Progress in Drug Research

Natural Compounds as Drugs, Volume I

Vol. 65

Frank Petersen
René Amstutz
Editors

Birkhäuser

Progress in Drug Research

Founded by Ernst Jucker

Progress in Drug Research

Natural Compounds as Drugs

Volume I

Vol. 65

Edited by
Frank Petersen and René Amstutz

Birkhäuser
Basel · Boston · Berlin

Editors

Frank Petersen
René Amstutz
Novartis Pharma AG
Lichtstrasse 35
4056 Basel
Switzerland

Library of Congress Control Number: 2007934728

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the internet at http://dnb.ddb.de

ISBN 978-3-7643-8098-4 Birkhäuser Verlag AG, Basel – Boston – Berlin

Basel · Boston · Berlin
P.O. Box 133, CH-4010 Basel, Switzerland
Part of Springer Science+Business Media
Printed on acid-free paper produced from chlorine-free pulp. TCF ∞
Cover design and layout: Micha Lotrovsky, CH-4106 Therwil, Switzerland
Printed in Germany
ISBN 978-3-7643-8098-4 e-ISBN 978-3-7643-8117-2

9 8 7 6 5 4 3 2 1 www.birkhauser.ch

Contents

Foreword

The use of herbal medicines, their preparation and application in human therapy, was described for the first time in history by the Sumerians approximately 5000 years ago. Even earlier in India and China, a highly sophisticated medical system with plant-based therapies had been developed. Along with the rise of the later empires, this knowledge was further expanded and institutionalized, while moving westwards to Egypt, Greece, and to the Arabic world, having a considerable influence on human therapy – until today.

Not all of these traditional medicines, although partially in medical practice still today, can be considered as efficacious or useful according to Western therapeutic standards, often due to the fact that the traditional disease categories cannot be easily translated into the "language" of Western medicine. The identification of active principles from plant preparations often failed, as the biological activity could not be enriched. However, there are numerous examples like morphine, quinine, salicylic acid, rivastigmine, digitalis glycosides, reserpine, or artemisinine, where the correlation between plant extract and the modern medicine, based on a single entity drug, was successfully established. All of these natural products originated from traditional medical practice and opened the door to today's medical indication areas, fundamentally enriching our knowledge of the pathophysiology and underlying biochemistry of diseases.

The successful market introduction of a drug requires a continuous financial engagement over approx. 12 years and finally an investment of US $ 1 to 1.5 billion, including capital costs. Although these figures steeply increased during the last decade and continue to follow this trend, the success of the pharmaceutical industry as mirrored in the numbers of filed new molecular entities (NMEs) and the development of new therapies for unmet medical needs can be considered at best modest. The investments during the last 15 years in innovative technologies like the "omics" world for target finding, target validation and the strongly improved capabilities to populate the chemical space by combinatorial concepts, still need to deliver on the financial figures.

During the same time period, research in natural products has largely been abandoned in drug discovery and development, mainly by the US

American and British pharmaceutical industry, and has been reduced in almost all other companies – at a first glance for some good reasons.

The molecules from nature are a weak starting point in today's high throughput screening and lead development processes. The extremely time-sensitive process from uHTS, the capability to validate millions of compounds on a given target and the identification of suitable starting points for medicinal chemists was squeezed together to a few weeks. Natural products were hardly adjustable to these new principles of lead identification. They are tested either as enriched mixtures, whereby the identification of the biologically active compound is achieved in intolerably repetitive and cumbersome process steps, or as pure compounds, often limited with respect to numbers and amounts. The vast majority of the natural products in the screening libraries are mainly singletons, unsuited for any hit explosion efforts. The chemical complexity of a natural product hampers also the engagement of medicinal chemists in industry. Synthetic molecules are often understood as a more auspicious alternative to reach the number of derivatives, as defined and accordingly rewarded in a project-driven industrial environment. A HTS campaign yields a variety of synthetic compound families interacting with the target of interest. Therefore a structure-activity relationship is immediately available with the submission of a hit list.

There are, however, solid arguments to leverage the molecules from nature in drug discovery. Besides the impressive success story in the pharmaceutical industry, natural products cover a unique breadth of biologically functionalized, complementary chemical diversity in comparison to other library types, are still sources for new pharmacophore classes, or represent a promising starting point to intervene with protein-protein interactions, to mention only a few opportunities to address current areas of pressing need in drug discovery.

With the books *Natural Products as Drugs Vol I and II* we aim to show not only the complex technology catalogue applicable to natural products in drug development-related science. Important additional motivations for the current edition have been to demonstrate how natural products can successfully be integrated in most technologies and concepts of modern drug discovery and to illustrate their potential for innovative treatments of human diseases.

The chapters cover topics from applied biodiversity considerations in natural products discovery, screening concepts, dereplication technologies

to the fermentation aspects of microbially and plant-derived secondary metabolites, *conditiones sine qua non*. These proven technologies require experts in ecology, botany, microbiology, genetics, physiology, fermentation technologies, chemistry, etc. to be effective and successful in the search for new drugs from nature. New developments and understandings thus contribute to an improved economy of the natural products deliverable in industrial research, the "magic triangle" of "faster, quicker, cheaper".

Efficient physiological investigations aiming at titer improvement of a microbially or plant-derived natural product or the generation of further derivatives play an essential role in avoiding a shortage in supply, often a motivation to disregard natural products in the lead selection discussions. Without a reliable and efficient supply of sufficient material, the prospects of natural products will be compromised and consequently, they will be uncoupled from the quickly emerging and pursued new directions in the science of drug discovery.

We compiled recent developments in the understanding of genetics principles, how nature created the generally accepted broad coverage of chemical space, and in the deciphering of silent gene clusters. Both research directions allow us an in-depth insight into the genetic organization of biosynthetic genes, generating new opportunities in pathway engineering to enlarge chemical diversity by biosynthetic manipulations.

Furthermore, the systematic application of whole animals such as zebrafish or yeast mutants as phenotype screens in combinations with biochemical approaches is increasingly used for target or pathway identification. This fascinating new avenue in drug discovery should improve the classical limitations of secondary metabolites, phenotypic activity of unknown underlying mechanisms of action, and will ultimately attract more medicinal chemists.

In the last years, natural products have been mainly omitted from virtual screening, due to their complexity, the frequent lack of exact stereochemical description, or insufficient computing power. However, more and more scientists from academia find ways out of these initial restrictions and successfully apply natural products to the development of new pharmacophore models.

The diversity of natural products is increasingly used as examples to generate "natural product-like" libraries. Herewith, the inclusion of natural products in cheminformatics investigations is essential to understand

underlying principles of structural characteristics of natural products. The interpretation of these analyses yields new ideas to mimic arrangements, combinations, and chemistry of functionalities of a natural product, representing a fascinating new facet in most recent chemical library design.

It is a surprising and encouraging phenomenon that even after their relegation in big pharma, many new drugs or drug candidates still originate from natural products or derivatives thereof. We count an impressive number of new clinical candidates and an almost stable number of patent applications over the years (*Nature Reviews Drug Discovery* March 2005, volume 5), contradicting the general notion of their non-druglike features and confirming that natural products are exempt from the Ro5, as repeatedly communicated by Chris Lipinsky.

The edition *Natural Products as Drugs* is concluded with dedicated monographs of natural products or compounds based on them, representing potential or already proven ways for innovative medical intervention.

Hence most recent enabling technologies and innovative chemical derivation concepts have substantially changed the classical natural products research during the last decade, giving us new possibilities to exploit the huge diversity designed and generated by nature, of which only the tip of the iceberg seems to have been discovered yet. Natural products expand the molecular armamentarium for therapeutic choices and may improve the moderate success of today's applied chemical diversity. Together with the understanding of their target proteins, they offer chances of complementary routes in drug discovery, great opportunity to better understand the causality of diseases, and will ultimately lead to new drugs for medical needs.

We are grateful to Birkhäuser Verlag for editing this volume. We also wish to acknowledge the excellent cooperation with Beatrice Menz, Anke Brosius, and Yvonne Schlereth for their assistance and seamless coordination. Most importantly, we would like to thank all chapter authors for their enthusiasm in the idea of the book, for their valuable time and knowing patience.

Frank Petersen
René Amstutz

Glossary

ACE	angiotensin converting enzyme
AKBA	acetyl-11-keto-β-boswellic acid
ALA	AIDS Research Alliance of America
API	atmospheric pressure ionization
BBE	berberine bridge enzyme
BIA	benzylisoquinoline alkaloids
BVDV	bovine viral diarrhea virus
CAP	community-acquired pneumonia
CBD	Convention on Biological Diversity
CCD	central composite design
CDAD	*Clostridium difficile*-associated diarrhoea
CGRP	calcitonin gene-related peptide
CML	chronic myeloid leukaemia
CMV	cauliflower mosaic virus
CNRS	Centre de la Recherche Scientifique
COPD	chronic obstructive pulmonary disorder
cSSSi	complicated skin and skin structure infection
CTCL	cutaneous T-cell lymphoma
DAOC	deacetoxycephalosporin
DHA	docosahexaenoic acid
DO	dissolved oxygen
EGCG	epigallocatechin gallate
ELSD	evaporative light scattering detector
EMEA	European Medicines Agency
EMS	ethylmethane sulfonate
ESI	electrospray ionization
FAC	frontal affinity chromatography
FAS	fatty acid synthase
FTICR	Fourier transform ion cyclotron resonance
FTMS	Fourier transform mass spectroscopy
FXR	farnesoid X receptor
GPCR	G-protein coupled receptor
GSK	GlaxoSmithKline
H6H	hyoscyamine 6B-hydroxylase
HA	hydroxylamine
HCV	hepatitis C virus
HGT	horizontal gene transfer
HMG-CoA	hydroxymethylglutaryl coenzyme A
HPI	high pathogenicity island
HSQC	heteronuclear single quantum correlation spectroscopy
HSV	herpes simplex virus

HTS	high throughput screening
ICR	ion cyclotron resonance
LOX	lipoxygenase
MAA	marketing authorization approval
MASS	multitarget affinity/specificity screening
MetAP2	methionine aminopeptidase 2
MMS	methylmethane sulfonate
MMV	Medicines for Malaria Venture
MOA	mechanism of action
MRSA	methicillin-resistant *Staphylococcus aureus*
NCI	National Cancer Institute
ND	natural-derived
NDA	new drug application
NME	new molecular entity
NP	natural product
NRP	non-ribosomal peptide
NRPS	non-ribosomal peptide synthase
NTG	nitrosoguanidine
OFAGE	orthogonal field alteration gel electrophoresis
ORF	once functional cluster
PABA	*para*-aminobenzoic acid
PASS	prediction of activity spectra for substances
PCA	principal component analysis
PDT	photodynamic therapy
PKS	polyketide synthase
PMT	putrescine N-methyltransferase
PRSP	penicillin-resistant *Streptococcus pneumoniae*
QHTS	quantitative high throughput screening
RCC	renal cell carcinoma
RTX	resiniferatoxin
SAHA	suberoylanilide hydroxamic acid
SAR	structure activity relationship
SERM	selective estrogen receptor modulator
STD	saturation transfer difference
TCMD	Traditional Chinese Medicinal Database
TCTP	translationally controlled tumor protein
tNOX	tumor-associated NADH oxidase
TROSY	transverse relaxation-optimized spectroscopy
TRP	transient receptor potential
VEGF	vascular endothelial growth factor
VOC	volatile organic substances
VRE	vancomycin-resistant enterococci
XIAP	X-linked inhibitor of apoptosis

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Mother Nature's gifts to diseases of man: the impact of natural products on anti-infective, anticholestemics and anticancer drug discovery

By Mark S. Butler[1] and David J. Newman[2]

[1]MerLion Pharmaceuticals, 1 Science Park Road, The Capricorn #05-01, Singapore Science Park II, Singapore 117528

[2]Natural Products Branch, Developmental Therapeutics Program, NCI-Frederick, Fairview Center Suite 206, P.O. Box B, Frederick, MD 21702, USA
<dn22a@nih.gov>

Abstract

This chapter is designed to demonstrate that compounds derived from nature are still in the forefront of drug discovery in diseases such as microbial and parasitic infections, carcinomas of many types and control of cholesterol/lipids in man. In each disease area we have provided short discussions of past, present and future agents, in general only considering compounds currently in clinical Phase II or later, that were/are derived from nature's chemical skeletons. Finishing with a discussion of the current and evolving role(s) of microbes (bacteria and fungi) in the production of old and new agents ostensibly produced by higher organisms.

1 Introduction to natural products

Throughout the ages humans have relied on nature for their basic needs and not least, their medicines. Plants have formed the basis of sophisticated traditional medicine systems that have been in existence for thousands of years. The first records, written on clay tablets in cuneiform, are from Mesopotamia and date from about 2600 BCE, while Egyptian medicine dates from about 2900 BCE, with the best known Egyptian pharmaceutical record being the *Ebers Papyrus* dating from 1500 BCE [1, 2]. The Chinese *Materia Medica* has been extensively documented over the centuries, with the first record (Wu Shi Er Bing Fang), containing 52 prescriptions, dating from about 1100 BCE [3, 4] though records from the *Pent'sao* are reputed to be even earlier (~2700 BCE) and documentation of the Indian Ayurvedic system dates from about 1000 BCE (Susruta and Charaka) [5, 6]. In the ancient Western world, the Greeks contributed substantially to the rational development of the use of herbal drugs. The philosopher and natural scientist, Theophrastus (~300 BCE), in his *History of Plants*, dealt with the medicinal qualities of herbs and Dioscorides, a Greek physician (100 CE), during his travels with Roman armies, recorded the collection, storage and use of medicinal herbs. Galen (130–200 CE), who practiced and taught pharmacy and medicine in Rome, published no less than 30 books on these subjects and is well known for his complex prescriptions and formulas used in compounding drugs, sometimes containing dozens of ingredients ('galenicals').

During the Dark and Middle Ages (5th to 12th Centuries), it was the Arabs who were responsible for the preservation of much of the Greco-Roman expertise and for expanding it to include the use of their own resources, together with Chinese and Indian herbs unknown to the Greco-Roman world. The Arabs were the first to establish privately owned drug

stores in the 8th Century and the Persian pharmacist, physician, philosopher and poet, Avicenna, contributed much to the sciences of pharmacy and medicine through works such as *Canon Medicinae*, regarded as "the final codification of all Greco-Roman medicine". A comprehensive review of the history of medicine may be found on the National Library of Medicine's 'History of Medicine' homepage [7].

That natural products (NPs) are still 'alive and well' as both direct and indirect sources of leads to drugs against all classes of disease is shown quite dramatically in Figure 1, which is taken from the third review in the series by Newman et al., covering sources of drugs approved against all diseases by the Food and Drug Administration (FDA) or their equivalents in other countries from 1 January 1981 to 30 June 2006 [8]. The influence of natural products directly (N) or slight modifications (ND) are quite obvious, and when their use as sources of pharmacophores (or privileged structures that may be utilized as isosteres of naturally occurring chemical skeletons, S*; S*/NM and S/NM being the classifications for these [9]) then the continued influence is quite striking.

Rather than the customary chemical usage of discussing compounds subdivided by chemical structures (skeletons) we have elected to classify by disease classes. One major reason for this is that nowadays it is quite common for a drug to be developed for one disease and then to find its 'niche' as a treatment for another separate disease or even to spawn derivatives that have yet a third or fourth potential area in which to be useful.

An excellent example of this is in the case of the fungal secondary metabolite, sirolimus (rapamycin) (**1**). Originally developed as an antifungal agent by Ayerst (now Wyeth), it was approved as an immunosuppressive drug (Rapamune®) and a derivative, temsirolimus (**2**), is in a multiplicity of trials against a variety of cancers (see Section 4).

As a result, we are discussing natural products and their derivatives, including modified nucleosides, some peptides depending upon the disease area and a number of drugs that contain the 'warhead of the natural product' but have modified lipophilic structures giving different pharmacodynamics/pharmacokinetics (i.e., the variations on mevastatin (**3**) that led to the formally synthetic anticholesterolemics such as atorvastatin (**4**) (see Section 3).

The format of each section will provide a brief historical coverage, followed by drugs in current use and a short description of each NP or NP-

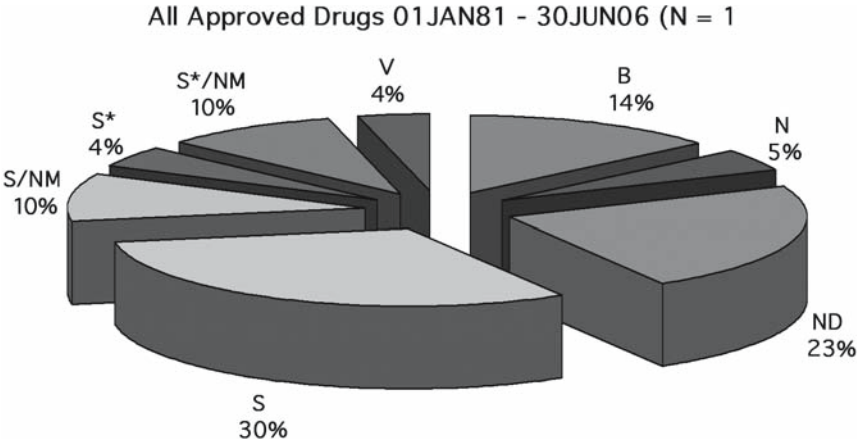


Figure 1.
Drugs approved against all diseases by the FDA or their equivalents in other countries from the 1 January 1981 to 30 June 2006.

derived compound in Phase II or Phase III clinical trials or undergoing drug registration. Due to space limitations, only NP or NP-derived compounds in Phase III clinical trials or undergoing drug registration are described for oncology (Section 4). The clinical status of each compound is correct to the end of October 2006. For those people interested in an in-depth listing of NP-derived compounds in clinical trials in all therapeutic areas should consult the reviews by Butler [10] and Kinghorn and co-workers [11]. Also worth reading are the two reviews on the role of NPs in drug discovery today [12, 13] and Sneader's book on the history of drug discovery [14].

2 Anti-infectives (antibacterials, antifungals and antiparasitics)

2.1 Antibacterials

The advent of the antibacterial era is often reckoned to be from the discovery of penicillin by Fleming in 1928 (and reported in the British Medical Literature in 1929), though there were anecdotal reports of earlier workers (Tyndall, Roberts and Pasteur in the 1870s) recognizing antagonism

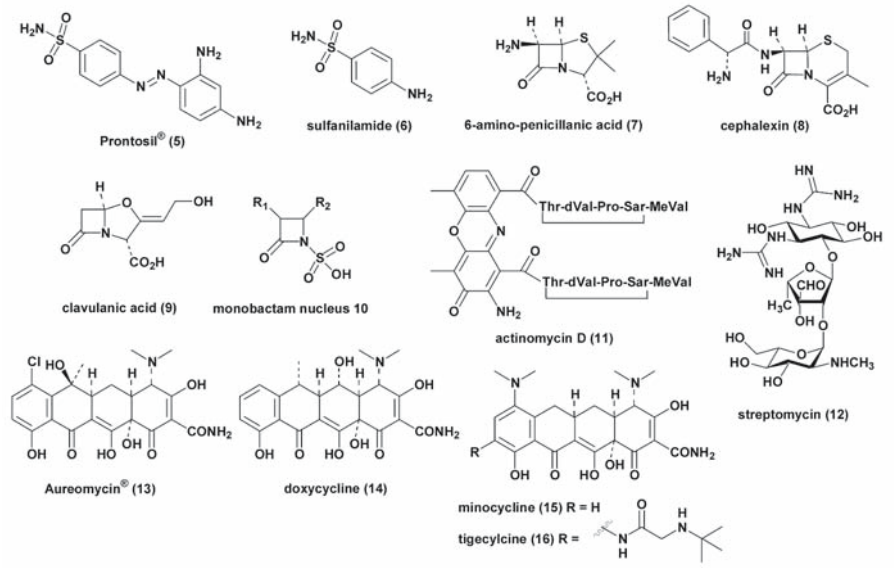
between various bacteria. The advent of the sulfonamides exemplified by Prontosil® (**5**) led to the introduction of synthetic antibacterials with the first clinical efficacy report in 1933 and ultimately leading to the award of the Nobel Prize for Medicine in 1938 to Domagk. This could also be thought of as the first formal prodrug in the antibiotic field as the active principle sulfanilamide (**6**) is a structural analogue of *para*-aminobenzoic acid (PABA). PABA competitively inhibits dihydropteroate synthase, thus leading to inhibition of folic acid and bacterial death. So although synthesized in the absence of such knowledge, and for an entirely different purpose, it was in retrospect an isostere of a NP.

We will briefly discuss the major chemical classes of natural antibiotics and give suitable references to articles which will go into much greater detail for the interested reader in each section.

2.1.1 β-Lactams

Following the isolation and identification of penicillin G and then penicillin V in the UK and the USA in the early 1940s which was covered in detail in 1998 by Mateles in an excellent reprint entitled *History of Penicillin Production* [15], the number of penicillin-based molecules that have been produced by semi- and total synthesis to date is well over the 15,000 level. Most of these compounds have started with modification of the fermentation product, 6-amino-penicillanic acid (**7**), which also can be produced by a simple chemical or biochemical deacylation from penicillins. The number above is only indicative as a significant proportion of materials were never published, particularly from industry, as they had marginal or no significant activity over those that had been reported previously.

