbiology* 191, 95–104.
- Korneli, C., Biedendieck, R., David, F., Jahn, D., Wittmann, C., 2013. High yield production of extracellular recombinant levansucrase by *Bacillus megaterium*. *Applied Microbiology and Biotechnology* 97, 3343–3353.
- Kumar, S., H., A., Asif, M., Goyal, A., 2013. Molecular mechanisms controlling dormancy and germination in barley, in: Goyal, A. (Ed.), *Crop Production*. InTech.
- Kuzuyama, T., Seto, H., 2012. Two distinct pathways for essential metabolic precursors for isoprenoid biosynthesis. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences* 88, 41–52.
- Labaree, D., Hoyte, R.M., Hochberg, R.B., 1997. A direct stereoselective synthesis of 7 β -hydroxytestosterone. *Steroids* 62, 482–486.
- Land, M., Hauser, L., Jun, S.-R., Nookaew, I., Leuze, M.R., Ahn, T.-H., Karpinets, T., Lund, O., Kora, G., Wassenaar, T., Poudel, S., Ussery, D.W., 2015. Insights from 20 years of bacterial genome sequencing. *Functional and Integrative Genomics* 15, 141–161.
- Lawson, A.J., Walker, E.A., White, S.A., Dafforn, T.R., Stewart, P.M., Ride, J.P., 2009. Mutations of key hydrophobic surface residues of 11 beta-hydroxysteroid dehydrogenase type 1 increase solubility and monodispersity in a bacterial expression system. *Protein Science*. 18, 1552–1563.
- Lednicer, D., 2011. *Steroid chemistry at a glance*. Wiley, Hoboken.

- Lee, G.-Y., Kim, D.-H., Kim, D., Ahn, T., Yun, C.-H., 2015. Functional characterization of steroid hydroxylase CYP106A1 derived from *Bacillus megaterium*. *Archives of Pharmacal Research* 38, 98–107.
- Lee, J.W., Kim, H.U., Choi, S., Yi, J., Lee, S.Y., 2011. Microbial production of building block chemicals and polymers. *Current Opinion in Biotechnology* 22, 758–767.
- Li, L., Furubayashi, M., Hosoi, T., Seki, T., Otani, Y., Kawai-Noma, S., Saito, K., Umeno, D., 2019. Construction of a nonnatural C60 carotenoid biosynthetic pathway. *ACS Synthetic Biology* 8, 511–520.
- Lichtenthaler, H.K., 1999. The 1-deoxy-d-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 50, 47–65.
- Liu, H., Xu, W., Chang, X., Qin, T., Yin, Y., Yang, Q., 2016. 4,4'-Diaponeurosporene, a C30 carotenoid, effectively activates dendritic cells via CD36 and NF- κ B signaling in a ROS independent manner. *Oncotarget* 7, 40978–40991.
- Liu, H., Xu, W., Yu, Q., Yang, Q., 2017. 4,4'-Diaponeurosporene-producing *Bacillus subtilis* increased mouse resistance against *Salmonella typhimurium* infection in a CD36-dependent manner. *Frontiers in Immunology* 8, 483.
- Liu, L., Li, Y., Zhang, J., Zou, W., Zhou, Z., Liu, J., Li, X., Wang, L., Chen, J., 2011. Complete genome sequence of the industrial strain *Bacillus megaterium* WSH-002. *Journal of Bacteriology* 193, 6389–6390.
- Liu, X., Wang, H., Wang, B., Pan, L., 2018. High-level extracellular protein expression in *Bacillus subtilis* by optimizing strong promoters based on the transcriptome of *Bacillus subtilis* and *Bacillus megaterium*. *Protein Expression and Purification* 151, 72–77.
- Lobo, G.P., Amengual, J., Palczewski, G., Babino, D., von Lintig, J., 2012. Mammalian carotenoid-oxygenases: key players for carotenoid function and homeostasis. *Biochimica et Biophysica Acta* 1821, 78–87.
- Ludwig, H., Stülke, J., 2001. The *Bacillus subtilis* catabolite control protein CcpA exerts all its regulatory functions by DNA-binding. *FEMS Microbiology Letters* 203, 125–129.
- Ma, B.-X., Ke, X., Tang, X.-L., Zheng, R.-C., Zheng, Y.-G., 2018. Rate-limiting steps in the *Saccharomyces cerevisiae* ergosterol pathway: towards improved ergosta-5,7-dien-3 β -ol accumulation by metabolic engineering. *World Journal of Microbiology and Biotechnology* 34, 55.
- Machado, D., Herrgård, M.J., 2015. Co-evolution of strain design methods based on flux balance and elementary mode analysis. *Metabolic Engineering Communications* 2, 85–92.