In 1948, the ring-expanded version of penicillin, cephalosporin C, was reported from *Cephalosporium* sp. by Brotzu and its structure determined in 1961 by the Oxford Group [16, 17]. As with the penicillin nucleus, this ring expanded molecule also served as the building block (as its 7-amino-cephalosporanic acid homologue) for many thousands of cephalosporins, with the first orally-active molecule, cephalexin (**8**) being introduced in 1970. Since that time, a multitude of cephalosporins have been synthesized with the aim of producing molecules that are more resistant to β-lactamases.



In order to give extra 'medicinal life' to β-lactams that were no longer resistant to the common β-lactamases, in the late 1960s and early 1970s, efforts were made, particularly by Beecham (now part of GlaxoSmithKline) and Pfizer to find molecules that would have similar pharmacokinetics to the β-lactams but would inhibit the 'regular' β-lactamases that were part of the pathogenic microbe's defense systems. Beecham discovered the clavulanate family with clavulanic acid (**9**) being incorporated into the combination known as Augmentin® a 1:1 mixture of amoxicillin and clavulanic acid (**9**) launched in 1981, thus extending the franchise of this particular β-lactam well beyond its original patent date.

Along with the search for the β-lactamase inhibitors, efforts were underway to produce the simplest β-lactam, the monobactam. Following many years of unsuccessful research at major pharmaceutical houses, predominately in the synthetic areas, came the reports from Imada et al. in 1981 [18] and a Squibb group led by Sykes [19], who both demonstrated the same basic monobactam nucleus (**10**). What is important to realize is that no molecules synthesized before the discoveries of these NPs had a sulfonyl group attached to the lactam nitrogen, which is an excellent method for stabilizing the single ring. Since that time a significant number of variations upon that theme have been placed into clinical trials and in some cases, into commerce.

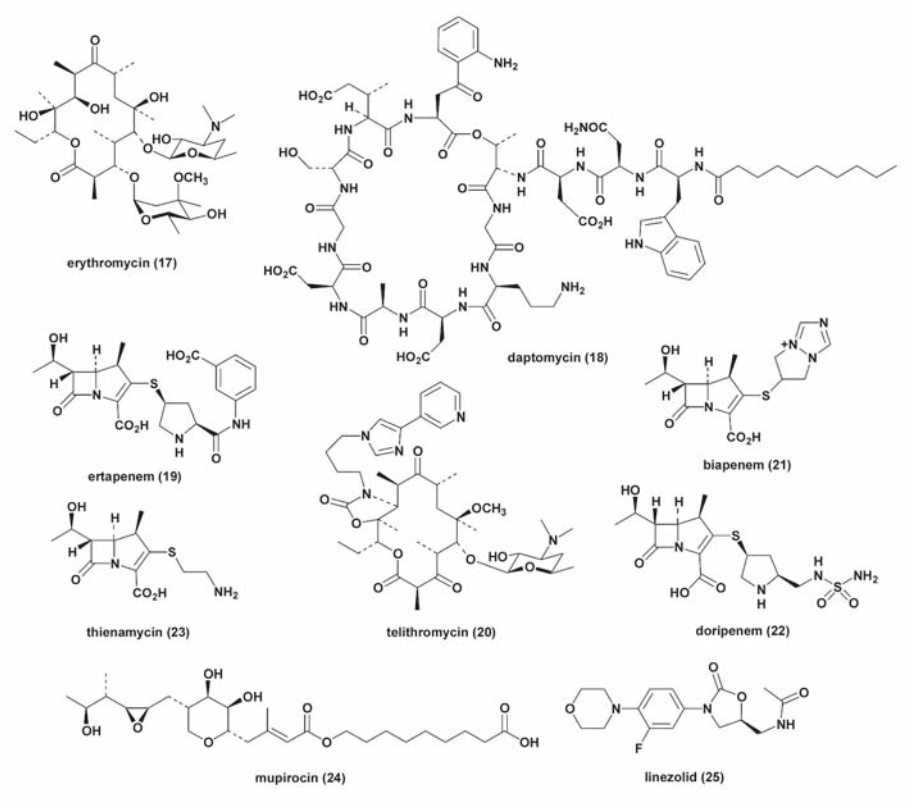
2.1.2 Actinomycins, aminoglycosides, tetracyclines and erythromycins

Concomitantly with the early development of the penicillins, Waksman was working at Rutgers University in New Jersey in the late 1930s/early 1940s, specializing in investigation of the actinomycetes (which at that time were considered to be fungi), with the aim of finding a treatment for tuberculosis. His initial finding in 1940, however, was the identification of chromooligopeptides of the actinomycin class (e.g., actinomycin D (**11**)), which though not useful as antibacterials, led to what was the first use of such secondary metabolites as a treatment for cancer (Wilms' tumor) [20].

In 1943, the aminoglycoside antibiotic streptomycin (**12**) was isolated from *Streptomyces griseus* and, in addition to being active against *Mycobacterium tuberculosis*, was active against a wide range of other pathogenic organisms. Further work over the next twenty or so years yielded a large number of similar glycosidic-based antibacterials. Unfortunately, as a group they have a major biological Achilles heel in the sense that they are easily inactivated by plasmid-mediated acetylation or phosphorylation and multiply resistant organisms have evolved. However, aminoglycosides still have utility particularly in conjunction with β-lactams with whom they exhibit true synergy.

The fourth series of molecules to be reported was a previously unknown molecule with four fused rings (a tetracycline). The parent molecule was not used to any great extent as an antibacterial but the naturally occurring chlorinated analogue, Aureomycin® (**13**), was. This tetracycline skeleton has given rise to a large number of semi-synthetic molecules with three of these, doxycycline (**14**), minocycline (**15**) and tigecycline (**16**), being used today, particularly against the causative agent of Lyme disease.

The macrolide antibiotics, exemplified by erythromycin (**17**), are as equally famous and long-lived as the other classes previously discussed. Even today, erythromycin (**17**) is still prescribed, particularly for pediatric patients. This class of antibiotics has yet another claim to fame as it was one of the first molecules for which the biogenetic system was described in 1990 using classical mutation studies [21], which later developed into the system known as combinatorial biosynthesis whereby non-naturally occurring metabolites are made by 'mixing and matching' gene clusters [22]. More current details can be found in the



recent overview by Demain and the references therein [23], together with the excellent review by Baltz et al. on the use of genetic constructs in developing further congeners of daptomycin (**18**) [24] and from a historical to current perspective, the excellent review by von Nussbaum and co-workers covering the older and modern literature from a medicinal chemistry/lead discovery and optimization aspect should definitely be consulted [25].

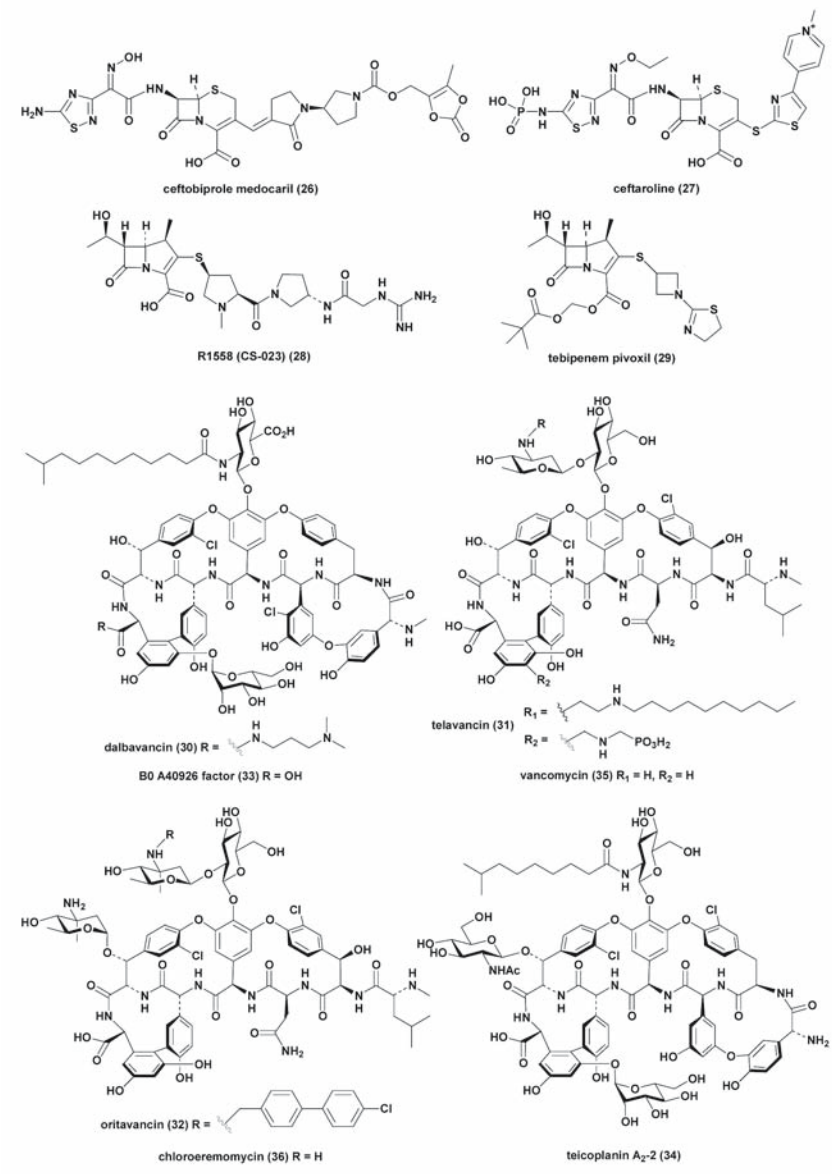
2.1.3 Antibacterials: current status

Since 2000, six new NP-derived drugs have been launched: ertapenem (2001, Invanz®, Merck) (**19**) [26, 27], telithromycin (2001, Ketek®, Sanofi-

Aventis) (**20**) [28, 29], biapenem (2002, Omegacin®, Meiji) (**21**) [30, 31], daptomycin (2003, Cubicin®, Cubist) (**18**) [24, 32], doripenem (2005, Finibax®, Shionogi & Co; Phase III (US), J&J) (**22**) [33, 34] and tigecycline (2005, Tygacil®, Wyeth) (**16**) [35–39]. Ertapenem (**19**), biapenem (**21**) and doripenem (**22**) are carbapenem antibiotics (part of the β-lactam family), which are produced synthetically but their lead structure was the NP thienamycin (**23**). Tigecycline (**16**) is a semi-synthetic derivative of tetracycline, while telithromycin (**20**) is a semi-synthetic derivative of erythromycin (**17**). Daptomycin (**18**) is a lipopeptide NP used for the treatment of complicated skin and skin structure infections (cSSSi) and *Staphylococcus aureus* bloodstream infections or bacteremia including right-sided infective endocarditis. In terms of sales, daptomycin (**18**) has had the most successful launch for an IV antibiotic in US history. Daptomycin (**18**) represents only one of three new antibiotic classes launched since 1970; the other two being the topical antibiotic NP mupirocin (**24**) in 1985 and the synthetic oxazolidinone linezolid (**25**) in 2000.

There are total of four β-lactams, two cephalosporins, ceftobripole medocaril (**26**) and ceftaroline acetate (**27**), and two carbapenems, R1558 (**28**) and tebipenem (**29**), in Phase II or Phase III clinical trials or undergoing drug registration. Ceftobiprole medocaril (**26**) is a fourth generation cephalosporin that has potent bactericidal activity against methicillin resistant *Staphylococcus aureus* (MRSA) and penicillin resistant *Streptococcus pneumoniae* (PRSP) [40]. Basilea and Johnson and Johnson Pharmaceutical Research and Development LLC (J&J) are evaluating ceftobiprole medocaril (**26**) for the treatment of cSSSi, nosocomial pneumonia and hospitalized community acquired pneumonia (CAP) in various Phase III trials. Ceftaroline (PPI-0903, TAK-599) (**27**) is being evaluated by Cerexa in Phase II trials and both ceftobiprole (**26**) and ceftaroline (**27**) have been granted FDA fast-track status [40, 41]. The carbapenems R1558 (Ro4908463, CS-023, Sankyo and Roche) (**28**) [40, 42] and tebipenem pivoxil (ME-1211, Meiji Seika) (**29**) are being evaluated in Phase II clinical trials as a broad spectrum antibiotics [43].

There are three semi-synthetic glycopeptides [32, 44, 45], dalbavancin (**30**), telavancin (**31**) and oritavancin (**32**), in late stage clinical investigation and their antibacterial mechanism is through inhibition of cell wall production. Dalbavancin (Zeven®) is a semi-synthetic derivative of B0-A40926 factor (**33**) [46], a glycopeptide related to teicoplanin (**34**), and a New Drug Application (NDA) for the treatment of skin and soft tissue



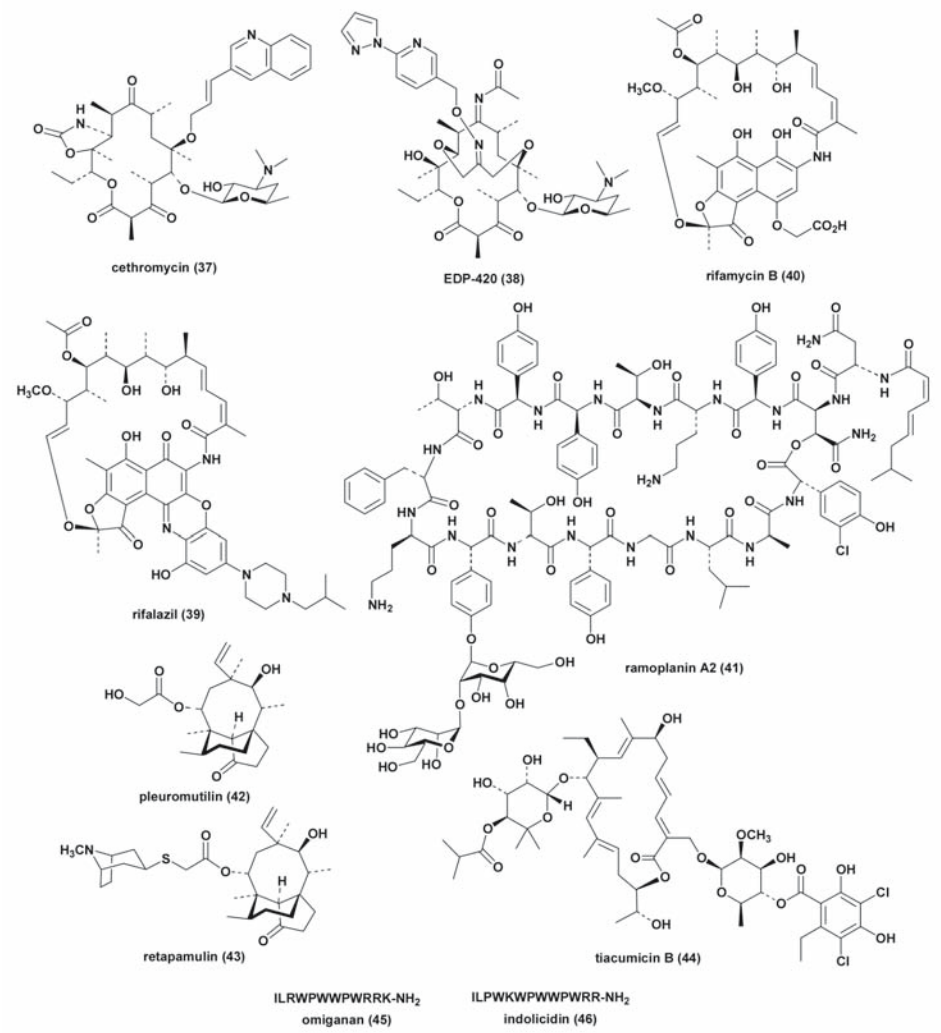
infections was filed in February 2005 by Vicuron Pharmaceuticals (now part of Pfizer). Pfizer received an Approvable Letter on 21 June 2006 from the FDA for dalbavancin (**30**) and its launch has been delayed until 2007. Telavancin (TD-6424) (**31**) [47], which is a semi-synthetic derivative of

vancomycin (**35**), is being evaluated in Phase III trials for treatment of patients whose infections are due to MRSA in both Gram-positive cSSSI and CAP by Theravance and Astellas. Oritavancin (Nuvocid™, LY-333328) (**32**) is a semi-synthetic derivative of the vancomycin analogue chloroeremomycin (**36**) [48], which Targanta Therapeutics acquired the rights to in December 2005 from InterMune, who had licensed it originally from Eli Lilly. Phase III clinical trials of oritavancin (**32**) for the treatment of cSSSI have been completed and an NDA is planned for 2007. Targanta also has Phase II clinical trials for the treatment of catheter-related bacteremia and nosocomial pneumonia in progress.

There are two semi-synthetic derivatives of the macrolide erythromycin (**17**), the ketolides cethromycin (ABT-773) (**37**) and EP-420 (EP-013420, S-013420) (**38**), in active clinical development [28, 49]. Advanced Life Sciences have been evaluating cethromycin (**37**) [50, 51] in Phase III trials for the treatment of CAP. Advanced Life Sciences licensed the worldwide rights to cethromycin (**37**) from Abbott Laboratories, except in Japan where it had been licensed previously by Taisho. EP-420 (**38**) [52, 53] is a novel, bridged bicyclic derivative developed by Enanta Pharmaceuticals, which is being evaluated in Phase II trials for treatment of CAP by both Enanta and Shionogi.

The only rifamycin derivative under late stage evaluation is rifalazil (ABI-1648, KRM-1648) (**39**), a semi-synthetic of rifamycin B (**40**) [54] developed by Kaneda Corporation [55] and licensed to ActivBiotics [56]. Rifalazil (**39**) and related compounds are bactericidal and can effectively penetrate target tissues and cells, which may enable them to be used to eliminate intracellular pathogens such as Chlamydia [57]. ActivBiotics is evaluating rifalazil (**39**) in Phase II trials for carotid artery atherosclerosis and Phase III trials for intermittent claudication associated with peripheral arterial disease, in which rifalazil (**39**) has FDA Fast Track status.

Ramoplanin is a lipopeptide antibiotic complex isolated from *Actinoplanes* sp. ATCC33076, which consists of factors A1, A2 and A3 [58, 59]. Factor A2 ('ramoplanin') (**41**) is the major component of the complex and is being evaluated in Phase II trials by Oscient Pharmaceuticals for the treatment of *Clostridium difficile*-associated diarrhoea (CDAD) [60, 61]. Ramoplanin (**41**) has been evaluated for the treatment of vancomycin-resistant *Enterococci* (VRE) but no trials are ongoing. Ramoplanin (**41**) exerts its antibacterial activity by binding to the peptidoglycan interme-



diate Lipid II (C_{35}-MurNAc-peptide-GlcNAc) and disrupting bacterial cell wall synthesis [62–64].

GlaxoSmithKline (GSK) has filed an NDA in February 2006 for the pleuromutilin (**42**) derivative, retapamulin (Altabax®, SB-275833) (**43**), for use as a topical antibiotic to treat skin infections [65, 66]. Pleuromutilins are fun-

gal-derived compounds that exert their antimicrobial activity by binding to the 50S bacterial ribosome [67]. GSK has two other pleuromutilin derivatives, 565154 and 742510, in Phase I clinical trials as oral antibiotics.

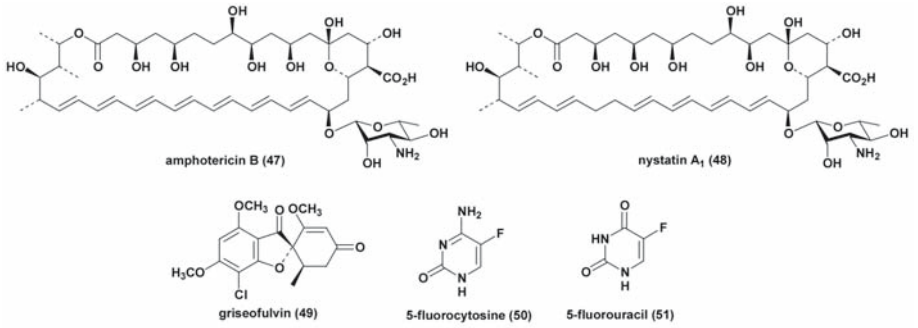
Tiacumicin B (PAR-101, OPT-80) (**44**) is the major component of the tiacumicin antibiotic complex produced by *Dactylosporangium aurantiacum* subsp. *hamdenensis* NRRL 18085, which originally was isolated by Abbott [68, 69], and is identical to lipiarmycin A3 [70] and clostomicin B1 [71]. Tiacumicin B (**44**) displays potent antibacterial activity and is being evaluated in Phase IIb/III clinical trials by Par Pharmaceuticals (and Opitmer Pharmaceuticals) for the treatment of CDAD [72–75]. Tiacumicin B (**44**) has been proposed to exert its antibacterial activity through inhibition of RNA synthesis [76].

Finally, the cationic peptide [77, 78], omiganan (**45**), which was developed by MIGENIX, has been licensed to Cadence Pharmaceuticals for catheter-related infections and Cutanea Life Sciences for dermatological diseases [79, 80]. The structure of omiganan (**45**) is based on indolicidin (**46**), a potent antibacterial and antiviral peptide originally purified from the cytoplasmic granules of bovine neutrophils. Omiganan (coded as Omigard™, CPI-226, MBI-226) (**45**) is being developed by Cadence for the treatment of catheter-related infections. Omiganan (**45**) showed promising results in a Phase III trial but its primary endpoint of a reduced rate of infections was not achieved and additional Phase III trials are underway in both the USA and Europe using a gel-based formulation. Omiganan (coded as CLS001, MX-594AN) (**45**) has completed two Phase II trials for the treatment of acne and Cutanea plan to initiate a Phase II trial for the treatment of Rosacea in 2007.

2.2 Antifungals

As fungi are eukaryotes, the number of potential targets where there are significant differences between the microbe and its host (the human or animal/plant) are much smaller compared to bacteria, as they have intrinsic differences in primary metabolic processes, as well as the architecture of their cell walls and membranes. As a consequence, a significant number of bacterial targets do not have eukaryotic counterparts.

In general, the targets are the cell walls of the fungi or subtle differences in membrane structure or in the biosynthesis of the sterol precursors lead-



ing to membrane synthesis. Although a very considerable amount of time and effort was expended in the early days of antibiotic discovery, only three general use antifungal agents were developed from that work.

Perhaps the best known clinically is the heptaene polyene, amphotericin B (**47**), originally isolated from *Streptomycetes nodosus* collected in Venezuela and reported in 1956, whose full structure was reported in two contemporaneous papers in 1970, by X-ray crystallography [81] where the absolute configuration was determined using the iodo-derivative and by mass spectroscopy [82].

Though many polyenes with varying numbers of conjugated double bonds have been reported since those early days, only one other compound of this class, in fact the first identified (in 1950) of this general structure class, the tetraene nystatin (**48**), has gone into general clinical use and like amphotericin B (**47**), its primary indication is for candidiasis. It was first reported from *Streptomyces noursei* and, as with **47**, its structure was reported in the 1970 time frame by two groups, one using classical chemical degradation plus proton NMR [83] and the other *via* mass spectroscopy [84]. The proposed hemiketal structures of amphotericin B (**47**) and nystatin (**48**) were later confirmed in Rinehart's laboratory in 1976 [85].

Probably the first clinically used antifungal NP (originally reported in 1939), has also survived the test of time, and this is griseofulvin (**49**) whose non-polyene structure was defined in a series of papers in 1952 using classical techniques [86]. Even today, over 70 years after it was first described, griseofulvin (**49**) is still in clinical use against dermatophytes. In fact,

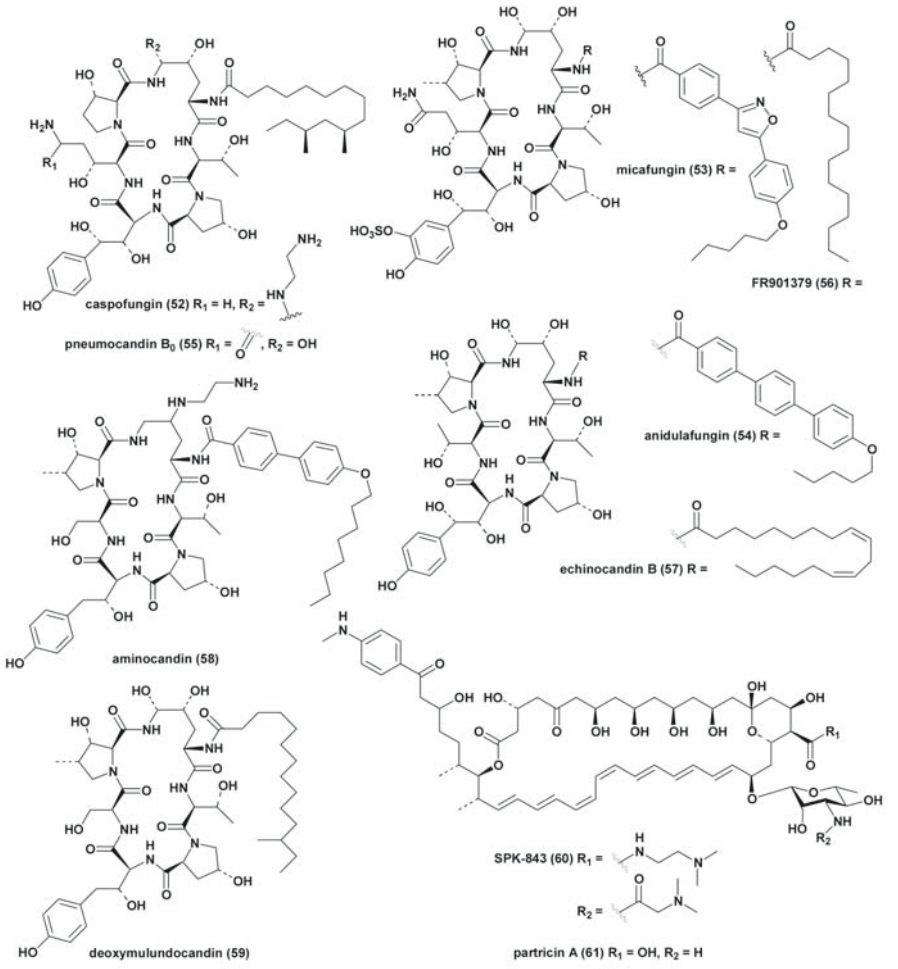
griseofulvin (**49**) is only active against this class of fungi and long-term treatment is necessary due to its insolubility.

One other NP-derived molecule also in clinical use is the modified pyrimidine nucleoside, 5-fluorocytosine (flucytosine) (**50**), which although made synthetically, can be considered to be derived from a NP. Following transport into *Candida* or *Cryptococcus via* a cytosine permease, 5-fluorocytosine (**50**) is deaminated to give 5-fluorouracil (**51**), which can interfere with both RNA and DNA metabolism. It is relatively selective for fungal cells as a result of the lack or very low levels of cytosine deaminase in mammalian cells.

2.2.1 Antifungals: current status

Three NP-derived antifungal drugs from the echinocandin class [87, 88], caspofungin (2001, Cancidas®, Merck) (**52**) [89], micafungin (2002, Mycamine®/Funguard®, Astellas) (**53**) [90] and anidulafungin (2006, Eraxis®, Pfizer) (**54**) [91], have been launched since 2000. These drugs are semi-synthetic derivatives of NPs (caspofungin (**52**)/pneumocandin B_0 (**55**), micafungin (**53**)/FR901379 (**56**) and anidulafungin (**54**)/echinocandin B (**57**)) and their antifungal activity is caused by inhibition of 1,3-β-D-glucan synthesis in the fungal cell wall. A very recent publication by Aperis and co-workers [92] details the clinical utility of these echinocandins and how they may be used for individualized therapy in conjunction with other agents (some of which are synthetic). Finally, another echinocandin, aminocandin (HMR-3270) (**58**), which is a semi-synthetic derivative of deoxymulundocandin (**59**), is currently undergoing Phase I evaluation for treatment of systemic fungal infections [93, 94].

There is only one NP or NP-derived compound in Phase II antifungal clinical trials or above. SPK-843 (**60**), the diascorbate salt of a semi-synthetic derivative of patricin A (**61**), is in Phase II clinical trials for treatment of systemic mycosis by Kaken Pharmaceuticals [95–98] and detailed information on SPK-843 is available on the Aparts BV web site [99]. Patricin A (**61**) is a 38-membered heptaene polyene member first described as part of the aureofacin antibiotic complex isolated from *Streptomyces aureofaciens*. SPK-843 (**60**) has comparable activity to amphotericin B (**47**) but is considerably more water soluble and its antifungal activity is caused by destabilization of fungal cell membrane in a similar fashion to other polyenes.



2.3 Antiparasitics

Perhaps the most important parasitic infection worldwide is malaria. One may argue that there are many other debilitating diseases caused by parasites but overall, the number of deaths annually due to malaria is over one million, with the vast majority (80% plus) being children under the age of 5 (Centers for Disease Control statistics for 2005).

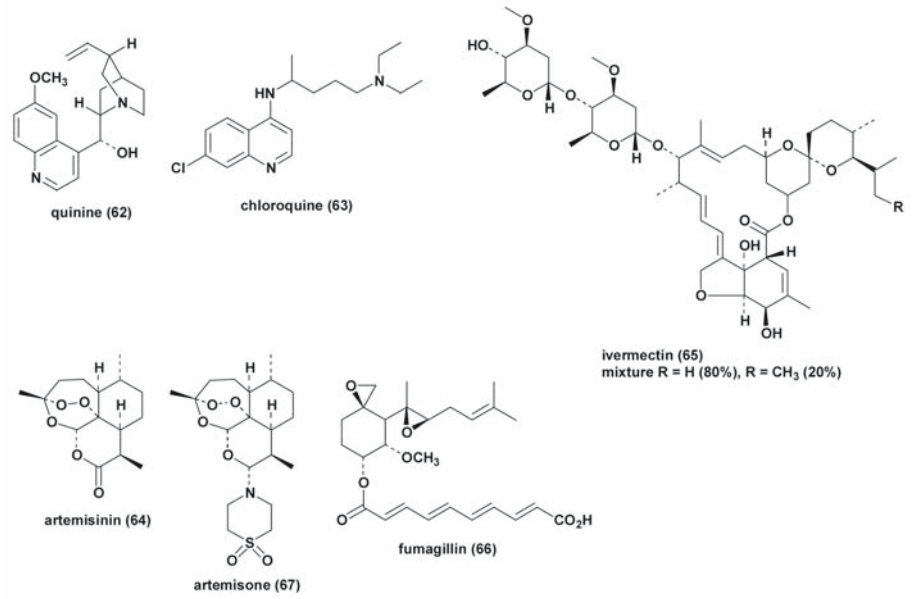
The first discovered NP used against malaria was quinine (**62**), which is isolated from the bark of *Cinchona* spp [100]. Synthetic molecules, which

contain the isoquinoline core present in quinine (**62**), have been used worldwide. The first of these analogues, chloroquine (**63**), was introduced in 1943 by the Allies, as quinine (**62**) could not be obtained for use in the Indo-Pacific theatres. Ironically, the same molecule had been synthesized in Germany in 1934 under the name Resochin but had been dropped due to toxicity. However, roughly 400 years after the recognition of quinine (**62**) came the next NP antimalarial, which was identified directly from the Chinese *Materia Medica*. This was the previously unknown molecule artemisinin (**64**), which was isolated from extracts of the Wormwood tree, *Artemisia annua*. Artemisinin (**64**) originally was crystallized by Chinese scientists in 1972 and was reported in a 1998 review to be present in other species in the same genus [101]. Over the last few years, this compound and more soluble derivatives have altered the treatment of resistant malaria. In addition, Keasling's group at Berkeley have transferred the genes from the producing plant into *E. coli* and have successfully expressed the base terpene (amorpha-4,11-diene) [102] with modification into the base structure both chemically and to some extent, biochemically *via* P450 enzymes. Though not yet economically viable when compared to direct plant collection and extraction, the process is yet to be optimized for production, and with the experience in manipulation of *E. coli*, this is definitely a feasible project. A very large amount of this work is directly funded by the Gates Foundation.

There is one other famous example of the use of a slightly modified NP as an antiparasitic and that is the use of ivermectin (Stromectol®) (**65**) in the treatment of onchocerciasis (with approximately 18 million people being affected by this disease in 1995). The donation of ivermectin (**65**) by Merck & Co since 1988 has helped to control this debilitating disease, although a yearly treatment with further ivermectin (**65**) is required for further control [103].

2.3.1 Antiparasitics: current status

Fumagillin (**66**) was isolated from *Aspergillus fumigatus* in 1949 and used shortly after its discovery to treat intestinal amoebiasis. In 2005, Sanofi-Aventis was granted approval in France for fumagillin (2005, Flisint®, Sanofi-Aventis) (**66**) for use in the treatment of intestinal microsporidiosis



[104, 105]. Intestinal microsporidiosis, which is caused by the spore-forming unicellular parasite *Enterocytozoon bieneusi*, causes chronic diarrhea and is a major concern to immunocompromised patients [106]. The anti-microsporidiosis and angiogenic activities of fumagillin (**66**) and related compounds are thought to be due to its strong methionine aminopeptidase 2 (MetAP2) binding [107, 108]. It has been shown using zebra fish models that MetAP-2 inhibition blocks noncanonical Wnt signaling, which plays a critical role in development, cell differentiation, and tumorigenesis [109, 110].

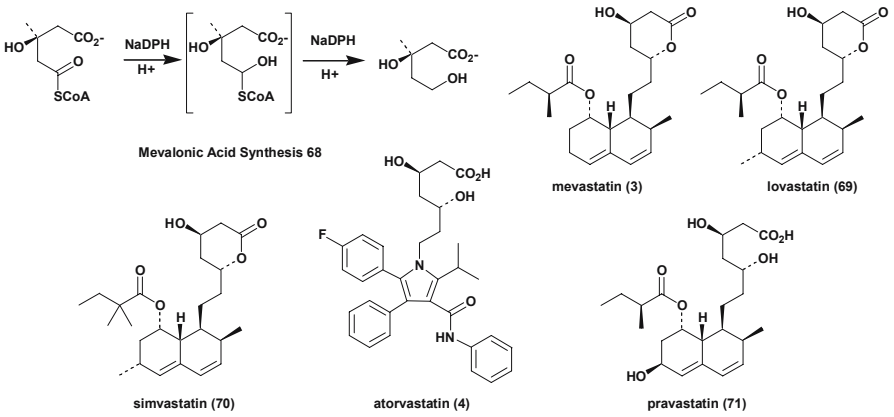
Artemisone (BAY 44-9585) (**67**), which is a semi-synthetic derivative of artemisinin (**64**) first synthesized at Hong Kong University, is being evaluated in Phase II antimalarial trials by Bayer [111, 112]. Artemisinin (**64**) and related compounds have been proposed to inhibit the parasite-encoded sarco-endoplasmic reticulum Ca^{2+}-ATPase (SERCA) PfATP6 [113], which is supported by polymorphism in the gene encoding PfATP6 that is associated with *in vitro* resistance to artemether in field [114, 115]. Further work into the mechanism of action of artemisinin (**64**) is ongoing in many laboratories around the world [116, 117].

3 Anticholestemics

A major cause of elevated blood pressure is due to the physical blockage of the arteries by plaques of cholesterol/lipoproteins (atherosclerotic plaque). Since a human usually synthesizes about 50% of its requirement for cholesterol, with the rest coming from diet, if the synthesis can be inhibited, a reduction in overall cholesterol levels may reduce its deleterious effects.

A potential site for inhibition of cholesterol biosynthesis in eukaryotes is at the rate limiting step in the system, the reduction of hydroxymethylglutaryl coenzyme A (HMG-CoA) to produce mevalonic acid (**68**). By following inhibition of sterol production and using fungal fermentation broths as the source of NPs, Endo first reported the activity of a fungal metabolite, mevastatin (**3**), in 1975 [118]. Mevastatin (**3**) (as compactin) had just been reported as an antifungal agent by Brown et al. [119]. Mevastatin (**3**) was shown to be a competitive inhibitor of the enzyme with K_is in the nanomolar range but was not further developed due to toxicity. Subsequently, using a similar assay, a 7-methyl derivative was reported by Endo as monacolin K (mevinolin, now known as lovastatin) (**69**) following isolation from *Monascus ruber* [120, 121]. This agent was patented in Japan but without a structure. Concomitantly, workers at Merck in the USA discovered the same material from *Aspergillus terreus*, using an isolated HMG-CoA reductase assay and microbial broths as their source of agents (having tried their synthetic library to no avail). After submission of both structure and findings to the US Patent Office, a US patent was issued in late 1980 and lovastatin (Mevacor®) (**69**) subsequently became the first commercialized HMG-CoA reductase inhibitor in 1987 [122]. Further work by Sankyo and Merck led to the entry of two more slightly modified versions. Firstly, the 2-methylbutanoate side-chain of lovastatin (**69**) was converted to 2,2-dimethyl butanoate, which was launched in 1988 as simvastatin (Zocor®) (**70**) by Merck. Secondly, biotransformation of mevastatin (**3**) led to the production of a lactone ring-opened, 7-hydroxy derivative, pravastatin (**71**), which was launched in 1989 by Sankyo and subsequently licensed to Bristol-Myers Squibb. In 2005, these latter two compounds had sales of just under $5 billion each. By that time, lovastatin (**69**) was off patent and so its sales were significantly reduced.

Following the success of these agents, other pharmaceutical companies used the information from mevastatin (**3**) and lovastatin (**69**) (effectively



the NP's 'warhead plus a lipophilic attachment') to develop a series of molecules, some of which have been approved for use. These include the best-selling drug of all time, atorvastatin (Lipitor®) (**4**), whose sales in 2005 were $7.4 billion in the USA and $12.2 billion worldwide. There are a number of agents similar in concept to atorvastin (**4**) in clinical use today.

3.1 Anticholestemics: current status

To the best of our knowledge there are no NP-derived compounds undergoing late stage clinical development as anticholestemics.

4 Anticancer

Since the early 1940s, the search for agents that may treat or ameliorate the scourge of cancer have involved all aspects of chemistry and pharmacology and throughout all these years, NPs have played an extremely important role. Initially, as the major source of drugs used for direct treatment; secondly as scaffolds upon which chemists would practice their skill and currently as modulators of specific cellular pathways in the tumor cell. Currently, the 175 agents available to the West and Japan can be categorized as follows: B (18; 10%), N (25; 14%), ND (48; 28%), S (42; 24%), S/NM (14;

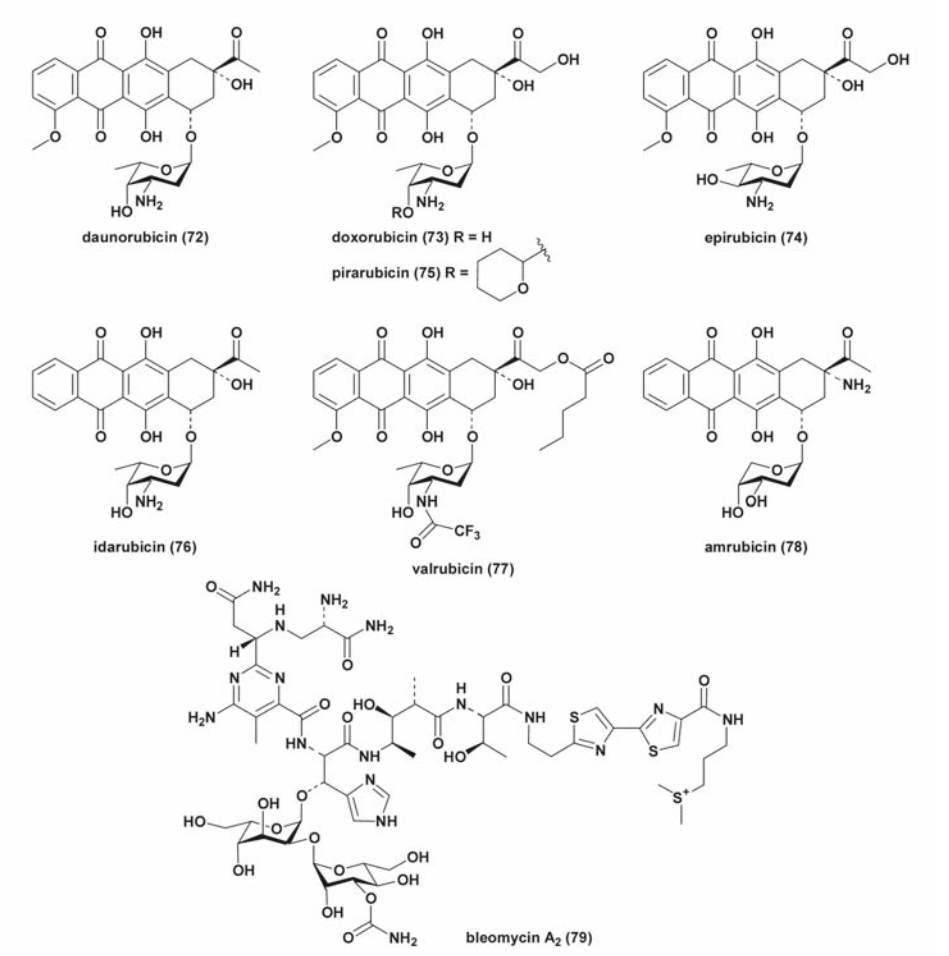
8%), S* (20; 11%), S*/NM (6; 4%) and V (2; 1%). If one removes the biologicals and vaccines, thus reducing the overall number to 155, the number of non-synthetic agents (i.e., N, ND, S/NM, S*, S*/NM) is 113 (73%), or if one removes the S/NM definition (14 in total), then the figure drops to 99 or expressed another way, 64% of the 155 agents approved for use in Japan and the West can trace their provenance to a NP source [8, 9].

Following WWII it became clear that a significant number of microbial products with nominal antibiotic activity against bacteria were effective also against eukaryotic cells from both fungi and mammals. In addition, there was some anecdotal evidence that plant secondary metabolites also demonstrated activity against mammalian cells, which effectively were tumor cells. Thus, from roughly the early 1950s to date, the systematic investigation of plant and microbial products for their potential as antitumor agents has been carried out worldwide, with a large amount of the funding coming from the US National Cancer Institute (NCI) in one form or another.

The first of the microbial compounds to be used systematically was named generically as 'Dactinomycin', or chemically as *D*-actinomycin C_1 with common usage as actinomycin D (**11**). As mentioned in Section 2.1, it was first introduced in the early 1960s and has been the subject of very significant chemical synthetic and semi-synthetic programs since its introduction.