- Markley, A.L., Begemann, M.B., Clarke, R.E., Gordon, G.C., Pfleger, B.F., 2015. Synthetic biology toolbox for controlling gene expression in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *ACS Synthetic Biology* 4, 595–603.
- Marshall, J.H., Wilmoth, G.J., 1981. Proposed pathway of triterpenoid carotenoid biosynthesis in *Staphylococcus aureus*: evidence from a study of mutants. *Journal of Bacteriology* 147, 914–919.
- Martín, L., Prieto, M.A., Cortés, E., García, J.L., 1995. Cloning and sequencing of the pac gene encoding the penicillin G acylase of *Bacillus megaterium* ATCC 14945. *FEMS Microbiology Letters* 125, 287–292.

- McCloskey, D., Palsson, B.Ø., Feist, A.M., 2013. Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Molecular Systems Biology* 9, 661.
- McCool, G., 1996. Polyhydroxyalkanoate inclusion-body growth and proliferation in *Bacillus megaterium*. *FEMS Microbiology Letters* 138, 41–48.
- McCool, G.J., Cannon, M.C., 2001. PhaC and PhaR are required for polyhydroxyalkanoic acid synthase activity in *Bacillus megaterium*. *Journal of Bacteriology* 183, 4235–4243.
- McDevitt, T.M., Tchao, R., Harrison, E.H., Morel, D.W., 2005. Carotenoids normally present in serum inhibit proliferation and induce differentiation of a human monocyte/macrophage cell line (U937). *Journal of Nutrition* 135, 160–164.
- McGarvey, D.J., 1995. Terpenoid Metabolism. *The Plant Cell Online* 7, 1015–1026.
- Meng, F., Zhu, X., Nie, T., Lu, F., Bie, X., Lu, Y., Trouth, F., Lu, Z., 2018. Enhanced expression of pullulanase in *Bacillus subtilis* by new strong promoters mined from transcriptome data, both alone and in combination. *Frontiers in Microbiology* 9.
- Metz, R.J., Allen, L.N., Zeman, N.W., 1988. Nucleotide sequence of an amylase gene from *Bacillus megaterium*. *Nucleic Acids Research* 16, 5203.
- Mialoundama, A.S., Heintz, D., Jadid, N., Nkeng, P., Rahier, A., Deli, J., Camara, B., Bouvier, F., 2010. Characterization of plant carotenoid cyclases as members of the flavoprotein family functioning with no net redox change. *Plant Physiology* 153, 970–979.
- Milhim, M., Putkaradze, N., Abdulmughni, A., Kern, F., Hartz, P., Bernhardt, R., 2016. Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. *Journal of Biotechnology* 240, 68–75.
- Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K., Harashima, K., 1990. Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. *Journal of Bacteriology* 172, 6704–6712.
- Mitchell, C., Iyer, S., Skomursi, J.F., Vary, J.C., 1986. Red pigment in *Bacillus megaterium* spores. *Applied and Environmental Microbiology* 52, 64–67.
- Miyada, C.G., Stoltzfus, L., Wilcox, G., 1984. Regulation of the *araC* gene of *Escherichia coli*: catabolite repression, autoregulation, and effect on *araBAD* expression. *Proceedings of the National Academy of Sciences* 81, 4120–4124.
- Miyake, Y., Fukushima, W., Tanaka, K., Sasaki, S., Kiyohara, C., Tsuboi, Y., Yamada, T., Oeda, T., Miki, T., Kawamura, N., Sakae, N., Fukuyama, H., Hirota, Y., Nagai, M., Fukuoka Kinki Parkinson's disease study group, 2011. Dietary intake of antioxidant vitamins and risk of Parkinson's disease: a case-control study in Japan. *European Journal of Neurology* 18, 106–113.
- Moeller, R., Horneck, G., Facius, R., Stackebrandt, E., 2005. Role of pigmentation in protecting *Bacillus* sp. endospores against environmental UV radiation. *FEMS Microbiology Ecology* 51, 231–236.
- Moise, A.R., Al-Babili, S., Wurtzel, E.T., 2014. Mechanistic aspects of carotenoid biosynthesis. *Chemical Reviews* 114, 164–193.
- Montigny, C., Penin, F., Lethias, C., Falson, P., 2004. Overcoming the toxicity of membrane peptide expression in bacteria by upstream insertion of Asp-Pro sequence. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1660, 53–65.

- Moore, S.J., Mayer, M.J., Biedendieck, R., Deery, E., Warren, M.J., 2014. Towards a cell factory for vitamin B12 production in *Bacillus megaterium*: bypassing of the cobalamin riboswitch control elements. *New Biotechnology, Synthetic Biotechnology* 31, 553–561.
- Morita, M., Tomita, K., Ishizawa, M., Takagi, K., Kawamura, F., Takahashi, H., Morino, T., 1999. Cloning of oxetanocin A biosynthetic and resistance genes that reside on a plasmid of *Bacillus megaterium* strain NK84-0128. *Bioscience, Biotechnology, and Biochemistry* 63, 563–566.
- Mortensen, A., 2006. Carotenoids and other pigments as natural colorants. *Pure and Applied Chemistry* 78, 1477–1491.
- Müller, D., Stelling, J., 2009. Precise regulation of gene expression dynamics favors complex promoter architectures. *PLOS Computational Biology* 5, e1000279.
- Müller, P., Li, X.-P., Niyogi, K.K., 2001. Non-photochemical quenching. A response to excess light energy. *Plant Physiology* 125, 1558–1566.
- Munck, A., Guyre, P.M., Holbrook, N.J., 1984. Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Reviews* 5, 25–44.
- Neuman, H., Galpaz, N., Cunningham, F.X., Zamir, D., Hirschberg, J., 2014. The tomato mutation *nxd1* reveals a gene necessary for neoxanthin biosynthesis and demonstrates that violaxanthin is a sufficient precursor for abscisic acid biosynthesis. *Plant Journal* 78, 80–93.
- Neunzig, J., Milhim, M., Schiffer, L., Khatri, Y., Zapp, J., Sánchez-Guijo, A., Hartmann, M.F., Wudy, S.A., Bernhardt, R., 2017. The steroid metabolite 16(β)-OH-androstenedione generated by CYP21A2 serves as a substrate for CYP19A1. *Journal of Steroid Biochemistry and Molecular Biology* 167, 182–191.
- Nevoigt, E., 2008. Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* 72, 379–412.
- Nguyen, A.D.Q., Kim, S.-W., Kim, S.B., Seo, Y.-G., Chung, I.-Y., Kim, D.H., Kim, C.-J., 2012. Production of β -carotene and acetate in recombinant *Escherichia coli* with or without mevalonate pathway at different culture temperature or pH. *Biotechnology and Bioprocess Engineering* 17, 1196–1204.
- Nielsen, J., Keasling, J.D., 2016. Engineering cellular metabolism. *Cell* 164, 1185–1197.
- Nocadello, S., Swennen, E.F., 2012. The new pLAI (lux regulon based auto-inducible) expression system for recombinant protein production in *Escherichia coli*. *Microbial Cell Factories* 11, 3.
- Norris, S.R., Barrette, T.R., DellaPenna, D., 1995. Genetic dissection of carotenoid synthesis in arabidopsis defines plastoquinone as an essential component of phytoene desaturation. *The Plant Cell* 7, 2139–2149.
- North, H.M., De Almeida, A., Boutin, J.-P., Frey, A., To, A., Botran, L., Sotta, B., Marion-Poll, A., 2007. The *Arabidopsis* ABA-deficient mutant *aba4* demonstrates that the major route for stress-induced ABA accumulation is via neoxanthin isomers. *Plant Journal* 50, 810–824.
- Obulesu, M., Dowlathabad, M.R., Bramhachari, P.V., 2011. Carotenoids and Alzheimer's disease: an insight into therapeutic role of retinoids in animal models. *Neurochemistry International* 59, 535–541.

- Palmerín-Carreño, D.M., Rutiaga-Quiñones, O.M., Verde-Calvo, J.R., Prado-Barragán, A., Huerta-Ochoa, S., 2016. Whole cell bioconversion of (+)-valencene to (+)-nootkatone in 100 % organic phase using *Yarrowia lipolytica* 2.2ab. *International Journal of Chemical Reactor Engineering* 14, 939–944.
- Parekh, S., Vinci, V.A., Strobel, R.J., 2000. Improvement of microbial strains and fermentation processes. *Applied Microbiology and Biotechnology* 54, 287–301.
- Partow, S., Siewers, V., Daviet, L., Schalk, M., Nielsen, J., 2012. Reconstruction and evaluation of the synthetic bacterial MEP pathway in *Saccharomyces cerevisiae*. *PLOS ONE* 7, e52498.
- Pauly, H.E., Pfeleiderer, G., 1975. D-glucose dehydrogenase from *Bacillus megaterium* M 1286: purification, properties and structure. *Hoppe-Seyler's Zeitschrift für physiologische Chemie* 356, 1613–1623.
- Pelz, A., Wieland, K.-P., Putzbach, K., Hentschel, P., Albert, K., Gotz, F., 2005. Structure and biosynthesis of staphyloxanthin from *Staphylococcus aureus*. *Journal of Biological Chemistry* 280, 32493–32498.
- Perez-Fons, L., Steiger, S., Khaneja, R., Bramley, P.M., Cutting, S.M., Sandmann, G., Fraser, P.D., 2011. Identification and the developmental formation of carotenoid pigments in the yellow/orange *Bacillus* spore-formers. *Biochimica et Biophysica Acta* 1811, 177–185.