From an antitumor perspective, one of the most important chemical classes of compounds derived from the bacterial order *Actinomycetales* are known generically as the 'anthracyclines', with one of the most useful being daunorubicin (**72**) and its derivative doxorubicin (adriamycin) (**73**). Even today, doxorubicin (**73**) is still a major component of the treatment regimen for breast cancer. Although there have been many similar molecules isolated and described in the literature, it is doxorubicin and its more modern derivatives such as epirubicin (**74**), pirirubicin (**75**), idarubicin (**76**) and, more recently, valrubicin (**77**) and amrubicin (**78**) (see below) that have been approved for cancer treatment [123]. The mechanism of action (MOA) of these molecules, aside from their formal identification as intercalators into the DNA helix, is now known to be inhibition of topoisomerase II, one of the important enzymes in the replication pathway of DNA during cell cycling [124].

Another series of extremely important molecules, also from the *Actinomycetales*, are the family of glycopeptolide antibiotics known as the



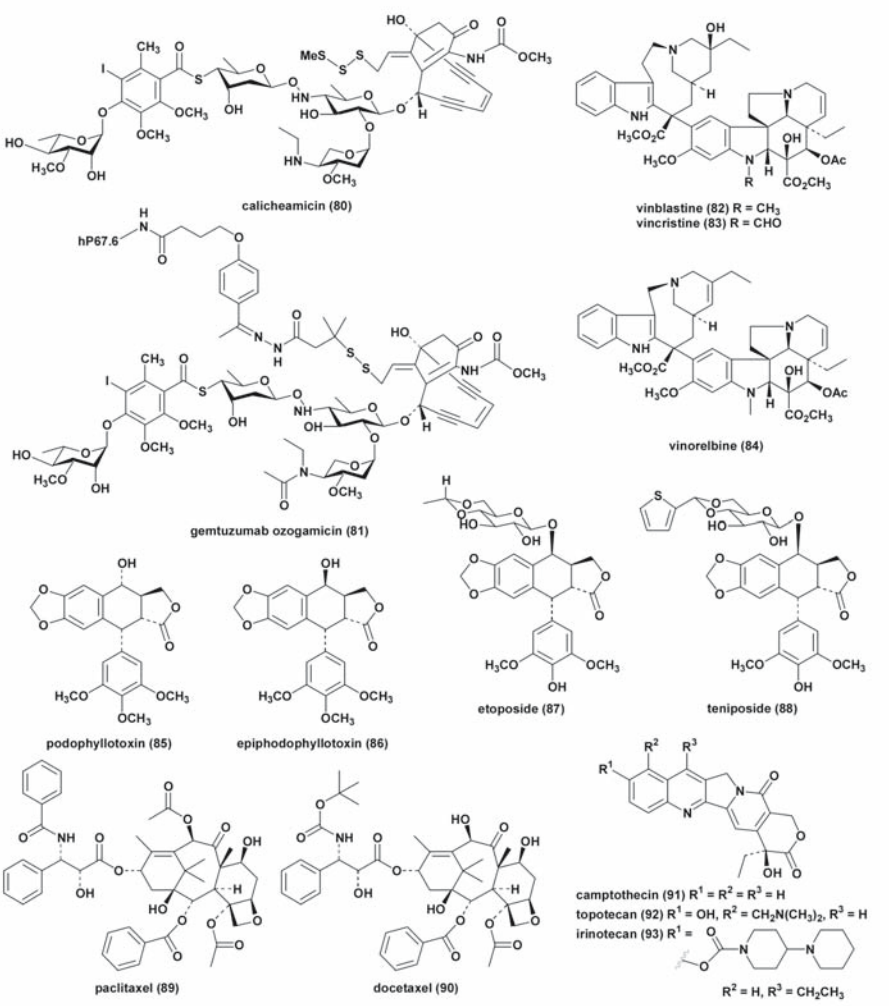
bleomycins (particularly bleomycin A_2 (**79**), Blenoxane®) and initially, the closely related structural class, the phleomycins. These molecules were originally reported by Umezawa's group at the Institute of Microbial Chemistry in Tokyo, and were developed as antitumor agents by Bristol-Myers. Their 'original' MOA was elucidated by Hecht and co-workers (MIT and the University of Virginia) who demonstrated that a metal ion (Cu^{2+} or Fe^{2+}) was required to activate the oxidative breakage of the DNA helix, once binding to the helices occurred, though there is evidence that they may also interact with RNA at lower concentrations [125].

Other microbial products that are still in use include the mitosanes, such as mitomycin C, the glycosylated anthracenone, mithramycin, streptozotocin and pentostatin, and, most importantly, calicheamicin γ_1^I (**80**). Calicheamicin γ_1^I (**80**) has *in vitro* activity at the sub-picomolar level and for a significant number of years was not developed as it was just too toxic to pursue, in spite of its exquisite activity [126]. It also was the progenitor of a new chemical class, the enediynes, which upon activation undergo an unprecedented rearrangement and interaction with DNA to produce double-stranded DNA cleavage. In 2000, Wyeth gained FDA approval for gemtuzumab ozogamicin (Mylotarg®) (**81**), an antibody-warhead construct for use against chronic myologenous leukemia [36, 127]. Gemtuzumab ozogamicin is possibly the most potent antitumor compound to be approved for clinical use.

Plants have a long history of use in the treatment of cancer [128] and the first agents to advance into clinical use were the so-called *Vinca* alkaloids, vinblastine (**82**) and vincristine (**83**), isolated from the Madagascar periwinkle, *Catharanthus roseus*, discovered during an investigation of the plant as a source of potential oral hypoglycemic agents [129]. Two semi-synthetic analogs of vincristine (**83**), vinorelbine (**84**) and vindesine, were approved subsequently for clinical use.

Podophyllotoxin (**85**) was isolated as the active antitumor agent from the roots of various species of the genus *Podophyllum*, and although podophyllotoxin (**85**) is still used as a topical treatment for warts (USP formulary) it was too toxic for use as an injectable antitumor agent. Following extensive semi-synthetic modifications of the isomeric natural product, epipodophyllotoxin (**86**), by Sandoz (now part of Novartis), two clinically-approved agents, etoposide (**87**) and teniposide (**88**), were developed. Interestingly, podophyllotoxin (**85**) recently was reported to be produced by epiphytic fungi isolated from plants that produce (**85**), so the real producer may well not be the plant itself [130].

More recent additions to the armamentarium of plant-derived chemotherapeutic agents are the taxanes and camptothecins. Paclitaxel (Taxol®) (**89**) initially was isolated from the bark of *Taxus brevifolia*, collected in Washington State as part of a random collection program by the US Department of Agriculture for the NCI [131]. Paclitaxel (**89**) was a pivotal discovery from a mechanistic aspect in the treatment of cancer as it was the first compound identified that caused the stabilization of microtubules



in the mitotic cycle. Paclitaxel (**89**) entered clinical use in 1993 and a semi-synthetic derivative, docetaxel (Taxotere®) (**90**), was approved in 1995, initially for treatment of refractory ovarian cancer and then breast cancer, though now these taxanes are being used in a large number of cancer treatments. In January 2005, Abraxane®, albumin-bound paclitaxel nanoparticles, was approved for the treatment of breast cancer. The albumin-bound paclitaxel eliminates the use of co-solvents such as Cremaphor®, which can cause irritation or toxicity [132, 133].

Wall and Wani, the discoverers of paclitaxel (**89**), also were responsible for the identification, isolation and purification of camptothecin (**91**) from the Chinese ornamental tree *Camptotheca acuminata*. It entered into clinical trials as its sodium salt by the NCI in the late 1970s, but it turned out to be too toxic due to reformation of the lactone ring in the bladder. Subsequent work involving the NCI, Johns Hopkins University and SmithKline Beecham (now GlaxoSmithKline) led to the synthesis of the semi-synthetic derivative, topotecan (**92**), which was approved in the USA in 1996 with another semi-synthetic derivative, irinotecan (**93**) approved approximately 18 months earlier in Japan [134].

4.1 Anticancer: current status

Despite the large numbers of NP and NP-derived compounds undergoing clinical evaluation [11, 12], only two NP-derived drugs, amrubicin hydrochloride (2002, Calsed®, Sumitomo Pharmaceuticals) (**78**) and vorinostat (2006, Zolinza™, Merck) (**94**) have been launched in Europe, USA or Japan since 2000. However, belotecan (2004, Camptobell®, CDK-602) (**95**) [135–137], a semi-synthetic camptothecin derivative, was launched in Korea by Chong Kun Dang. Amrubicin (**78**) [138–140] is the first totally synthetic doxorubicin (**73**) derivative used in the clinic and Cabrellis Pharmaceuticals recently started a Phase II trial of amrubicin (**78**) for the treatment of small-cell lung cancer in the USA. In October 2006, Merck announced that the FDA had approved the use of vorinostat (Zolinza™, suberoylanilide hydroxamic acid, SAHA) (**94**) [141, 142] for the treatment of cutaneous manifestations in patients with cutaneous T-cell lymphoma (CTCL), a form of non-Hodgkin's lymphoma. Vorinostat (**94**) is a histone deacetylase (HDAC) inhibitor [143] whose structure is related to the NP HDAC inhibitor trichostatin (**96**), which was originally isolated from *Streptomyces hygroscopicus* as an antifungal agent [144, 145].

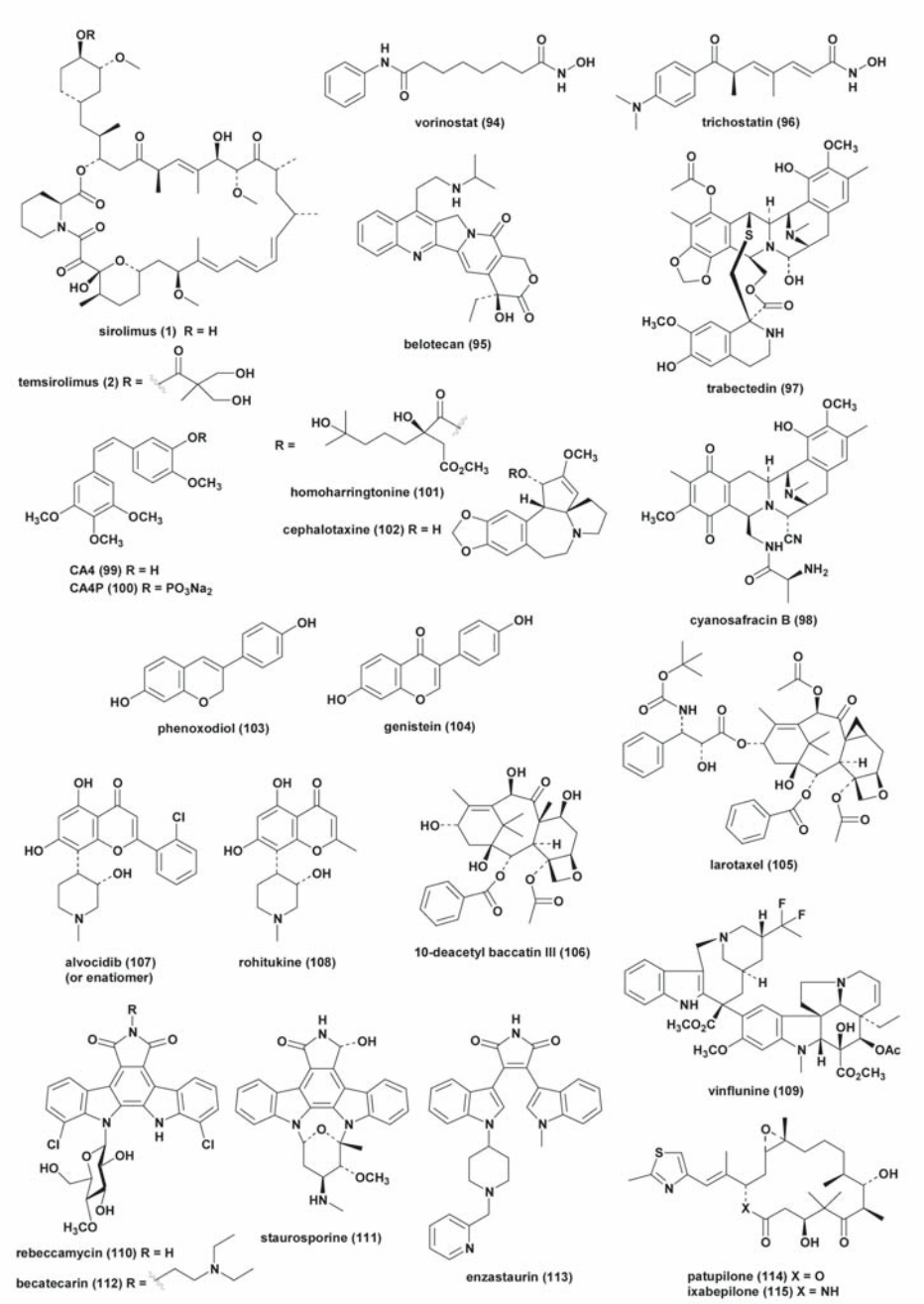
In October 2006, Wyeth Pharmaceuticals announced that they simultaneously filed an NDA to the FDA in the USA and a marketing authorization application to the European Medicines Agency (EMEA) for temsirolimus (Torisel®, CCI-779) (**2**) for the treatment of patients with advanced renal cell carcinoma (RCC). Temsirolimus (**2**) [36, 146] is a semi-synthetic derivative of sirolimus (**1**) that specifically inhibits mTOR (mammalian target

of rapamycin), a kinase that regulates cell proliferation, cell growth and cell survival [147, 148]. Temsirolimus (**2**) has orphan drug status for RCC in both the EU and the USA and is also in Phase II trials against multiple sclerosis and rheumatoid arthritis.

Trabectedin (Yondelis®, ET-743) (**97**) [149, 150] is a NP first isolated from the ascidian *Ecteinascidia turbinata* [151, 152] and is being developed by PharmaMar in partnership with J&J. Trabectedin can be produced by semi-synthesis from cyanosafracin B (**98**), which is obtained by fermentation of the bacterium *Pseudomonas fluorescens* [153]. The anticancer activity of trabectedin (**97**) is primarily due to DNA interactions but also involves additional non-DNA targets [154]. Trabectedin (**97**) is undergoing Phase II trials for the treatment of soft tissue sarcomas, prostate and breast cancers and Phase III trials for ovarian cancer. Orphan Drug status for soft tissue sarcomas and ovarian cancer indications has been granted by both the EMEA and FDA and in late July/early August 2006 application was made to the EMEA for Marketing Authorization Approval (MAA) for the treatment of sarcoma. If successful, this will be the first 'direct from marine source' drug for cancer treatment.

Combretastatin A-4 (CA4) (**99**) was first reported in 1989 by Pettit and co-workers from the African medicinal plant *Combretum caffrum* [155–159]. Although CA4 (**99**) showed potent cell growth and tubulin inhibition, it was not sufficiently water soluble and, as a consequence, a prodrug, combretastatin A-4 phosphate (CA4P) (**100**) [160], was advanced to clinical trials. CA4P (**100**) has shown the ability in both preclinical and clinical studies to drastically reduce blood flow in tumors, which may be useful in treating cancer where the abnormal growth of blood vessels is an essential component to the disease's progression, and OXiGENE is evaluating CA4P (**100**) in various Phase I, II and III clinical trials.

ChemGenex Pharmaceuticals began a Phase II/III study in September 2006 to evaluate the potential of homoharringtonine (Ceflatonin®) (**101**) [161–163] to treat patients with chronic myeloid leukemia (CML) who have the T315I bcr-abl point mutation. This mutation has become increasingly prevalent among patients with CML and is associated with resistance to imatinib mesylate (Gleevec®/Glivec®) and other tyrosine kinase inhibitors. Homoharringtonine (**101**) is an alkaloid from the Chinese evergreen tree *Cephalotaxus harringtonia*, which is isolated along with many closely related ester analogs [164]. These alkaloids are esters of an inactive alcohol,


sirolimus (1) R = H
temsirolimus (2) R =
vorinostat (94)
belotecan (95)
trichostatin (96)
trabectedin (97)
homoharringtonine (101)
R =
cephalotaxine (102) R = H
CA4 (99) R = H
CA4P (100) R = PO3Na2
cyanosafracin B (98)
phenoxodiol (103)
genistein (104)
larotaxel (105)
alvocidib (107)
(or enatiomer)
rohitukine (108)
10-deacetyl baccatin III (106)
vinflunine (109)
rebeccamycin (110) R = H
becatecarin (112) R =
staurosporine (111)
enzastaurin (113)
patupilone (114) X = O
ixabepilone (115) X = NH

cephalotaxine (**102**), which can be used to semi-synthetically produce homoharringtonine (**101**) in an ecologically sustainable way [165].

Phenoxodiol (**103**) is a synthetic analog of genistein (**104**), an isoflavone found in many plants, being evaluated by Marshall Edwards (Novogen) for the treatment of ovarian (Phase III), prostate (Phase II) and cervical cancer (Phase I) [166, 167]. Although phenoxodiol (**103**) induces cancer cell death through inhibition of multiple anti-apoptotic proteins, tumor-associated NADH oxidase (tNOX) has been identified as its primary target. tNOX is essential for growth and survival of cancers and, as a consequence, is a promising oncology target [168].

As previously discussed, the taxanes paclitaxel (**89**) and docetaxel (**90**) are in use for the treatment of various cancers. Larotaxel (XRP-9881, RPR 109881A) (**105**) [169–171], which is a semi-synthetic derivative of the taxoid 10-deacetyl baccatin III (**106**), is active in tumors with a high level of MDR1 expression and is undergoing Phase III trials with Sanofi-Aventis.

Alvocidib (flavopiridol, HMR 1275) (**107**) is a cyclin-dependent kinase inhibitor being evaluated in Phase III trials for the treatment of chronic lymphocytic leukemia also by Sanofi-Aventis [172–175]. Alvocidib (**107**) is a synthetic derivative of rohitukine (**108**), an alkaloid first isolated from *Amoora rohituka* [176] and later isolated from *Dysoxylum binectariferum* [177].

Vinflunine ditartrate (Javlor®) (**109**) [178, 179] is semi-synthetically derived from vinorelbine (**84**) (Navelbine®) [180], which in turn is semi-synthetically derived from the NP vincristine (**83**), using fluorination in superacid [181, 182]. Vinflunine (**109**) is being evaluated by Bristol-Myers Squibb and Pierre Fabre in various Phase III trials for the treatment of breast, bladder and lung cancers. In the USA, there is a pressing need for new drugs to treat bladder cancer as cisplatin is the only drug approved for this indication and this approval was back in 1981. Although all of these alkaloids block mitosis through a suppression of spindle microtubule dynamic, vinorelbine (**84**) and vinflunine (**109**) have been shown to affect microtubule dynamics differently to vinblastine (**82**) [183].

Rebeccamycin (**110**) is an *N*-glycoside indolocarbazole related to staurosporine (**111**) originally isolated from the actinomycete *Saccharothrix aerocolonigenes* [184], which inhibited topisomerase I and II and displayed potent *in vitro* cancer cell activity. However, poor water solubility of rebeccamycin (**110**) precluded further development and a more water soluble bis-

ethylamine derivative, becatecarin (DEAE rebbecamycin, XL-119) (**112**) [185–189], is in Phase III trials for the treatment of bile duct cancer. Exelixis licensed the worldwide rights of becatecarin (**112**) to Helsinn Healthcare in June 2005 but have retained the right to reacquire the USA license at a later date.

Enzastaurin (LY317615) (**113**) an orally bioavailable, synthetically-derived analog of staurosporine (**111**), is a potent inhibitor of serine/threonine protein kinase Cβ (PKCβ) [190–193]. Activation of PKCβ, one of at least eleven isoforms of PKC, is involved with upregulation of vascular endothelial growth factors (VEGF) receptors, which can lead to neovascularization and VEGF driven angiogenesis [194]. Eli Lilly has initiated a Phase III trial to study enzastaurin (**113**) as a maintenance therapy to prevent relapse in patients with diffuse large B-cell lymphoma, the most common form of non-Hodgkin's lymphoma, as well as Phase III trials for the treatment of relapsed glioblastoma multiforme, an aggressive and malignant form of brain cancer.

The epothilones [195, 196] were first reported by Höfle and co-workers from the myxobacterium *Sorangium cellulosum* as potent cytotoxics in a 1991 patent [197–199]. In 1995, workers at Merck showed that the epothilones had tubulin-stablizing activity [200] similar to that of paclitaxel (**89**) and this discovery was the catalyst for significant interest in this class of anticancer agents. Patupilone (epothilone B, EPO-906) (**114**) [201], which is a NP produced by fermentation, is being evaluated in Phase III trials for the treatment of solid tumors and Novartis expect to file an NDA for this indication in 2008. Ixabepilone (BMS-247550) (**115**) [202, 203] is a semi-synthetic lactam analog of patupilone (**114**) currently under investigation in Phase III trials for the treatment of metastatic breast cancer, as well as Phase I and II trials with other tumor types and combinatorial studies with other agents.

5 Conclusion

From the examples given above, it can be seen that Mother Nature's chemistry is still a fundamental source of inspiration, whether acknowledged or inferred, in drug discovery in the areas covered by this chapter and also in a variety of other diseases that are covered by other authors/chapters.

Although a very large number of compounds have been synthesized by combinatorial methods and tested in the disease areas covered in this chapter, to date, only one antitumor drug (sorafenib or BAY-43-9006) from *de novo* combinatorial chemistry has been approved (FDA, 2005), although many compounds whether derived from Nature or made synthetically have been optimized by combinatorial techniques and are in clinical trials or in some cases such as the antibiotic linezolid, clinical use.