- Peterson, D.H., Murray, H.C., 1952. Microbiological oxygenation of steroids at carbon 11. *Journal of the American Chemical Society* 74, 1871–1872.
- Pitera, D.J., Paddon, C.J., Newman, J.D., Keasling, J.D., 2007. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metabolic Engineering* 9, 193–207.
- Pötter, M., Steinbüchel, A., 2005. Poly(3-hydroxybutyrate) granule-associated proteins: impacts on poly(3-hydroxybutyrate) synthesis and degradation. *Biomacromolecules* 6, 552–560.
- Priest, F.G., 1977. Extracellular enzyme synthesis in the genus *Bacillus*. *Bacteriological Reviews*. 41, 711–753.
- Putkaradze, N., Kiss, F.M., Schmitz, D., Zapp, J., Hutter, M.C., Bernhardt, R., 2017a. Biotransformation of prednisone and dexamethasone by cytochrome P450 based systems - Identification of new potential drug candidates. *Journal of Biotechnology* 242, 101–110.
- Putkaradze, N., Litzenburger, M., Abdumughni, A., Milhim, M., Brill, E., Hannemann, F., Bernhardt, R., 2017b. CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*. *Applied Microbiology and Biotechnology* 101, 8379–8393.
- Putkaradze, N., Litzenburger, M., Hutter, M.C., Bernhardt, R., 2019. CYP109E1 from *Bacillus megaterium* acts as a 24- and 25-hydroxylase for cholesterol. *Chembiochem* 20, 655–658.
- Racine, F.M., Vary, J.C., 1980. Isolation and properties of membranes from *Bacillus megaterium* spores. *J. Bacteriol.* 143, 1208–1214.
- Radha, S., Gunasekaran, P., 2007. Cloning and expression of keratinase gene in *Bacillus megaterium* and optimization of fermentation conditions for the production of keratinase by recombinant strain. *Journal of Applied Microbiology* 103, 1301–1310.
- Raffy, S., Teissié, J., 1999. Control of lipid membrane stability by cholesterol content. *Biophysical Journal* 76, 2072–2080.

- Rao, A.V., Agarwal, S., 2000. Role of antioxidant lycopene in cancer and heart disease. *Journal of the American College of Nutrition* 19, 563–569.
- Ray, R.C., V K Joshi, 2014. Fermented foods: past, present and future.
- Redden, H., Morse, N., Alper, H.S., 2015. The synthetic biology toolbox for tuning gene expression in yeast. *FEMS Yeast Research* 15, 1–10.
- Reetz, M.T., Kahakeaw, D., Lohmer, R., 2008. Addressing the numbers problem in directed evolution. *ChemBioChem* 9, 1797–1804.
- Rehm, B.H.A., 2003. Polyester synthases: natural catalysts for plastics. *Biochemistry Journal* 376, 15–33.
- Rodríguez-Contreras, A., Koller, M., Miranda-de Sousa Dias, M., Calafell-Monfort, M., Braunegg, G., Marqués-Calvo, M.S., 2013. High production of poly(3-hydroxybutyrate) from a wild *Bacillus megaterium* bolivian strain. *Journal of Applied Microbiology* 114, 1378–1387.
- Rodríguez-García, A., Combes, P., Pérez-Redondo, R., Smith, M.C.A., Smith, M.C.M., 2005. Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces*. *Nucleic Acids Research*. 33, e87.
- Rohmer, M., 1999. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nature Product Reports* 16, 565–574.
- Rouch, D.A., Brown, N.L., 1997. Copper-inducible transcriptional regulation at two promoters in the *Escherichia coli* copper resistance determinant *pco*. *Microbiology* 143, 1191–1202.
- Ruiz-Sola, M.Á., Rodríguez-Concepción, M., 2012. Carotenoid biosynthesis in *Arabidopsis*: a colorful pathway. *Arabidopsis Book* 10.
- Rygun, T., Hillen, W., 1992. Catabolite repression of the *xyl* operon in *Bacillus megaterium*. *Journal of Bacteriology* 174, 3049–3055.
- Rygun, T., Scheler, A., Allmansberger, R., Hillen, W., 1991. Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. *Archives of Microbiology* 155, 535–542.
- Rygun, T., Hillen, W., 1991. Inducible high-level expression of heterologous genes in *Bacillus megaterium* using the regulatory elements of the xylose-utilization operon. *Applied Microbiology and Biotechnology* 35, 594–599.
- Saecker, R.M., Record, M.T., deHaseth, P.L., 2011. Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *Journal of Molecular Biology, Mechanisms of Transcription* 412, 754–771.
- Sarett, L.H., 1946. Partial synthesis of pregnene-4-triol-17(β), 20(β), 21-dione-3,11 and pregnene-4-diol-17(β), 21-trione-3,11,20 monoacetate. *Journal of Biological Chemistry* 162, 601–631.
- Scalcinati, G., Knuf, C., Partow, S., Chen, Y., Maury, J., Schalk, M., Daviet, L., Nielsen, J., Siewers, V., 2012. Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene α -santalene in a fed-batch mode. *Metabolic Engineering* 14, 91–103.

- Schiffer, L., Anderko, S., Hobler, A., Hannemann, F., Kagawa, N., Bernhardt, R., 2015. A recombinant CYP11B1 dependent *Escherichia coli* biocatalyst for selective cortisol production and optimization towards a preparative scale. *Microbial Cell Factories* 14, 25.
- Schmitz, D., Janocha, S., Kiss, F.M., Bernhardt, R., 2018. CYP106A2-A versatile biocatalyst with high potential for biotechnological production of selectively hydroxylated steroid and terpenoid compounds. *Biochimica et Biophysica Acta Proteins and Proteomics* 1866, 11–22.
- Schmitz, D., Zapp, J., Bernhardt, R., 2014. Steroid conversion with CYP106A2 - production of pharmaceutically interesting DHEA metabolites. *Microbial Cell Factories* 13, 81.
- Schulz, S., Girhard, M., Gaßmeyer, S.K., Jäger, V.D., Schwarze, D., Vogel, A., Urlacher, V.B., 2015. Selective enzymatic synthesis of the grapefruit flavor (+)-nootkatone. *ChemCatChem* 7, 601–604.
- Schwartz, S.H., Qin, X., Zeevaart, J.A., 2001. Characterization of a novel carotenoid cleavage dioxygenase from plants. *Journal of Biological Chemistry* 276, 25208–25211.
- Sengun, I.Y., Karabiyikli, S., 2011. Importance of acetic acid bacteria in food industry. *Food Control* 22, 647–656.
- Seshasayee, A.S., Bertone, P., Fraser, G.M., Luscombe, N.M., 2006. Transcriptional regulatory networks in bacteria: from input signals to output responses. *Current Opinion in Microbiology, Antimicrobials/Genomics* 9, 511–519.
- Sharma, P., Jha, A.B., Dubey, R.S., Pessarakli, M., 2012. Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *Journal of Botany* 2012, 1–26.
- Shimoda, C., Itadani, A., Sugino, A., Furusawa, M., 2006. Isolation of thermotolerant mutants by using proofreading-deficient DNA polymerase delta as an effective mutator in *Saccharomyces cerevisiae*. *Genes and Genetic Systems* 81, 391–397.
- Sicard, D., Legras, J.-L., 2011. Bread, beer and wine: yeast domestication in the *Saccharomyces sensu stricto* complex. *Comptes Rendus Biologies* 334, 229–236.
- Sikora, P., Chawade, A., Larsson, M., Olsson, J., Olsson, O., 2011. Mutagenesis as a tool in plant genetics, functional genomics, and breeding. *International Journal of Plant Genomics* 2011, 1–13.
- Sindhu, I., Chhibber, S., Capalash, N., Sharma, P., 2006. Production of cellulase-free xylanase from *Bacillus megaterium* by solid state fermentation for biobleaching of pulp. *Current Microbiology* 53, 167–172.
- Smith, D.R., 2017. Goodbye genome paper, hello genome report: the increasing popularity of “genome announcements” and their impact on science. *Briefings in Functional Genomics* 16, 156–162.
- Solera, D., Arengi, F.L.G., Woelk, T., Galli, E., Barbieri, P., 2004. TouR-mediated effector-independent growth phase-dependent activation of the 54 Ptou promoter of *Pseudomonas stutzeri* OX1. *Journal of Bacteriology* 186, 7353–7363.
- Sowden, R.J., Yasmin, S., Rees, N.H., Bell, S.G., Wong, L.-L., 2005. Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450cam and P450BM-3. *Organic and Biomolecular Chemistry* 3, 57–64.
- St. John, P.C., Bomble, Y.J., 2019. Approaches to computational strain design in the multiomics era. *Frontiers in Microbiology* 10., 597.

- Stammen, S., Müller, B.K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E., Jahn, D., 2010. High-yield intra- and extracellular protein production using *Bacillus megaterium*. *Applied and Environmental Microbiology* 76, 4037–4046.
- Steiger, S., Perez-Fons, L., Cutting, S.M., Fraser, P.D., Sandmann, G., 2015. Annotation and functional assignment of the genes for the C30 carotenoid pathways from the genomes of two bacteria: *Bacillus indicus* and *Bacillus firmus*. *Microbiology* 161, 194–202.
- Steiger, S., Perez-Fons, L., Fraser, P.D., Sandmann, G., 2012. Biosynthesis of a novel C30 carotenoid in *Bacillus firmus* isolates. *Journal of Applied Microbiology* 113, 888–895.