What is becoming of great interest is the question "what is the actual producing organism?" when one isolates a natural product from a plant or marine invertebrate, or even from a microorganism such as a dinoflagellate or a fungus.

Within the last two or so years, natural product researchers have isolated endo(epi)phytic fungi from 'producing plants' that when fermented in a laboratory, produce the compound isolated from the plant, with examples being podophyllotoxin [130] and camptothecin [204], to add to the previously reported paclitaxel. In addition, there is little doubt that maytansine is in fact the product of a bacterium in the rhizosphere of the producing plant where the plant may make a very minor modification to the compound [205]. If one moves to the microbial arena, then the work reported from Ecopia on the isolation of a novel antifungal agent whose genes were identified in the organism and then expressed [206], can be added to the earlier work reported from investigation of the *S. coelicolor* genome by Challis and Hopwood [207], where previously unrecognized biosynthetic gene clusters were identified and expressed. These reports can now be added to those that unequivocally demonstrate that a significant number of marine natural products are produced by microorganisms contained within the invertebrate [208].

Perhaps the most unusual recent report was one demonstrating that the tubulin interactive 'fungal' product rhizoxin is in fact produced by an endosymbiotic microbe in the *Rhizopus* sp., with the ability to grow the isolated bacterium and demonstrate that it produces the metabolite, plus on removal of the bacterium from the 'fungal producing strain' no rhizoxin was produced under conditions where the untreated fungus will produce this compound [209].

Finally, just to demonstrate that even very well known sources can now be mined and their biosynthetic clusters expressed was the report in early 2006 where as a result of data-mining of the published *Aspergillus nidulans*

genome, Bok et al. [210] reported the existence of over 40 putative biosynthetic clusters of all classes, demonstrated that contrary to perceived dogma, that in at least this organism the secondary metabolic genes were clustered as in the prokaryotes, and demonstrated that they had located a master control gene by switching on production of a metabolite that had never been reported from *A. nidulans* but whose biosynthetic genes had been identified.

All of this data implies that there are tremendous opportunities for identification of novel agents/chemical skeletons against a multiplicity of disease targets from materials that we already have available and in addition, that researchers should actively consider genomic DNA extractions/curation performed as the 'host organisms' are collected, as new techniques will enable further investigation of the previously collected and stored genetic materials.

Acknowledgements

The authors used the Prous Integrity® database many times throughout this work in order to check the status of current drugs and those in clinical trials. The work would have been immeasurably more difficult and time consuming without this resource.

References

1 Anonymous (1998) A pictorial history of herbs in medicine and pharmacy. *Herbalgram* 42: 33–47
2 Borchardt JK (2002) The beginnings of drug therapy: Ancient mesopotamian medicine. *Drug News Perspect* 15: 187–192
3 Chang HM, But PPH (1986) *Pharmacology and applications of Chinese Materia Medica*. World Scientific Publishing, Singapore
4 Huang KC (1999) *The pharmacology of Chinese herbs*. CRC Press, Boca Raton, FL
5 Dev S (1999) Ancient-modern concordance in Ayurvedic plants: Some examples. *Environ Health Persp* 107: 783–789
6 Kapoor LD (1990) *CRC handbook of ayurvedic medicinal plants*. CRC Press, Boca Raton, FL
7 NLM, http://www.nlm.nih.gov/hmd/
8 Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the period 01/1981–06/2006. *J Nat Prod* 70: 461–477

9 Newman DJ, Cragg GM, Snader KM (2003) Natural products as sources of new drugs over the period 1981–2002. *J Nat Prod* 66: 1022–1037

10 Butler MS (2005) Natural products to drugs: Natural product derived compounds in clinical trials. *Nat Prod Rep* 22: 162–195

11 Chin YW, Balunas MJ, Chai HB, Kinghorn AD (2006) Drug discovery from natural sources. *AAPS J* 8: E239–253

12 Butler MS (2004) The role of natural products in drug discovery. *J Nat Prod* 67: 2141–2153

13 Koehn FE, Carter GT (2005) The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* 4: 206–220

14 Sneader W (2005) *Drug discovery: A history*. John Wiley & Sons, Chichester, UK

15 Mateles RI (1998) *Penicillin: A paradigm for biotechnology*. Candida Corporation, Chicago, IL

16 Newton GGF, Abraham EP (1956) Isolation of cephalosporin C, a penicillin-like antibiotic containing D-α-aminoadipic acid. *Biochem J* 62: 651–158

17 Abraham EP, Newton GGF (1961) The structure of cephalosporin C. *Biochem J* 79: 377–393

18 Imada A, Kitano K, Kintaka K, Muroi M, Asai M (1981) Sulfazecin and isosulfazecin, novel B-lactam antibiotics of bacterial origin. *Nature* 289: 590–591

19 Sykes RB, Cimarusti CM, Bonner DP, Bush K, Floyd DM, Georgopapadakou NH, Koster WH, Liu WC, Parker WL, Principe PA et al (1981) Monocyclic B-lactam antibiotics produced by bacteria. *Nature* 291: 489–491

20 Mauger AB, Lackner H (2005) The actinomycins. In: GM Cragg, DGI Kingston, DJ Newman (eds): *Anticancer agents from natural products*. Taylor and Francis, Boca Raton, Fl, 281–297

21 Weber JM, Leung JO, Maine GT, Potenz RHB, Paulus TJ, DeWitt JP (1990) Organization of a cluster of erythromycin genes in *Saccharopolyspora erythraea*. *J Bacteriol* 172: 2372–2383

22 Katz L, Donadio S (1993) Polyketide synthesis: Prospects for hybrid antibiotics. *Ann Revs Microbiol* 47: 875–912

23 Demain AL (2006) From natural products discovery to commercialization: A success story. *J Ind Microbiol Biotechnol* 33: 486–495

24 Baltz RH, Miao V, Wrigley SK (2005) Natural products to drugs: Daptomycin and related lipopeptide antibiotics. *Nat Prod Rep* 22: 714–741

25 von Nussbaum F, Brands M, Hinzen B, Weigand S, Habich D (2006) Antibacterial natural products in medicinal chemistry – exodus or revival? *Angew Chemie Int Ed* 45: 5072–5129

26 Keating GM, Perry CM (2005) Ertapenem: A review of its use in the treatment of bacterial infections. *Drugs* 65: 2151–2178

27 Shah PM, Isaacs RD (2003) Ertapenem, the first of a new group of carbapenems. *J Antimicrob Chemother* 52: 538–542

28 Nilius AM, Ma Z (2002) Ketolides: The future of the macrolides? *Curr Opin Pharmacol* 2: 493–500

29 Nguyen M, Chung EP (2005) Telithromycin: The first ketolide antimicrobial. *Clin Ther* 27: 1144–1163

30 Ubukata K, Hikida M, Yoshida M, Nishiki K, Furukawa Y, Tashiro K, Konno M, Mitsu-

hashi S (1990) *In vitro* activity of LJC10,627, a new carbapenem antibiotic with high stability to dehydropeptidase I. *Antimicrob Agents Chemother* 34: 994–1000
31 Perry CM, Ibbotson T (2002) Biapenem. *Drugs* 62: 2221–2234
32 Kahne D, Leimkuhler C, Lu W, Walsh C (2005) Glycopeptide and lipoglycopeptide antibiotics. *Chem Rev* 105: 425–448
33 Iso Y, Irie T, Iwaki T, Kii M, Sendo Y, Motokawa K, Nishitani Y (1996) Synthesis and modification of a novel 1 beta-methyl carbapenem antibiotic, S-4661. *J Antibiot* 49: 478–484
34 Anderson DL (2006) Doripenem. *Drugs Today (Barc)* 42: 399–404
35 Sum PE, Lee VJ, Testa RT, Hlavka JJ, Ellestad GA, Bloom JD, Gluzman Y, Tally FP (1994) Glycylcyclines. 1. A new generation of potent antibacterial agents through modification of 9-aminotetracyclines. *J Med Chem* 37: 184–188
36 Abou-Gharbia M (2002) In: B. Şener (ed): *Biodiversity: Biomolecular aspects of biodiversity and innovation utilization*. Kluwer Academic, New York, 63–70
37 Jones CH, Petersen PJ (2006) Tigecycline: First class glycylcycline a new choice for empiric therapy. *Drug Discov Today: Therap Strat* 3: 137–144
38 Stein GE, Craig WA (2006) Tigecycline: A critical analysis. *Clin Infect Dis* 43: 518–524
39 Sum PE (2006) Case studies in current drug development: 'glycylcyclines'. *Curr Opin Chem Biol* 10: 374–379
40 Page MG (2006) Anti-MRSA beta-lactams in development. *Curr Opin Pharmacol* 6: 480–485
41 Ishikawa T, Matsunaga N, Tawada H, Kuroda N, Nakayama Y, Ishibashi Y, Tomimoto M, Ikeda Y, Tagawa Y, Iizawa Y et al (2003) TAK-599, a novel N-phosphono type prodrug of anti-MRSA cephalosporin T-91825: synthesis, physicochemical and pharmacological properties. *Bioorg Med Chem* 11: 2427–2437
42 Kawamoto I, Shimoji Y, Kanno O, Kojima K, Ishikawa K, Matsuyama E, Ashida Y, Shibayama T, Fukuoka T, Ohya S (2003) Synthesis and structure-activity relationships of novel parenteral carbapenems, CS-023 (R-115685) and related compounds containing an amidine moiety. *J Antibiot* 56: 565–579
43 Kobayashi R, Konomi M, Hasegawa K, Morozumi M, Sunakawa K, Ubukata K (2005) *In vitro* activity of tebipenem, a new oral carbapenem antibiotic, against penicillin-nonsusceptible *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* 49: 889–894
44 Pace JL, Yang G (2006) Glycopeptides: Update on an old successful antibiotic class. *Biochem Pharmacol* 71: 968–980
45 Van Bambeke F (2006) Glycopeptides and glycodepsipeptides in clinical development: A comparative review of their antibacterial spectrum, pharmacokinetics and clinical efficacy. *Curr Opin Investig Drugs* 7: 740–749
46 Malabarba A, Ciabatti R, Kettenring J, Ferrari P, Scotti R, Goldstein BP, Denaro M (1994) Amides of de-acetylglucosaminyl-deoxy teicoplanin active against highly glycopeptide-resistant enterococci. Synthesis and antibacterial activity. *J Antibiot* 47: 1493–1506
47 Leadbetter MR, Adams SM, Bazzini B, Fatheree PR, Karr DE, Krause KM, Lam BM, Linsell MS, Nodwell MB, Pace JL (2004) Hydrophobic vancomycin derivatives with improved ADME properties: Discovery of telavancin (TD-6424). *J Antibiot* 57: 326–336
48 Cooper RD, Snyder NJ, Zweifel MJ, Staszak MA, Wilkie SC, Nicas TI, Mullen DL, Butler

TF, Rodriguez MJ, Huff BE et al (1996) Reductive alkylation of glycopeptide antibiotics: synthesis and antibacterial activity. *J Antibiot* 49: 575–581

49 Katz L, Ashley G (2005) Translation and protein synthesis: Macrolides. *Chem Rev* 105: 499–528

50 Azoulay-Dupuis E, Mohler J, Bedos JP, Barau C, Fantin B (2006) Efficacy of cethromycin, a new ketolide, against *Streptococcus pneumoniae* susceptible or resistant to erythromycin in a murine pneumonia model. *Antimicrob Agents Chemother* 50: 3033–3038

51 Or YS, Clark RF, Wang S, Chu DT, Nilius AM, Flamm RK, Mitten M, Ewing P, Alder J, Ma Z (2000) Design, synthesis, and antimicrobial activity of 6-O-substituted ketolides active against resistant respiratory tract pathogens. *J Med Chem* 43: 1045–1049

52 Wang G, Niu D, Qiu YL, Phan LT, Chen Z, Polemeropoulos A, Or YS (2004) Synthesis of novel 6,11-O-bridged bicyclic ketolides *via* a palladium-catalyzed bis-allylation. *Org Lett* 6: 4455–4458

53 Xiong L, Korkhin Y, Mankin AS (2005) Binding site of the bridged macrolides in the *Escherichia coli* ribosome. *Antimicrob Agents Chemother* 49: 281–288

54 Floss HG, Yu T-W (2005) Rifamycin – mode of action, resistance, and biosynthesis. *Chem Rev* 105: 621–632

55 Yamane T, Hashizume T, Yamashita K, Konishi E, Hosoe K, Hidaka T, Watanabe K, Kawaharada H, Yamamoto T, Kuze F (1993) Synthesis and biological activity of 3'-hydroxy-5'-aminobenzoxazinorifamycin derivatives. *Chem Pharm Bull (Tokyo)* 41: 148–155

56 Rothstein DM, Shalish C, Murphy CK, Sternlicht A, Campbell LA (2006) Development potential of rifalazil and other benzoxazinorifamycins. *Expert Opin Investig Drugs* 15: 603–623

57 Suchland RJ, Bourillon A, Denamur E, Stamm WE, Rothstein DM (2005) Rifampin-resistant RNA polymerase mutants of *Chlamydia trachomatis* remain susceptible to the ansamycin rifalazil. *Antimicrob Agents Chemother* 49: 1120–1126

58 Ciabatti R, Kettenring JK, Winters G, Tuan G, Zerilli L, Cavalleri B (1989) Ramoplanin (A-16686), a new glycolipodepsipeptide antibiotic. III. Structure elucidation. *J Antibiot* 42: 254–267

59 Walker S, Chen L, Hu Y, Rew Y, Shin D, Boger DL (2005) Chemistry and biology of ramoplanin: A lipoglycodepsipeptide with potent antibiotic activity. *Chem Rev* 105: 449–476

60 Freeman J, Baines SD, Jabes D, Wilcox MH (2005) Comparison of the efficacy of ramoplanin and vancomycin in both *in vitro* and *in vivo* models of clindamycin-induced *Clostridium difficile* infection. *J Antimicrob Chemother* 56: 717–725

61 Sunenshine RH, McDonald LC (2006) *Clostridium difficile*-associated disease: New challenges from an established pathogen. *Cleve Clin J Med* 73: 187–197

62 Breukink E, de Kruijff B (2006) Lipid II as a target for antibiotics. *Nat Rev Drug Discov* 5: 321–332

63 Fang X, Tiyanont K, Zhang Y, Wanner J, Boger D, Walker S (2006) The mechanism of action of ramoplanin and enduracidin. *Mol BioSyst* 2: 69–76

64 Tiyanont K, Doan T, Lazarus MB, Fang X, Rudner DZ, Walker S (2006) Imaging peptidoglycan biosynthesis in *Bacillus subtilis* with fluorescent antibiotics. *Proc Natl Acad Sci USA* 103: 11033–11038

65 Pankuch GA, Lin G, Hoellman DB, Good CE, Jacobs MR, Appelbaum PC (2006) Activity of retapamulin against *Streptococcus pyogenes* and *Staphylococcus aureus* evaluated

by agar dilution, microdilution, E-test, and disk diffusion methodologies. *Antimicrob Agents Chemother* 50: 1727–1730

66 Yan K, Madden L, Choudhry AE, Voigt CS, Copeland RA, Gontarek RR (2006) Biochemical characterization of the interactions of the novel pleuromutilin derivative retapamulin with bacterial ribosomes. *Antimicrob Agents Chemother* 50: 3875–3881

67 Schlunzen F, Pyetan E, Fucini P, Yonath A, Harms JM (2004) Inhibition of peptide bond formation by pleuromutilins: The structure of the 50S ribosomal subunit from *Deinococcus radiodurans* in complex with tiamulin. *Mol Microbiol* 54: 1287–1294

68 Hochlowski JE, Swanson SJ, Ranfranz LM, Whittern DN, Buko AM, McAlpine JB (1987) Tiacumicins, a novel complex of 18-membered macrolides. II. Isolation and structure determination. *J Antibiot* 40: 575–588

69 Theriault RJ, Karwowski JP, Jackson M, Girolami RL, Sunga GN, Vojtko CM, Coen LJ (1987) Tiacumicins, a novel complex of 18-membered macrolide antibiotics. I. Taxonomy, fermentation and antibacterial activity. *J Antibiot* 40: 567–574

70 Coronelli C, White RJ, Lancini GC, Parenti F (1975) Lipiarmycin, a new antibiotic from *Actinoplanes*. II. Isolation, chemical, biological and biochemical characterization. *J Antibiot* 28: 253–259

71 Omura S, Imamura N, Oiwa R, Kuga H, Iwata R, Masuma R (1986) Clostomicins, new antibiotics produced by *Micromonospora echinospora* subsp. *armeniaca* subsp. nov. I. Production, isolation, and physico-chemical and biological properties. *J Antibiot* 39: 1407–1412

72 Swanson RN, Hardy DJ, Shipkowitz NL, Hanson CW, Ramer NC, Fernandes PB, Clement JJ (1991) *In vitro* and *in vivo* evaluation of tiacumicins B and C against *Clostridium difficile*. *Antimicrob Agents Chemother* 35: 1108–1111

73 Ackermann G, Loffler B, Adler D, Rodloff AC (2004) *In vitro* activity of OPT-80 against Clostridium difficile. *Antimicrob Agents Chemother* 48: 2280–2282

74 Credito KL, Appelbaum PC (2004) Activity of OPT-80, a novel macrocycle, compared with those of eight other agents against selected anaerobic species. *Antimicrob Agents Chemother* 48: 4430–4434

75 Finegold SM, Molitoris D, Vaisanen ML, Song Y, Liu C, Bolanos M (2004) *In vitro* activities of OPT-80 and comparator drugs against intestinal bacteria. *Antimicrob Agents Chemother* 48: 4898–4902

76 Sergio S, Pirali G, White R, Parenti F (1975) Lipiarmycin, a new antibiotic from *Actinoplanes* III. Mechanism of action. *J Antibiot* 28: 543–549

77 Jenssen H, Hamill P, Hancock RE (2006) Peptide antimicrobial agents. *Clin Microbiol Rev* 19: 491–511

78 Marr AK, Gooderham WJ, Hancock RE (2006) Antibacterial peptides for therapeutic use: Obstacles and realistic outlook. *Curr Opin Pharmacol* 6: 468–472

79 MIGENIX Annual Report (2006). Available at http://www.migenix.com

80 Sader HS, Fedler KA, Rennie RP, Stevens S, Jones RN (2004) Omiganan pentahydrochloride (MBI 226), a topical 12-amino-acid cationic peptide: spectrum of antimicrobial activity and measurements of bactericidal activity. *Antimicrob Agents Chemother* 48: 3112–3118

81 Mechlinski W, Schaffner CP, Ganis P, Avitabile G (1970) Structure and absolute configuration of the polyene macrolide antibiotic amphotericin B. *Tetrahedron Lett* 11: 3873–3876