- Steinbüchel, A., 1995. Diversity of bacterial polyhydroxyalkanoic acids. *FEMS Microbiology Letters* 128, 219–228.
- Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 41, 207–234.
- Sudesh, K., Abe, H., Doi, Y., 2000. Synthesis, structure and properties of polyhydroxyalkanoates: biological polyesters. *Progress in Polymer Science* 25, 1503–1555.
- Suga, K.-I., Shiba, Y., Sorai, T., Shioya, S., Ishimura, F., 1990. Reaction kinetics and mechanism of immobilized penicillin acylase from *Bacillus megaterium*. *Annals of the New York Academy of Sciences* 613, 808–815.
- Szcebara, F.M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E., Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R., Pompon, D., Dumas, B., 2003. Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nature Biotechnology* 21, 143–149.
- Tabor, J.J., Levskaya, A., Voigt, C.A., 2011. Multichromatic control of gene expression in *Escherichia coli*. *Journal of Molecular Biology* 405, 315–324.
- Takaichi, S., Inoue, K., Akaike, M., Kobayashi, M., Oh-oka, H., Madigan, M.T., 1997. The major carotenoid in all known species of heliobacteria is the C30 carotenoid 4,4'-diaponeurosporene, not neurosporene. *Archives of Microbiology* 168, 277–281.
- Takaichi, S., Ishitsu, J., Seki, T., Fukada, S., 1990. Carotenoid pigments from *Rhodococcus rhodochromus* RNMS1: two monocyclic carotenoids, a carotenoid monoglycoside and carotenoid glycoside monoesters. *Agricultural and Biological Chemistry* 54, 1931–1937.
- Takaichi, S., Mochimaru, M., 2007. Carotenoids and carotenogenesis in cyanobacteria: unique ketocarotenoids and carotenoid glycosides. *Cellular and Molecular Life Sciences*. 64, 2607–2619.
- Takaichi, S., Shimada, K., 1992. Characterization of carotenoids in photosynthetic bacteria, in: *Methods in Enzymology*. Elsevier, pp. 374–385.
- Takaichi, S., Tsuji, K., Matsuura, K., Shimada, K., 1995. A monocyclic carotenoid glucoside ester is a major carotenoid in the green filamentous bacterium *Chloroflexus aurantiacus*. *Plant and Cell Physiology* 36, 773–778.
- Taniguchi, H., Okano, K., Honda, K., 2017. Modules for in vitro metabolic engineering: pathway assembly for bio-based production of value-added chemicals. *Synthetic and Systems Biotechnology* 2, 65–74.

- Tao, L., Schenzle, A., Odom, J.M., Cheng, Q., 2005. Novel carotenoid oxidase involved in biosynthesis of 4,4'-diapolycopene dialdehyde. *Applied and Environmental Microbiology* 71, 3294–3301.
- Tarshis, L.C., Proteau, P.J., Kellogg, B.A., Sacchettini, J.C., Poulter, C.D., 1996. Regulation of product chain length by isoprenyl diphosphate synthases. *Proceedings of the National Academy of Sciences of the U.S.A.* 93, 15018–15023.
- Taylor, W.E., Straus, D.B., Grossman, A.D., Burton, Z.F., Gross, C.A., Burgess, R.R., 1984. Transcription from a heat-inducible promoter causes heat shock regulation of the sigma subunit of *E. coli* RNA polymerase. *Cell* 38, 371–381.
- Tobias, A.V., Arnold, F.H., 2006. Biosynthesis of novel carotenoid families based on unnatural carbon backbones: a model for diversification of natural product pathways. *Biochimica et Biophysica Acta* 1761, 235–246.
- Tong, W.-Y., Dong, X., 2009. Microbial biotransformation: recent developments on steroid drugs. *Recent patents on biotechnology* 3, 141–153.
- Torregrosa-Crespo, J., Montero, Z., Fuentes, J.L., Reig García-Galbis, M., Garbayo, I., Vilchez, C., Martínez-Espinosa, R.M., 2018. Exploring the valuable carotenoids for the large-scale production by marine microorganisms. *Marine Drugs* 16.
- Tseng, C.K., Marquez, V.E., Milne, G.W., Wysocki, R.J., Mitsuya, H., Shirasaki, T., Driscoll, J.S., 1991. A ring-enlarged oxetanocin A analogue as an inhibitor of HIV infectivity. *Journal of Medicinal Chemistry* 34, 343–349.
- van der Voort, M., Kuipers, O.P., Buist, G., de Vos, W.M., Abee, T., 2008. Assessment of CcpA-mediated catabolite control of gene expression in *Bacillus cereus* ATCC 14579. *BMC Microbiology* 8, 62.
- van Vugt-Lussenburg, B.M.A., Damsten, M.C., Maasdijk, D.M., Vermeulen, N.P.E., Commandeur, J.N.M., 2006. Heterotropic and homotropic cooperativity by a drug-metabolising mutant of cytochrome P450 BM3. *Biochemical and Biophysical Research Communications* 346, 810–818.
- Vary, P., 1992. Development of genetic engineering in *Bacillus megaterium*. *Biotechnology* 22, 251–310.
- Vary, P.S., 1994. Prime time for *Bacillus megaterium*. *Microbiology* 140, 1001–1013.
- Vary, P.S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W.-D., Jahn, D., 2007. *Bacillus megaterium*--from simple soil bacterium to industrial protein production host. *Applied Microbiology and Biotechnology* 76, 957–967.
- Vattekkatte, A., Garms, S., Brandt, W., Boland, W., 2018. Enhanced structural diversity in terpenoid biosynthesis: enzymes, substrates and cofactors. *Organic and Biomolecular Chemistry* 16, 348–362.
- Vogl, T., Kickenweiz, T., Pitzer, J., Sturmberger, L., Weninger, A., Biggs, B.W., Köhler, E.-M., Baumschlager, A., Fischer, J.E., Hyden, P., Wagner, M., Baumann, M., Borth, N., Geier, M., Ajikumar, P.K., Glieder, A., 2018. Engineered bidirectional promoters enable rapid multi-gene co-expression optimization. *Nature Communications* 9, 3589.
- Wachtmeister, J., Rother, D., 2016. Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale. *Current Opinion in Biotechnology* 42, 169–177.

- Wackenroder, H., 1831. Ueber das Oleum radices Dauci aetherum, das Carotin, den Carotenzucker und den officinellen succus Dauci; so wie auch über das Mannit, welches in dem Möhrensaft durch eine besondere Art der Gährung gebildet wird. Geigers Mag. Pharm., pp. 144–172.
- Waldie, T., McCulloch, H., Leyser, O., 2014. Strigolactones and the control of plant development: lessons from shoot branching. *Plant Journal*. 79, 607–622.
- Wang, C., Zhao, S., Shao, X., Park, J.-B., Jeong, S.-H., Park, H.-J., Kwak, W.-J., Wei, G., Kim, S.-W., 2019. Challenges and tackles in metabolic engineering for microbial production of carotenoids. *Microbial Cell Factories* 18, 55.
- Wang, W., Hollmann, R., Fürch, T., Nimtz, M., Malten, M., Jahn, D., Deckwer, W.-D., 2005. Proteome analysis of a recombinant *Bacillus megaterium* strain during heterologous production of a glucosyltransferase. *Proteome Science* 3, 4.
- Wang, W., Sun, J., Hollmann, R., Zeng, A.-P., Deckwer, W.-D., 2006. Proteomic characterization of transient expression and secretion of a stress-related metalloprotease in high cell density culture of *Bacillus megaterium*. *Journal of Biotechnology* 126, 313–324.
- Weinhandl, K., Winkler, M., Glieder, A., Camattari, A., 2014. Carbon source dependent promoters in yeasts. *Microbial Cell Factories* 13, 5.
- Weusthuis, R.A., Lamot, I., van der Oost, J., Sanders, J.P.M., 2011. Microbial production of bulk chemicals: development of anaerobic processes. *Trends in Biotechnology* 29, 153–158.
- Whitehouse, C.J.C., Bell, S.G., Wong, L.-L., 2012. P450BM3 (CYP102A1): connecting the dots. *Chemical Society Reviews* 41, 1218–1260.
- Wieland, B., Feil, C., Gloria-Maercker, E., Thumm, G., Lechner, M., Bravo, J.M., Poralla, K., Götz, F., 1994. Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. *Journal of Bacteriology* 176, 7719–7726.
- Wierman, M.E., 2007. Sex steroid effects at target tissues: mechanisms of action. *Advances in Physiology Education* 31, 26–33.
- Wilson, D.M., Putzrath, R.M., Wilson, T.H., 1981. Inhibition of growth of *Escherichia coli* by lactose and other galactosides. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 649, 377–384.
- Winterhalter, P., Rouseff, R., 2001. Carotenoid-derived aroma compounds: an introduction, in: Winterhalter, P., Rouseff, R.L. (Eds.), *Carotenoid-Derived Aroma Compounds*. American Chemical Society, Washington, DC, pp. 1–17.
- Winterhalter, P., 1996. Carotenoid-derived aroma compounds: biogenetic and biotechnological aspects, in: Takeoka, G.R., Teranishi, R., Williams, P.J., Kobayashi, A. (Eds.), *Biotechnology for Improved Foods and Flavors*. American Chemical Society, Washington, DC, pp. 295–308.