82 Borowski E, Zielinski J, Ziminski T, Falowski L, Kolodziejczyk P, Golik J, Jereczek E (1970) Chemical studies with amphotericin B. *Tetrahedron Lett* 11: 3909–3914
83 Chong CN, Rickards RW (1970) Macrolide antibiotic studies. XVI. The structure of nystatin. *Tetrahedron Lett* 11: 5145–5148
84 Borowski E, Zielinski J, Falowski L, Ziminski T, Golik J, Kolodziejczyk P, Jereczek E, Gdulewicz M, Shenin Y, Kotienko T (1971) The complete structure of the polyene macrolide antibiotic nystatin A_1. *Tetrahedron Lett* 12: 685–690
85 Pandey RC, Rinehart KL (1976) Polyene antibiotics. 7. C-13 Nuclear magnetic resonance evidence for cyclic hemiketals in polyene antibiotics amphotericin-B, nystatin-A1, tetrin-A, tetrin-B, lucensomycin, and pimaricin. *J Antibiot* 29: 1035–1042
86 Grove JF, MacMillan J, Mulholland TPC, Rogers MAT (1952) Griseofulvin. Part IV. Structure. *J Chem Soc*: 3977–3987
87 Morrison VA (2006) Echinocandin antifungals: Review and update. *Expert Rev Anti Infect Ther* 4: 325–342
88 Turner MS, Drew RH, Perfect JR (2006) Emerging echinocandins for treatment of invasive fungal infections. *Exp Opin Emerg Drugs* 11: 231–250
89 McCormack PL, Perry CM (2005) Caspofungin: A review of its use in the treatment of fungal infections. *Drugs* 65: 2049–2068
90 Chandrasekar PH, Sobel JD (2006) Micafungin: A new echinocandin. *Clin Infect Dis* 42: 1171–1178
91 Vazquez JA, Sobel JD (2006) Anidulafungin: A novel echinocandin. *Clin Infect Dis* 43: 215–222
92 Aperis G, Myriounis N, Spanakis EK, Mylonakis E (2006) Developments in the treatment of candidiasis: More choices and new challenges. *Expert Opin Investig Drugs* 15: 1319–1336
93 Andes D, Marchillo K, Lowther J, Bryskier A, Stamstad T, Conklin R (2003) *In vivo* pharmacodynamics of HMR 3270, a glucan synthase inhibitor, in a murine candidiasis model. *Antimicrob Agents Chemother* 47: 1187–1192
94 Warn PA, Sharp A, Morrissey G, Denning DW (2005) Activity of aminocandin (IP960) compared with amphotericin B and fluconazole in a neutropenic murine model of disseminated infection caused by a fluconazole-resistant strain of *Candida tropicalis*. *J Antimicrob Chemother* 56: 590–593
95 Bruzzese T, Rimaroli C, Bonabello A, Ferrari E, Signorini M (1996) Amide derivatives of patricin A with potent antifungal activity. *Eur J Med Chem* 31: 965–972
96 Kantarcioglu AS, Yucel A, Vidotto V (2003) *In vitro* activity of a new polyene SPK-843 against *Candida* spp, *Cryptococcus neoformans* and *Aspergillus* spp clinical isolates. *J Chemother* 15: 296–298
97 Kasanah N, Hamann MT (2005) SPK-843 (Aparts/Kaken). *Curr Opin Investig Drugs* 6: 845–853
98 Kaken Pharmaceutical Annual Report (2006) Available at: http://www.kaken.co/jp
99 Aparts BV Web Site (2006) Available at http://www.apartsbv.com.
100 Kaufman TS, Rúveda EA (2005) The quest for quinine: Those who won the battles and those who won the war. *Angew Chem Int Ed Engl* 44: 854–885
101 Tan RX, Zheng WF, Tang HQ (1998) Biologically active substances from the genus *Artemisia*. *Planta Med* 64: 295–302
102 Newman JD, Marshall J, Chang M, Nowoozi F, Paradise E, Pitera D, Newman KL, Keasling JD (2006) High-level production of amorpha-4,11-diene in a two-phase

partitioning bioreactor of metabolically engineered *Escherichia coli*. *Biotecnol Bioeng* 95: 684–691
103 Boatin BA, Richards FO Jr (2006) Control of onchoceriasis. *Adv Parasitol* 61: 349–394
104 Sanofi-Aventis Annual Report (2005) Available at http://www.sanofiaventis.com
105 Didier PJ, Phillips JN, Kuebler DJ, Nasr M, Brindley PJ, Stovall ME, Bowers LC, Didier ES (2006) Antimicrosporidial activities of fumagillin, TNP-470, ovalicin, and ovalicin derivatives *in vitro* and *in vivo*. *Antimicrob Agents Chemother* 50: 2146–2155
106 Didier ES (2005) Microsporidiosis: An emerging and opportunistic infection in humans and animals. *Acta Trop* 94: 61–76
107 Griffith EC, Su Z, Niwayama S, Ramsay CA, Chang YH, Liu JO (1998) Molecular recognition of angiogenesis inhibitors fumagillin and ovalicin by methionine aminopeptidase 2. *Proc Natl Acad Sci USA* 95: 15183–15188
108 Upadhya R, Zhang HS, Weiss LM (2006) System for expression of microsporidian methionine amino peptidase type 2 (MetAP2) in the yeast *Saccharomyces cerevisiae*. *Antimicrob Agents Chemother* 50: 3389–3395
109 Peterson RT (2006) A noncanonical path to mechanism of action. *Chem Biol* 13: 924–926
110 Zhang Y, Yeh JR, Mara A, Ju R, Hines JF, Cirone P, Griesbach HL, Schneider I, Slusarski DC, Holley SA et al (2006) A chemical and genetic approach to the mode of action of fumagillin. *Chem Biol* 13: 1001–1009
111 Haynes RK (2006) From artemisinin to new artemisinin antimalarials: Biosynthesis, extraction, old and new derivatives, stereochemistry and medicinal chemistry requirements. *Curr Top Med Chem* 6: 509–537
112 Haynes RK, Fugmann B, Stetter J, Rieckmann K, Heilmann HD, Chan HW, Cheung MK, Lam WL, Wong HN, Croft SL et al (2006) Artemisone – a highly active antimalarial drug of the artemisinin class. *Angew Chem Int Ed* 45: 2082–2088
113 Eckstein-Ludwig U, Webb RJ, Van Goethem ID, East JM, Lee AG, Kimura M, O'Neill PM, Bray PG, Ward SA, Krishna S (2003) Artemisinins target the SERCA of *Plasmodium falciparum*. *Nature* 424: 957–961
114 Jambou R, Legrand E, Niang M, Khim N, Lim P, Volney B, Ekala MT, Bouchier C, Esterre P, Fandeur T et al (2005) Resistance of *Plasmodium falciparum* field isolates to *in vitro* artemether and point mutations of the SERCA-type PfATPase6. *Lancet* 366: 1960–1963
115 Uhlemann AC, Cameron A, Eckstein-Ludwig U, Fischbarg J, Serovich P, Zuniga FA, East M, Lee A, Brady L, Haynes RK et al (2005) A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. *Nat Struct Mol Biol* 12: 628–629
116 Golenser J, Waknine JH, Krugliak M, Hunt NH, Grau GE (2006) Current perspectives on the mechanism of action of artemisinins. *Int J Parasitol* 36: 1427–1441
117 Krishna S, Woodrow CJ, Staines HM, Haynes RK, Mercereau-Puijalon O (2006) Re-evaluation of how artemisinins work in light of emerging evidence of *in vitro* resistance. *Trends Mol Med* 12: 200–205
118 Endo A (1975) Compactin (ML-236B) and related compounds as potential cholesterol-lowering agents that inhibit HMG-CoA reductase. *J Med Chem* 28: 401–405
119 Brown AG, Smale TC, King TJ, Hasenkamp R, Thompson RH (1976) Crystal and molecular structure of compactin, a new antifungal metabolite from *Penicillium brevicompactum*. *J Chem Soc, Perkin Trans* 1: 1165–1170

120 Endo A (1979) Monacolin K, a new hypocholesterolemic agent. *J Antibiot* 32: 852–854
121 Endo A (1980) Monacolin K, a new hypocholesterolemic agent that specifically inhibits 3-hydroxy-3-methylglutaryl coenzyme A reductase. *J Antibiot* 33: 334–336
122 Vagelos RP (1991) Are prescription drug prices too high? *Science* 252: 1090–1084
123 Arcamone FM (2005) Anthracyclines. In: GM Cragg, DGI Kingston, DJ Newman (eds): *Anticancer agents from natural products*. Taylor and Francis, Boca Raton, FL, 299–320
124 Pommier Y (2006) Topoisomerase I inhibitors: Camptothecins and beyond. *Nat Rev Cancer* 6: 789–802
125 Hecht SM (2005) Bleomycin group antitumor agents. In: GM Cragg, DGI Kingston, DJ Newman (eds): *Anticancer agents from natural products*. Taylor and Francis, Boca Raton, FL, 357–381
126 Hamann PR, Upeslacis J, Borders DB (2005) Enediynes. In: GM Cragg, DGI Kingston, DJ Newman (eds): *Anticancer agents from natural products*. Taylor and Francis, Boca Raton, FL, 451–474
127 Hamann PR, Hinman LM, Hollander I, Beyer CF, Lindh D, Holocomb R, Hallett W, Tsou H-R, Upeslacis J, Shochat D et al (2002) Gemtuzumab ozogamicin, a potent and selective anti-CD33 antibody-calicheamicin conjugate for treatment of acute myeloid leukemia. *Bioconjugate Chem* 13: 47–58
128 Hartwell JL (1982) *Plants used against cancer*. Quarterman, Lawrence, MA
129 Cragg GM, Boyd MR, Cardellina II JH, Newman DJ, Snader KM, McCloud TG (1994) Ethnobotany and drug discovery: The experience of the US National Cancer Institute. In: DJ Chadwick, J Marsh (eds): *Ethnobotany and the search for new drugs Ciba Foundation Symposium*. Wiley & Sons, Chichester, UK, 178–196
130 Eyberger AL, Dondapati R, Porter JR (2006) Endophyte fungal isolates from *Podophyllum peltatum* produce podophyllotoxin. *J Nat Prod* 69: 1121–1124
131 Kingston DGI (2005) Taxol and its analogs. In: GM Cragg, DGI Kingston, DJ Newman (eds): *Anticancer agents from natural products*. Taylor and Francis, Boca Raton, FL, 89–122
132 Desai N, Trieu V, Yao Z, Louie L, Ci S, Yang A, Tao C, De T, Beals B, Dykes D et al (2006) Increased antitumor activity, intratumor paclitaxel concentrations, and endothelial cell transport of Cremophor-free, albumin-bound paclitaxel, ABI-007, compared with Cremophor-based paclitaxel. *Clin Cancer Res* 12: 1317–1324
133 Green MR, Manikhas GM, Orlov S, Afanasyev B, Makhson AM, Bhar P, Hawkins MJ (2006) Abraxane®, a novel Cremophor®-free, albumin-bound particle form of paclitaxel for the treatment of advanced non-small-cell lung cancer. *Ann Oncol* 17: 1263–1268
134 Rahier NJ, Thomas J, Hecht SM (2005) Camptothecin and its analogs. In: GM Cragg, DGI Kingston, DJ Newman (eds): *Anticancer agents from natural products*. Taylor and Francis, Boca Raton, FL, 5–21
135 Lee JH, Lee JM, Kim JK, Ahn SK, Lee SJ, Kim MY, Jew SS, Park JG, Hong CI (1998) Antitumor activity of 7-[2-(N-isopropylamino)ethyl]-(20S)-camptothecin, CKD602, as a potent DNA topoisomerase I inhibitor. *Arch Pharm Res* 21: 581–590
136 Chung MK, Han SS, Kim JC (2006) Evaluation of the toxic potentials of a new camptothecin anticancer agent CKD-602 on fertility and early embryonic development in rats. *Regul Toxicol Pharmacol* 45: 273–281
137 Crul M (2003) CKD-602. Chong Kun Dang. *Curr Opin Investig Drugs* 4: 1455–1459

138 Ishizumi K, Ohashi N, Tanno N (1987) Stereospecific total synthesis of 9-aminoanthracyclines: (+)-9-amino-9-deoxydaunomycin and related compound. *J Org Chem* 52: 4477–4485
139 Hanada M, Mizuno S, Fukushima A, Saito Y, Noguchi T, Yamaoka T (1998) A new antitumor agent amrubicin induces cell growth inhibition by stabilizing topoisomerase II-DNA complex. *Jpn J Cancer Res* 89: 1229–1238
140 Hanada M, Noguchi T, Yamaoka T (2006) Amrubicin induces apoptosis in human tumor cells mediated by the activation of caspase-3/7 preceding a loss of mitochondrial membrane potential. *Cancer Sci* 97: 1396–1403
141 Miller TA, Witter DJ, Belvedere S (2003) Histone deacetylase inhibitors. *J Med Chem* 46: 5097–5116
142 Richon VM, Webb Y, Merger R, Sheppard T, Jursic B, Ngo L, Civoli F, Breslow R, Rifkind RA, Marks PA (1996) Second generation hybrid polar compounds are potent inducers of transformed cell differentiation. *Proc Natl Acad Sci USA* 93: 5705–5708
143 Bolden JE, Peart MJ, Johnstone RW (2006) Anticancer activities of histone deacetylase inhibitors. *Nat Rev Drug Discov* 5: 769–784
144 Tsuji N, Kobayashi M, Nagashima K, Wakisaka Y, Koizumi K (1976) A new antifungal antibiotic, trichostatin. *J Antibiot* 29: 1–6
145 Yoshida M, Kijima M, Akita M, Beppu T (1990) Potent and specific inhibition of mammalian histone deacetylase both *in vivo* and *in vitro* by trichostatin A. *J Biol Chem* 265: 17174–17179
146 Gu J, Ruppen ME, Cai P (2005) Lipase-catalyzed regioselective esterification of rapamycin: Synthesis of temsirolimus (CCI-779). *Org Lett* 7: 3945–3948
147 Reddy GK, Mughal TI, Rini BI (2006) Current data with mammalian target of rapamycin inhibitors in advanced-stage renal cell carcinoma. *Clin Genitourin Cancer* 5: 110–113
148 Smolewski P (2006) Investigating mammalian target of rapamycin inhibitors for their anticancer properties. *Expert Opin Investig Drugs* 15: 1201–1227
149 Jimeno J, Faircloth G, Fernández-Sousa JM, Scheuer P, Rinehart K (2004) New marine derived anticancer therapeutics – A journey from the sea to clinical trials. *Mar Drugs* 2: 14–29
150 Fayette J, Coquard IR, Alberti L, Boyle H, Meeus P, Decouvelaere AV, Thiesse P, Sunyach MP, Ranchere D, Blay JY (2006) ET-743: A novel agent with activity in soft-tissue sarcomas. *Curr Opin Oncol* 18: 347–353
151 Sakai R, Rinehart KL, Guan Y, Wang AH (1992) Additional antitumor ecteinascidins from a Caribbean tunicate: Crystal structures and activities *in vivo*. *Proc Natl Acad Sci USA* 89: 11456–11460
152 Sakai R, Jares-Erijman EA, Manzanares I, Silva Elipe MV, Rinehart KL (1996) Ecteinascidins: Putative biosynthetic precursors and absolute stereochemistry. *J Am Chem Soc* 118: 9017–9023
153 Menchaca R, Martinez V, Rodriguez A, Rodriguez N, Flores M, Gallego P, Manzanares I, Cuevas C (2003) Synthesis of natural ecteinascidins (ET-729, ET-745, ET-759B, ET-736, ET-637, ET-594) from cyanosafracin B. *J Org Chem* 68: 8859–8866
154 David-Cordonnier MH, Gajate C, Olmea O, Laine W, de la Iglesia-Vicente J, Perez C, Cuevas C, Otero G, Manzanares I, Bailly C et al (2005) DNA and non-DNA targets in the mechanism of action of the antitumor drug trabectedin. *Chem Biol* 12: 1201–1210

155 Pettit GR, Singh SB, Hamel E, Lin CM, Alberts DS, Garcia-Kendall D (1989) Isolation and structure of the strong cell growth and tubulin inhibitor combretastatin A-4. *Experientia* 45: 209–211

156 Pettit GR, Singh SB, Hamel E, Lin CM, Alberts DS, Garcia-Kendall D (1989) Erratum, isolation and structure of the strong cell growth and tubulin inhibitor combretastatin A-4. *Experientia* 45: 680

157 Chaplin DJ, Horsman MR, Siemann DW (2006) Current development status of small-molecule vascular disrupting agents. *Curr Opin Investig Drugs* 7: 522–528

158 Mahindroo N, Liou J-P, Chang J-Y, Hsieh H-P (2006) Antitubulin agents for the treatment of cancer – a medicinal chemistry update. *Expert Opin Ther Patents* 16: 647–691

159 Tron GC, Pirali T, Sorba G, Pagliai F, Busacca S, Genazzani AA (2006) Medicinal chemistry of combretastatin A4: Present and future directions. *J Med Chem* 49: 3033–3044

160 Pettit GR, Temple Jr C, Narayanan VL, Varma R, Simpson MJ, Boyd MR, Rener GA, Bansal N (1995) Antineoplastic agents 322. Synthesis of combretastatin A-4 prodrugs. *Anticancer Drug Des* 10: 299–309

161 Tujebajeva RM, Graifer DM, Karpova GG, Ajtkhozhina NA (1989) Alkaloid homoharringtonine inhibits polypeptide chain elongation on human ribosomes on the step of peptide bond formation. *FEBS Lett* 257: 254–256

162 Kantarjian HM, Talpaz M, Santini V, Murgo A, Cheson B, O'Brien SM (2001) Homoharringtonine: history, current research, and future direction. *Cancer* 92: 1591–1605

163 Lévy V, Zohar S, Bardin C, Vekhoff A, Chaoui D, Rio B, Legrand O, Sentenac S, Rousselot P, Raffoux E et al (2006) A phase I dose-finding and pharmacokinetic study of subcutaneous semisynthetic homoharringtonine (ssHHT) in patients with advanced acute myeloid leukaemia. *Br J Cancer* 95: 253–259

164 Powell RG, Weisleder D, Smith Jr CR, Rohwedder WK (1970) Structures of harringtonine, isoharringtonine, and homoharringtonine. *Tetrahedron Lett* 11: 815–818

165 Robin J-P, Blanchard J, Cavoleau S, Chauviat L, Charbonnel S, Dhal R, Dujardin G, Fournier F, Gilet C, Girodier L et al (2001) *USA Patent No. US6,613,900*

166 Choueiri TK, Wesolowski R, Mekhail TM (2006) Phenoxodiol: Isoflavone analog with antineoplastic activity. *Curr Oncol Rep* 8: 104–107

167 Mor G, Fu HH, Alvero AB (2006) Phenoxodiol, a novel approach for the treatment of ovarian cancer. *Curr Opin Investig Drugs* 7: 542–548

168 Davies SL, Bozzo J (2006) Spotlight on tNOX: A tumor-selective target for cancer therapies. *Drug News Perspect* 19: 223–225

169 Gelmon KA, Latreille J, Tolcher A, Génier L, Fisher B, Forand D, D'Aloisio S, Vernillet L, Daigneault L, Lebecq A et al (2000) Phase I dose-finding study of a new taxane, RPR 109881A, administered as a one-hour intravenous infusion days 1 and 8 to patients with advanced solid tumors. *J Clin Oncol* 18: 4098–4108

170 Kurata T, Shimada Y, Tamura T, Yamamoto N, Hyodo I, Saeki T, Takashima S, Fujiwara K, Wakasugi H, Kashimura M (2000) Phase I and pharmacokinetic study of a new taxoid, RPR 109881A, given as a 1-hour intravenous infusion in patients with advanced solid tumors. *J Clin Oncol* 18: 3164–3171

171 Sessa C, Cuvier C, Caldiera S, Bauer J, Van den Bosch S, Monnerat C, Semiond D, Pérard D, Lebecq A, Besenval M et al (2002) Phase I clinical and pharmacokinetic studies of the taxoid derivative RPR 109881A administered as a 1-hour or a 3-hour infusion in patients with advanced solid tumors. *Ann Oncol* 13: 1140–1150

172 Kaur G, Stetler-Stevenson M, Sebers S, Worland P, Sedlacek H, Myers C, Czech J, Naik

R, Sausville E (1992) Growth inhibition with reversible cell cycle arrest of carcinoma cells by flavone L86-8275. *J Natl Cancer Inst* 84: 1736–1740

173 Blagosklonny MV (2004) Flavopiridol, an inhibitor of transcription: implications, problems and solutions. *Cell Cycle* 3: 1537–1542

174 Potapova TA, Daum JR, Pittman BD, Hudson JR, Jones TN, Satinover DL, Stukenberg PT, Gorbsky GJ (2006) The reversibility of mitotic exit in vertebrate cells. *Nature* 440: 954–958

175 Shapiro GI (2006) Cyclin-dependent kinase pathways as targets for cancer treatment. *J Clin Oncol* 24: 1770–1783

176 Harmon AD, Weiss U, Silverton JV (1979) The structure of rohitukine, the main alkaloid of *Amoora rohituka* (Syn. *Aphanamixis polystachya*) (Meliaceae). *Tet Letts* 20: 721–724

177 Naik RG, Kattige SL, Bhat SB, Alreja B, de Souza NJ, Rupp RH (1988) An antiinflammatory cum immunomodulatory piperidinylbenzopyranone from *Dysoxylum binectariferum*: Isolation, structure and total synthesis. *Tetrahedron* 44: 2081–2086

178 Bennouna J, Campone M, Delord JP, Pinel MC (2005) Vinflunine: A novel antitubulin agent in solid malignancies. *Expert Opin Investig Drugs* 14: 1259–1267

179 Kruczynski A, Poli M, Dossi R, Chazottes E, Berrichon G, Ricome C, Giavazzi R, Hill BT, Taraboletti G (2006) Anti-angiogenic, vascular-disrupting and anti-metastatic activities of vinflunine, the latest vinca alkaloid in clinical development. *Eur J Cancer* 42: 2821–2832

180 Johnson SAN (1996) Vinorelbine: An update and review of activity. *Clin Oncol* 8: 353–357

181 Fahy J, Duflos A, Ribet J-P, Jacquesy J-C, Berrier C, Jouannetaud M-P, Zunino F (1997) Vinca alkaloids in superacidic media: A method for creating a new family of antitumor derivatives. *J Am Chem Soc* 119: 8576–8577

182 Jacquesy J-C, Berrier C, Jouannetaud M-P, Zunino F, Fahy J, Duflos A, Ribet J-P (2002) Fluorination in superacids: A novel access to biologically active compounds. *J Fluor Chem* 114: 139–141

183 Okouneva T, Hill BT, Wilson L, Jordan MA (2003) The effects of vinflunine, vinorelbine, and vinblastine on centromere dynamics. *Mol Cancer Ther* 2: 427–436

184 Nettleton DE, Doyle TW, Krishnan B, Matsumoto GK, Clardy J (1985) Isolation and structure of rebeccamycin – a new antitumor antibiotic from *Nocardia aerocoligenes*. *Tetrahedron Lett* 26: 4011–4014

185 Krishnan BS, Moore ME, Lavoie CP, Long BH, Dalterio RA, Wong HS, Rosenberg IE (1994) Fluorescence polarization studies of the binding of BMS 181176 to DNA. *J Biomol Struct Dyn* 12: 625–636

186 Long BH, Rose WC, Vyas DM, Matson JA, Forenza S (2002) Discovery of antitumor indolocarbazoles: Rebeccamycin, NSC 655649, and fluoroindolocarbazoles. *Curr Med Chem Anticancer Agents* 2: 255–266

187 Prudhomme M (2003) Rebeccamycin analogues as anti-cancer agents. *Eur J Med Chem* 38: 123–140

188 Rewcastle GW (2005) Becatecarin (Helsinn Healthcare). *IDrugs* 8: 838–847

189 Ricart AD, Hammond LA, Kuhn JG, Takimoto CH, Goetz A, Forouzesh B, Forero L, Ochoa-Bayona JL, Berg K, Tolcher AW et al (2005) Phase I and pharmacokinetic study of sequences of the rebeccamycin analogue NSC 655649 and cisplatin in patients with advanced solid tumors. *Clin Cancer Res* 11: 8728–8736