- Wittchen, K.-D., Meinhardt, F., 1995. Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Applied Microbiology and Biotechnology* 42, 871–877.
- Wriessnegger, T., Augustin, P., Engleder, M., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Zellnig, G., Schwab, H., Pichler, H., 2014. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. *Metabolic Engineering* 24, 18–29.
- Xu, P., Li, L., Zhang, F., Stephanopoulos, G., Koffas, M., 2014. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *PNAS* 111, 11299–11304.

- Xue, D., Abdallah, I.I., de Haan, I.E.M., Sibbald, M.J.J.B., Quax, W.J., 2015. Enhanced C30 carotenoid production in *Bacillus subtilis* by systematic overexpression of MEP pathway genes. *Applied Microbiology and Biotechnology* 99, 5907–5915.
- Yabuzaki, J., 2017. Carotenoids database: structures, chemical fingerprints and distribution among organisms. Database (Oxford) 2017.
- Yang, C.-Y., Cheng, B.-H., Hsu, T.-Y., Chuang, H.-Y., Wu, T.-N., Chen, P.-C., 2002. Association between petrochemical air pollution and adverse pregnancy outcomes in Taiwan. *Archives of Environmental Health* 57, 461–465.
- Yang, J., Guo, L., 2014. Biosynthesis of β -carotene in engineered *E. coli* using the MEP and MVA pathways. *Microbial Cell Factories* 13, 160.
- Yao, K., Xu, L.-Q., Wang, F.-Q., Wei, D.-Z., 2014. Characterization and engineering of 3-ketosteroid- Δ 1-dehydrogenase and 3-ketosteroid-9 α -hydroxylase in *Mycobacterium neoaurum* ATCC 25795 to produce 9 α -hydroxy-4-androstene-3,17-dione through the catabolism of sterols. *Metabolic Engineering* 24, 181–191.
- Yatsunami, R., Ando, A., Yang, Y., Takaichi, S., Kohno, M., Matsumura, Y., Ikeda, H., Fukui, T., Nakasone, K., Fujita, N., Sekine, M., Takashina, T., Nakamura, S., 2014. Identification of carotenoids from the extremely halophilic archaeon *Haloarcula japonica*. *Frontiers in Microbiology* 5.
- Yokoyama, A., Sandmann, G., Hoshino, T., Adachi, K., Sakai, M., Shizuri, Y., 1995. Thermozeaxanthins, new carotenoid-glycoside-esters from thermophilic eubacterium *Thermus thermophilus*. *Tetrahedron Letters* 36, 4901–4904.
- Yoshida, K., Ueda, S., Maeda, I., 2009. Carotenoid production in *Bacillus subtilis* achieved by metabolic engineering. *Biotechnology Letters* 31, 1789.
- Young, A.J., Lowe, G.L., 2018. Carotenoids - Antioxidant Properties. *Antioxidants (Basel)* 7.
- Yu, X., Xu, J., Liu, X., Chu, X., Wang, P., Tian, J., Wu, N., Fan, Y., 2015. Identification of a highly efficient stationary phase promoter in *Bacillus subtilis*. *Scientific Reports* 5, 18405.
- Zehentgruber, D., Hannemann, F., Bleif, S., Bernhardt, R., Lütz, S., 2010. Towards preparative scale steroid hydroxylation with cytochrome P450 monooxygenase CYP106A2. *Chembiochem* 11, 713–721.
- Zhang, D., Zhang, R., Zhang, J., Chen, L., Zhao, C., Dong, W., Zhao, Q., Wu, Q., Zhu, D., 2014. Engineering a hydroxysteroid dehydrogenase to improve its soluble expression for the asymmetric reduction of cortisone to 11 β -hydrocortisone. *Applied Microbiology and Biotechnology* 98, 8879–8886.
- Zhang, J., Zhou, T., 2014. Promoter-mediated transcriptional dynamics. *Biophysical Journal* 106, 479–488.
- Zhang, Y., Cronan, J.E., 1998. Transcriptional analysis of essential genes of the *Escherichia coli* fatty acid biosynthesis gene cluster by functional replacement with the analogous *Salmonella typhimurium* gene cluster. *Journal of Bacteriology* 180, 3295–3303.
- Zhao, L., Chang, W., Xiao, Y., Liu, H., Liu, P., 2013. Methylerythritol phosphate pathway of isoprenoid biosynthesis. *Annual Review of Biochemistry* 82, 497–530.

- Zhou, S., Alper, H.S., 2019. Strategies for directed and adapted evolution as part of microbial strain engineering. *Journal of Chemical Technology & Biotechnology* 94, 366–376.
- Zile, M.H., Cullum, M.E., 1983. The function of vitamin A: current concepts. *Proceedings of the Society for Experimental Biology and Medicine* 172, 139–152.