190 Faul MM, Gillig JR, Jirousek MR, Ballas LM, Schotten T, Kahl A, Mohr M (2003) Acyclic N-(azacycloalkyl)bisindolylmaleimides: Isozyme selective inhibitors of PKCβ. *Bioorg Med Chem Lett* 13: 1857–1859

191 Faul MM, Grutsch JL, Kobierski ME, Kopach ME, Krumrich CA, Staszak MA, Udodong U, Vicenzi JT, Sullivan KA (2003) Strategies for the synthesis of N-(azacycloalkyl)bisin dolylmaleimides: Selective inhibitors of PKCβ. *Tetrahedron* 59: 7215–7229

192 Pearce HL, Miller MA (2005) The evolution of cancer research and drug discovery at Lilly Research Laboratories. *Adv Enzyme Regul* 45: 229–255

193 Rizvi MA, Ghias K, Davies KM, Ma C, Weinberg F, Munshi HG, Krett NL, Rosen ST (2006) Enzastaurin (LY317615), a protein kinase Cβ inhibitor, inhibits the AKT pathway and induces apoptosis in multiple myeloma cell lines. *Mol Cancer Ther* 5: 1783–1789

194 Verhoef C, de Wilt JH, Verheul HM (2006) Angiogenesis inhibitors: Perspectives for medical, surgical and radiation oncology. *Curr Pharm Des* 12: 2623–2630

195 Altmann K-H (2004) The merger of natural product synthesis and medicinal chemistry: On the chemistry and chemical biology of epothilones. *Org Biomol Chem* 2: 2137–2152

196 Bergstralh DT, Ting JP (2006) Microtubule stabilizing agents: Their molecular signaling consequences and the potential for enhancement by drug combination. *Cancer Treat Rev* 32: 166–179

197 Höfle G, Bedorf N, Gerth K, Reichenbach H (1993) *German Patent No. DE4,138,042*

198 Höfle G, Bedorf N, Steinmetz H, Schomburg D, Gerth K, Reichenbach H (1996) Epothilone A and B – novel 16-membered. Macrolides with cytotoxic activity: Isolation, crystal structure, and conformation in solution. *Angew Chem Int Ed* 35: 1567–1569

199 Gerth K, Bedorf N, Höfle G, Irschik H, Reichenbach H (1996) Epothilons A and B: Antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties. *J Antibiot* 49: 560–563

200 Bollag DM, McQueney PA, Zhu J, Hensens O, Koupal L, Liesch J, Goetz M, Lazarides E, Woods CM (1995) Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. *Cancer Res* 55: 2325–2333

201 Rubin EH, Rothermel J, Tesfaye F, Chen T, Hubert M, Ho YY, Hsu CH, Oza AM (2005) Phase I dose-finding study of weekly single-agent patupilone in patients with advanced solid tumors. *J Clin Oncol* 23: 9120–9129

202 Borzilleri RM, Zheng X, Schmidt RJ, Johnson JA, Kim SH, DiMarco JD, Fairchild CR, Gougoutas JZ, Lee FYF, Long BH et al (2000) A novel application of a Pd(0)-catalyzed nucleophilic substitution reaction to the regio- and stereoselective synthesis of lactam analogues of the epothilone natural products. *J Am Chem Soc* 122: 8890–8897

203 Lee FYF, Borzilleri R, Fairchild CR, Kim SH, Long BH, Reventos-Suarez C, Vite GD, Rose WC, Kramer RA (2001) BMS-247550: A novel epothilone analog with a mode of action similar to paclitaxel but possessing superior antitumor efficacy. *Clin Cancer Res* 7: 1429–1437

204 Puri SC, Verma V, Amna T, Qazi N, Spiteller M (2005) An endophytic fungus from *Nothapodytes foetida* that produces camptothecin. *J Nat Prod* 68: 1717–1719

205 Yu T-W, Floss HG (2005) Ansamitocins (Maytansinoids). In: GM Cragg, DGI Kingston, DJ Newman (eds): *Anticancer agents from natural sources*. Taylor and Francis, Boca Raton, FL, 321–337

206 McAlpine JB, Bachmann BO, Piraee M, Tremblay S, Alarco A-M (2005) Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel anitfungal agent, as an example. *J Nat Prod* 68: 493–496

207 Challis GL, Hopwodd DA (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci USA* 100: 14555–14561

208 Newman DJ, Hill RT (2006) New drugs from marine microbes: The tide is turning. *J Ind Microbiol Biotechnol* 33: 539–544

209 Partida-Martinez LP, Hertweck C (2005) Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* 437: 884–888

210 Bok JW, Hiffmeister D, Maggio-Hall LA, Murillo R, Glasner JD, Keller NP (2006) Genomic mining for *Aspergillus* natural products. *Chem Biol* 13: 31–37

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Drug discovery and development with plant-derived compounds

By Olivier Potterat and Matthias Hamburger

University of Basel,
Institute of Pharmaceutical Biology,
Klingelbergstrasse 50, 4056 Basel,
Switzerland
<matthias.hamburger@unibas.ch>

Abstract

An overview is given on current efforts in drug development based on plant-derived natural products. Emphasis is on projects which have advanced to clinical development. Therapeutic areas covered include cancer, viral infections including HIV, malaria, inflammatory diseases, nociception and vaccine adjuvants, metabolic disorders, and neurodegenerative diseases. Aspects which are specific to plant-based drug discovery and development are also addressed, such as supply issues in the commercial development, and the Convention on Biological Diversity.

1 Introduction

From prehistoric times man has used plants to alleviate and treat diseases. However, feeding patterns of African primates provide evidence that these apes selectively use non-nutritional plants with pharmacological properties in case of certain diseases such as intestinal parasites [1, 2], suggesting that the origins of herbal medicine have their roots within the animal kingdom. Archeological findings document early breeding of plants such as poppy (*Papaver somniferum*), hemp (*Cannabis sativa*), and coca (*Erythroxylum coca* and *E. novogranatense*) which are known as cultured species only. Earliest written documentation of knowledge on medicinal properties of plants is found on Assyrian clay tablets dated about 2000 B.C. The Egyptian culture already used a range of medicinal plants which were described in the Ebers papyrus (1550 B.C.). In India, the traditional medicine was documented in the Ayurveda in about 900 B.C., and the first written records on the system of Traditional Chinese Medicine appeared around the same time. In Europe, the knowledge on medicinal plants reached an early summit with Hippocrates (5th Century B.C.), Dioscorides (1st Century A.D.) and Galen (2nd Century A.D.). The knowledge was transmitted and expanded by the Arab scholars of the 10th to 12th Century A.D., such as Avicenna and Ibn al-Baitar. In Europe, this knowledge was recorded in numerous herbals that were published from the Middle Age onwards [3]. Paracelsus (1493–1541) developed the first, albeit alchemistic idea of active principles contained in a medicinal plant (the so-called *Arcanum*, which he considered as an immaterial principle), and the concept of dose dependency of drug action and toxicity (*sola dosis facit venenum*).

The origin of pharmaceutical natural product research can be traced to 1805, when the German Pharmacist Sertürner isolated morphine from

Opium latex and soon recognized the superior therapeutic properties of the pure compound. The isolation of morphine spurred the discovery of numerous important drug substances within short time, e.g., emetine (1817), atropine (1819), quinine (1820), caffeine (1820), and digitoxin (1841) [4, 5]. The French pharmacists Pelletier and Caventou were particularly prolific in isolating numerous important alkaloids, and Caventou established in the mid 1820s a factory for production of quinine which was to become the first commercial natural product and pure drug substance (1826). Soon, factories in other European countries were established, e.g., E. Merck company in Germany (1827), and early pharmaceutical industry developed mainly along with the discovery of an increasing number of plant alkaloids. In the 19th and well into the 20th Century, ethnopharmacology provided a number of compounds with unique pharmacological properties and provided significant advances in pharmacotherapy, e.g., quinine (1820) as an antimalarial, cocaine (1860), the first local anesthetic, tubocurarine (1935), a muscle relaxant enabling modern surgical procedures, and reserpine (1951) as a first effective antihypertensive.

Towards the end of the 19th Century, rapidly growing understanding of organic synthesis and chemical structures led to first derivatives of natural products. Diacetylmorphin (1898) and acetylsalicylic acid (1899) were among the first compounds to be commercialized as pharmaceuticals. The contribution of plant-derived natural products to modern pharmacotherapy is considerable. Out of 243 structures which Sneader [6] identified as the starting point for the development of our modern drugs, some 60 compounds are of plant origin (Fig. 1).

In some cases, plant-derived natural products have inspired generations of medicinal chemists in their quest for better drugs. A case in point is the anticholinergic tropane alkaloids. Following the isolation of atropine in 1819 (**1**) and scopolamine (**2**) in 1881, a series of aromatic esters of tropine were synthesized in 1884. Homatropine (**3**) proved to be a faster and shorter acting mydriatic than atropine and represents the first example of a drug which provided genuine improvement over the natural product. Atropine methonitrate (**4**), a quaternary ammonium salt of atropine, was originally introduced in 1902 as mydriatic, but was used as a treatment for pyloric spasms in infants when it became clear that the polar nature of the compound precluded central side effects [3]. Larger substituents

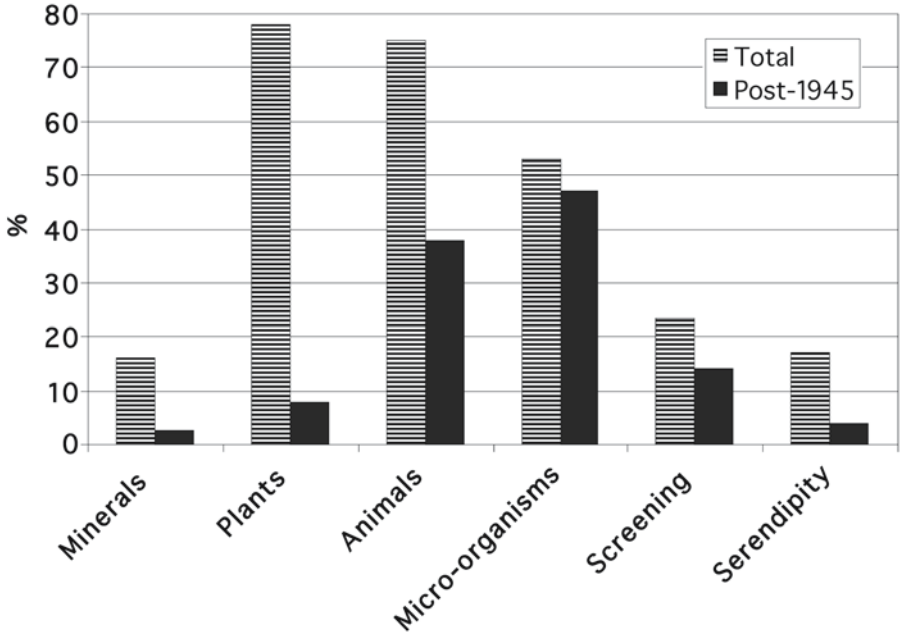
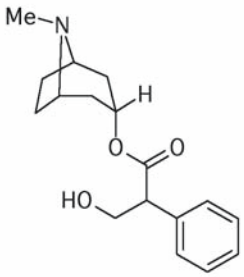


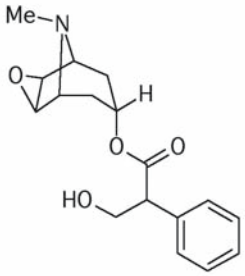
Figure 1.
Source of lead structures used as starting point for the development of modern drugs (adapted from [6])

were introduced at the nitrogen to prolong action, e.g., in N-butylscopolaminium bromide (Buscopan®) (**5**). On the other hand, medicinal chemists attempted to separate spasmolytic activity from antisecretory effects. They realized that the tropane ring could be replaced by a tertiary amine located two or three carbon distant to the ester moiety. As a result, spasmolytics such as amprotropine (**6**) and adiphenine (**7**) were developed in the 1930s. The most recent development in this class of compounds has been tiotropium bromide (Spiriva®) (**8**) which was introduced in 2002 as a long acting anticholinergic for inhalative therapy of chronic obstructive pulmonary disease (COPD).

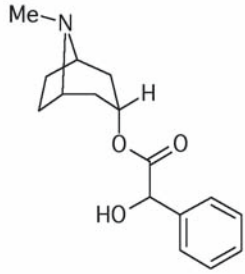
Recent examples of major plant-derived drugs based on new templates include paclitaxel (Taxol), and derivatives of artemisinin and ellipticine. In this chapter, we attempt to give an overview of ongoing efforts on plant-derived compounds and their derivatives which are being developed for a



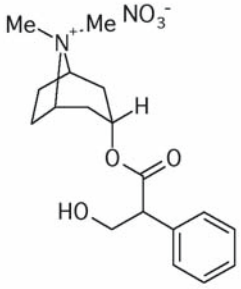
1 atropine



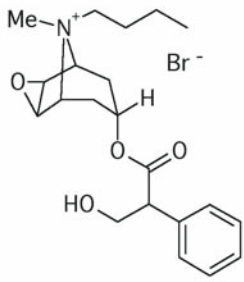
2 scopolamine



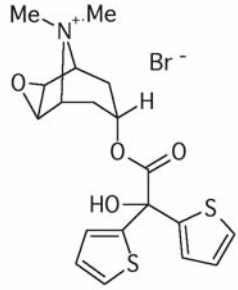
3 homatropine



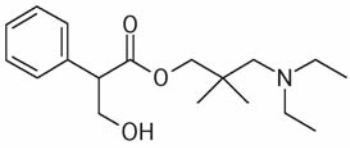
4 atropine methonitrate



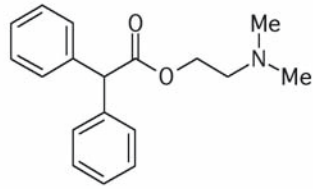
5 N-butylscopolaminium bromide



8 tiotropium bromide



6 amprotropine



7 adiphenine

number of therapeutic areas. The list of compounds discussed here is not meant to be exhaustive; rather we attempt to provide a selection of promising compounds with interesting and sometimes unique modes of action. We conclude the chapter with some considerations on issues which are quite specific to plant-derived compounds.

2 The discovery process of biologically active plant products

When screening for biologically active plant constituents, the selection of the plant species to be studied is obviously a crucial factor for the ultimate success of the investigation. Besides random collection of plant material, targeted collection based on consideration of chemotaxonomic relationships and the exploitation of ethnomedical information is currently pursued. Plants that have been used in traditional medicine are more likely to yield pharmacologically active compounds [7]. In the field of anticancer activity, a correlation between biological activity and plants used in folklore has been demonstrated [8].

The classical process leading from a bioactive extract to a pharmacologically active pure constituent has always been a long and tedious process requiring substantial material amount and financial resources [9]. It consists of several consecutive steps of preparative chromatographic separation, whereby each fraction has to be submitted to suitable bioassays to track the activity ultimately to a defined pure compound. While this procedure has led to the successful isolation of many bioactive molecules in the past, its weakness cannot be overlooked. Besides being slow and costly, the separation performance is poor, at least in the initial fractionation steps which are typically by open column chromatography. The loss of bioactivity in the course of the purification process is not uncommon and there is little means for early dereplication of known or otherwise uninteresting compounds. The approach described above obviously hardly matches the timelines and the workflow of modern drug discovery.

In the past decade, more effective strategies have been developed that are adequate for a high throughput environment. The concerted use of HPLC-DAD, -MS and -NMR has opened entirely new possibilities for the characterization of secondary metabolites in biological extracts by providing a wealth of structural information online with minute amounts of sample [10, 11]. Even absolute configuration can be established using HPLC-CD [12] or HPLC-NMR after Mosher's ester derivatization [11]. With the more recent emergence of mass spectrometry-controlled preparative high performance liquid chromatography (HPLC) [13], compound purification has also become straightforward, provided the compounds are sufficiently stable. The developments in the field of nuclear magnetic resonance (NMR)

[14], in particular the advent of new probe technology and higher magnetic fields, as well as the miniaturization process in crystallography rendered structural elucidation on amounts of less than a milligram becoming a rather routine process. While analysis, purification and structure elucidation of natural products have experienced a technological breakthrough over the last decade, tracking bioactivity in complex matrices such as plant extracts remains a highly challenging task. Extracts are complex mixtures, and one of the greatest challenges is the judicious interfacing of chemistry and biology in order to correlate chemical analysis with biochemical data. The development of high sensitivity and miniaturized assays has provided the technological basis to achieve this goal in novel ways. A variety of innovative methodologies measuring macromolecule-ligand interaction by means of various physicochemical methods or integrating immunochemical and enzymatic methods with online chemo-analytical systems have been recently implemented. New offline strategies such as HPLC-based activity profiling are directly applicable to a broad range of mechanism-based and cellular assays and have encountered increasing popularity in the context of industrial natural product screening [15].

A major issue when screening plant extracts is the comparably high number of false positives observed, due to the common presence of compounds with unspecific activities or interfering with the assay format. Tannins, forming tight complexes with metal ions and a wide array of proteins and polysaccharides often give a false-positive result in most assays involving a purified protein. Detergent-like compounds such as saponins, widely distributed in plants, and fatty acids have the tendency to disrupt membranes giving a misleading result. Compounds showing general cytotoxicity may give rise to false positives, in particular in cell-based assays. Strong metal chelators are susceptible to react with the assay components when nickel beads are used as linkers. Finally, in assays based on light measurement, pigments, ultraviolet (UV) quenchers such as the chlorophyll breakdown product phaeophorbide A, and autofluorescing compounds are prone to interfere with the readout [16]. Confronted to unrealistic hit rates in HTS and a slow and labour-intensive deconvolution process, many pharmaceutical companies have switched away from extract screening to prefractionated extracts or pure compounds libraries [17]. Large collections of compounds and semi-purified fractions have been generated using parallel fractionation and purification technology [17–20]. While these methods have the undeniable advantage to consid-

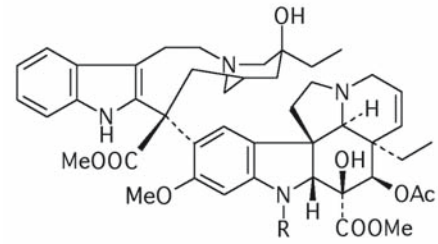
erably reduce or even skip the time-consuming follow-up process, they also have some intrinsic drawbacks. Thus, pure compound libraries will never be a real substitute for the huge structural diversity found in extracts. In particular, trace components potentially as interesting as major constituents are likely to be absent of such collections. The splitting of an extract into a large number of fractions on the other hand leads to a considerable increase of the number of samples to be screened, which can be an issue when working with expensive assay formats or costly targets, such as recombinant proteins. In fact, a well-balanced combination of pure compounds, fractions and extracts and a differential use thereof depending on the target and the screening format often proves to be the most judicious approach.

3 Anticancer agents

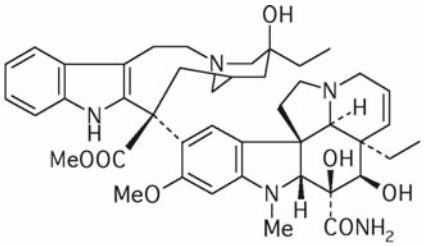
Efforts to obtain anticancer agents from higher plant started more than five decades ago and have been particularly successful [21]. Over 60% of all cancer drugs are of natural origin and a majority of these compounds are obtained from higher plants [22]. Up to the early 1990s, screening for new antitumor compounds was mainly based on cytotoxic testing against broad panels of cancer cell lines grown *in vitro* and subsequent testing *in vivo*. With the identification of an increasing number of targets associated with particular cancers, emphasis has meanwhile shifted towards more specific assays involving enzymes and receptors involved in cell division and tumor growth. Cyclin dependant kinases, tyrosine kinase receptors, topoisomerases are examples of such new molecular targets which have been subjected to high throughput screenings programs. The potential of these new approaches is illustrated below with a selection of the most promising lead compounds and drug candidates currently under development. In addition, well-established plant-derived anticancer drugs are also presented.

3.1 Vinca alkaloids

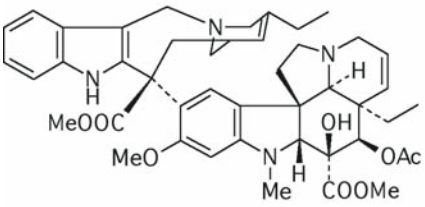
Vinblastine (**9**) (Velban®) and vincristine (**10**) (Oncovin®) were the first plant constituents to be used as antineoplastic agents. These alkaloids were isolated from the Madagascan periwinkle (*Catharanthus roseus*, Apocyna-



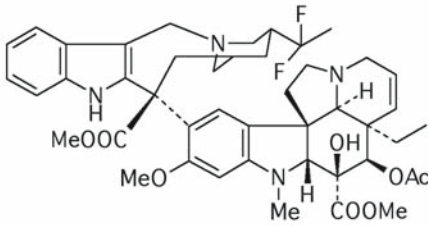
9 R = CH_3 vinblastine
10 R = CHO vincristine



11 vindesine



12 vinorelbine



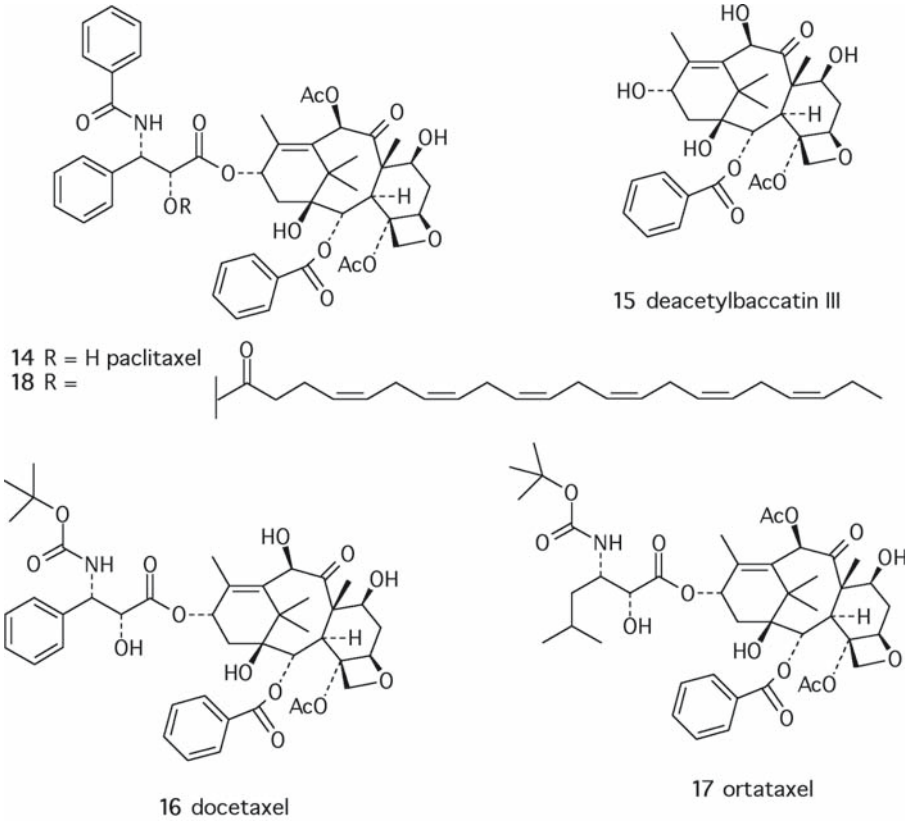
13 vinflunine

ceae) in the 1950s and developed at Eli Lilly. Vinca alkaloids inhibit cell growth by disrupting microtubules, causing the dissolution of the cell mitotic spindles and the arrest of the cells at metaphase finally leading to apoptotic cell death. They are widely use in the treatment of a wide variety of cancers including breast and lung cancer, as well as non-Hodgkin's lymphoma and leukemia. Two semisynthetic derivatives, vindesine (**11**) (Eldisine®) and vinorelbine (**12**), (navelbine®), were later prepared and are also in clinical use. Efforts to find further compounds with reduced toxicity and different activity profiles are still ongoing. Vinflunine (**13**), a bifluorinated derivative of vinorelbine is currently tested in Phase III clinical trials jointly by Bristol-Myers Squibb and Pierre Fabre [23–25].

3.2 Taxanes

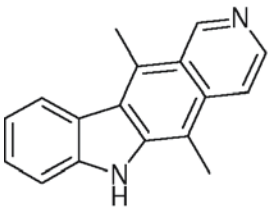
In 1971, Wall and co-workers at Research Triangle Institute reported the isolation of Paclitaxel (**14**) (Taxol®) from the stem bark of the pacific yew

Taxus brevifolia (Taxaceae) [26]. Yew has been reputed since ancient times for its toxicity and was used in particular as an arrow poison by the Celts [7]. The interest in Taxol was strongly stimulated by the discovery of its novel mode of action. In contrast to other antimicotic agents such as Vinca alkaloids or colchicin, Taxol does not inhibit tubulin polymerization but stabilizes the mitotic bundle by enhancing both the rate and speed of microtubule assembly resulting in cell death following arrest of cell cycle in M phase [27]. The clinical development of Taxol has long been hampered by its very limited supply from natural source. The compound occurs almost exclusively in the stem bark of various yew species at low concentration. Some 10,000 trees had to be sacrified to supply 2.5 Kg of Taxol for clinical testing which seriously threatened populations of Pacific yews. In the search for alternative strategies, French researchers at the Centre de la Recherche Scientifique (CNRS) developed a semisynthetic route to paclitaxel from a biogenetic precursor, deacetyl baccatin III (**15**), present in the leaves in substantial amount (1g/1kg fresh leaves). In addition, the semisynthetic approach enabled the preparation of Taxol analogs, in particular docetaxel (**16**) (Taxotere®), the second taxane currently used in clinics. Taxol and Taxotere are applied in the treatment of a large numbers of human cancers, including ovarian, breast and non-small cell lung cancers [28]. The clinical use of Taxol and Taxotere, even though being highly potent and effective cancer drugs, is restricted by low oral bioavailability, P-glycoprotein-mediated multidrug drug resistance and toxicity [29]. Extensive SAR studies have led to the development of a second generation of taxanes which overcome multidrug resistence and show better biovailability [30–32]. Modifications have been in particular introduced at the C-2, C-3' and C-10 positions. C-seco taxanes have also been prepared from seco-baccatin III and are claimed to have high oral bioavailability, reduced toxicity and high antiangiogenic activity [33, 34]. Several new taxoids are in clinical or advance preclinical development [23]. Probably one of the most promising is ortataxel (**17**) (IDN5180, BAY59-8862). It exhibits excellent acitivity against a variety of drug-sensitive and drug-resistant cancer lines and is currently in multi-centered Phase II clinical trials sponsored by Bayer and Indena [35]. The compound has an excellent therapeutic index, is orally active and inhibits the function of the Pgp efflux pump which brings much superior activity than Taxol and Taxotere against Pgp-expressing drug-resistant tumors. Targeted taxoid therapy is another inter-

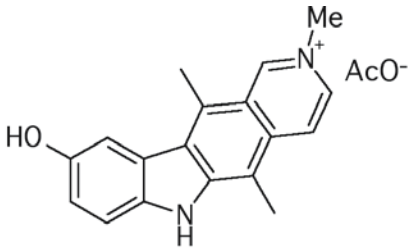


esting line of development to reduce toxic side effects. Taxoprexin® (**18**), a conjugate of docosahexaenoic acid (DHA) and paclitaxel is undergoing Phase III clinical evaluation. DHA-paclitaxel is well tolerated in patients. DHA efficiently targets paclitaxel to the tumor resulting in much higher drug concentration in tumors [36]. Very recently fatty acids have been also coupled to taxoids of the second generation [37], resulting in strong activity against drug-resistant colon and ovarian cancer xenografts and low systemic toxicity.

The huge therapeutical and economical importance of taxanes has led to extensive efforts to find alternative methods to produce paclitaxel or its precursor deacetyl-baccatin III. Since the report of the production of paclitaxel by the endophytic fungus *Taxomyces andreanae* in 1993 [38], other microorganisms have been reported to produce paclitaxel. The yields



19 ellipticine



20

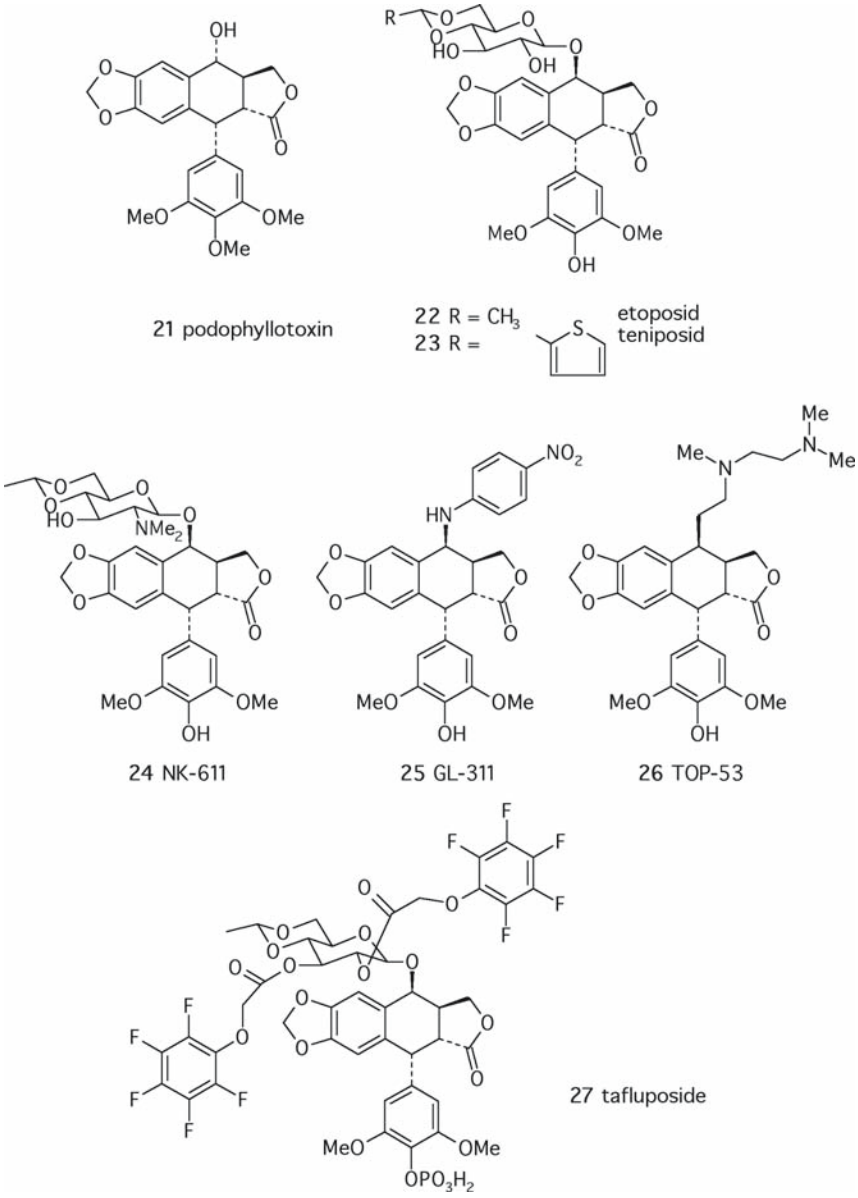
of microbial production are extremely low and thus not competitive with the extraction of deacetylbaccatin III from the plant and hemisynthesis. Cell cultures and genetic engineering of *Taxus* species are other approaches which appear of high interest for the future production of taxoids [30].

3.3 Ellipticines

The pyridocarbazole alkaloid ellipticine (**19**) was first isolated in 1959 from the leaves of *Ochrosia elliptica* (Apocynaceae). Ellipticine and its derivatives have been subsequently also found in other Apocynaceae, notably in species belonging to the genera *Aspidosperma*, *Tabernaemontana* and *Strychnos*. In the presence of cytochrome P450, ellipticin is oxidized in 8-hydroxyellipticine which is far more potent that the parent compound. In the course of extensive studies on the mechanism of action, the highly active and less toxic derivative 2-N-methyl-9-hydroxyellipticinium acetate (**20**) was prepared [7]. This molecule has been developed by Sanofi and is now marketed in France under the trade name Celiptium® for the treatment of breast cancer [21]. Drugs of the ellipticine series exert their action at the DNA level through intercalation and the induction of DNA strand breaks as a result of altered topoisomerase II activity [39].

3.4 Podophyllotoxin derivatives

Podophyllotoxin (**21**) is a major constituent of podophyllin, the resin obtained by aqueous extraction of May apple (*Podophyllum peltatum*, Ber-

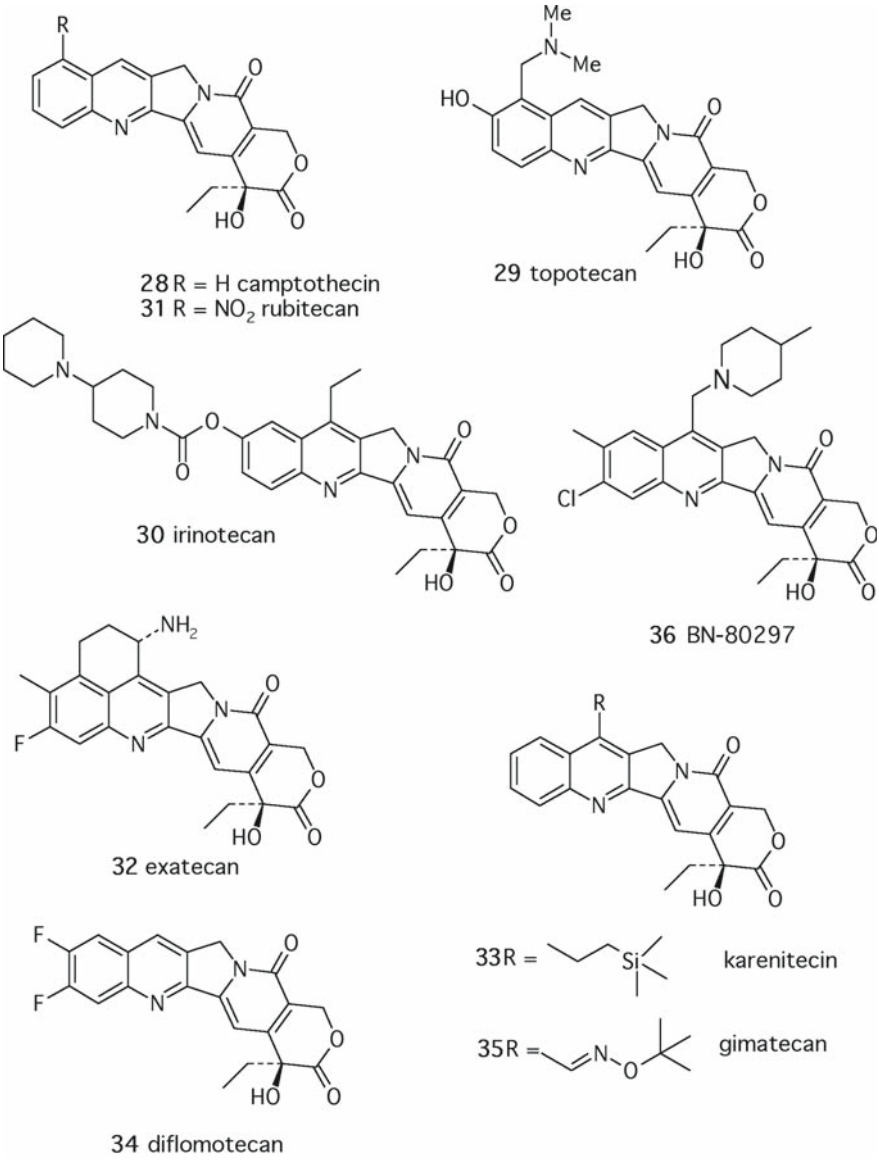


beridaceae) and the taxonomically close Indian species *P. emodii*. The antimitotic properties of podophyllin were discovered in the 1940s. The main active constituent, podophyllotoxin, and several congeners were

subsequently isolated and submitted to clinical evaluation. Therapeutic application of podophyllotoxins initially failed due to severe side effects. However, extensive research led to the development of two semisynthetic glucosidic acetals of epipodophyllotoxin, etoposide (**22**) (Vepesid®) and teniposide (**23**) (Vumon®), which have found clinical use for the treatment of various cancers. In contrast to podophyllotoxin, a classical spindle poison, etoposide and teniposide inhibit topoisomerase II activity and show few effects on tubulin polymerization. The mechanism, which is shared by many 4'-demethylepipodophyllotoxin derivatives, involves the stabilization of the DNA-enzyme cleavable complex and inhibition of the strand-rejoining activity of the enzyme, thus inducing double-strand DNA breaks [40]. The clinical efficacy of etoposide and teniposide has stimulated a wide range of research programs to find new derivatives with better solubility and able to overcome drug resistance. Etophos® (Etoposide phosphate), a water soluble prodrug of etoposide, was launched by Bristol-Myers Squibb in 1996. New molecules with promising activity currently in development include NK 611 (**24**), GL-331 (**25**), TOP 53 (**26**) and tafluposide (**27**). Of particular interest is the latter compound, since it is a dual inhibitor of both topoisomerases I and II with a unique molecular mechanism of action. Tafluposide impairs the binding of the enzyme to DNA but does not stabilize the cleavage complex like the classical epipodophyllotoxin derivatives [41]. Tafluposide is currently undergoing Phase I clinical trials at Pierre Fabre.

3.5 Camptothecin

The alkaloid camptothecin (**28**) was first isolated in 1966 from the wood of the Chinese ornamental tree *Camptotheca acuminata* (Nyssaceae), in the course of the antitumor screening program under the auspices of the NCI. Meanwhile it has been obtained also from several plants belonging notably to the families Apocynaceae and Icacinaceae. Animal studies revealed a potent antitumor activity, and camptothecin was clinically evaluated in patients with gastrointestinal cancers [42]. However, poor water solubility and severe side effects, in particular bladder toxicity, resulted in the discontinuation of Phase II trials in 1972. The discovery of its mode of action in the 1980s revived the interest in camptothecin. Camptothecin inhibits

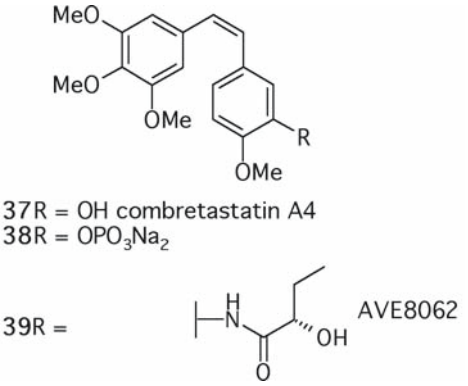


topoisomerase I by binding to the topoisomerase I-DNA binary complex, thereby inducing single strand breaks of cellular DNA [43]. Intensified efforts in medicinal chemistry eventually led to the development of more effective analogs with increased water solubility. Two of them, topotecan

(**29**) (Hycamtin®) and irinotecan (**30**) (Camptosar®) have found their way into the clinic and have become essential drugs with nearly $750 million in annual sales. Topotecan is primarily used for the treatment of ovarian and small cell lung cancer, while irinotecan is in use to treat colorectal cancers [21]. Further synthetic efforts are ongoing. A particular line of development is in the search for compounds with improved stability of the lactone function at physiological pH, which is essential for the activity. The most advanced derivative is rubitecan (**31**, 9-nitro camptothecin) which has recently reached pre-registration for the treatment of pancreatic cancer [42]. Further derivatives, including exatecan (**32**) (Phase III), karenitecin (**33**), diflomotecan (**34**) and gimatecan (**35**) are at various stages of clinical development [23]. Of particular interest is the homocamptothecin derivative BN 80297 (**36**) which acts as a dual topoisomerases I/II inhibitor [44]. On an industrial scale, camptothecin is still obtained by extraction from the plant, mainly from *Camptotheca acuminata* and *Nothapodytes foetida* (Icacinaceae). However, efforts are being made to produce camptothecin in cell and organ cultures. While low yields were reported in suspension cultures of *C. acuminata* or *N. foetida* (0.0003–0.01%), good production (0.1–0.3% dry wt) was achieved in root and hairy root cultures of *Ophiorrhiza* species (Rubiaceae) [45, 46]. In addition, the biosynthesis of camptothecin by an endophytic fungus of *N. foetida* has been recently reported [47]. While the yields are still low, strain optimization as well as improvement of the fermentation conditions could possibly enable a biotechnological production of camptothecin capable of meeting the ever-increasing demand for camptothecin.

3.6 Combretastatins

Combretastatins belong to the vascular disrupting agents and are one of the latest additions to the anticancer drug candidates derived from natural products. Combretastatin A4 (**37**) is the most potent compound of a series of cis-stilbene derivatives which were originally isolated from the African tree *Combretum caffrum* (Combretaceae) [48]. Like Vinca alkaloids, combretastatin A4 is a potent inhibitor of tubulin polymerization. However the interest for this compound is not due to its antimitotic properties but to its ability to induce vascular dysfunction in tumors. Neovasculariza-

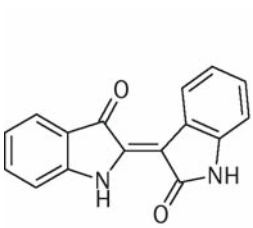


tion is essential for solid tumor to grow beyond a threshold size and is also critical for the spread of tumor cells to distant organs [48]. Combretastatin selectively disrupts the vascular system of tumors leading to extensive tumor necrosis at doses less than one tenth of the maximum tolerated dose in mice [49]. The mechanism for the selective vascular collapse observed in tumors is not yet completely elucidated, but it has been shown that combretastatin has dramatic effects on the three dimensional shape of immature endothelial cells. Combretastatin 4A causes disruption of the tubulin cytoskeleton and remodelling of the actin cytoskeleton [50]. Clinical trials are being performed with the water-soluble prodrug combretastatin 4A phosphate (**38**), which is now produced synthetically in a five step process. Despite extensive tumor cell killing, significant delay of tumor growth is rarely obtained when combretastatin 4A phosphate is used as a single therapy. This has been attributed to the rapid re-growth from the rim of cells at the tumor periphery. On the other hand, results in experimental tumor models with combretastatin 4A phosphate in combination with conventional cytotoxic chemotherapy or radiotherapy are promising. Such combinations are currently being tested in Phase II clinical trials against a range of tumor types. Combretastatin 4A phosphate is developed by OxiGene but is not the only compound under investigation. The straightforward synthesis has facilitated the preparation of hundreds of combrestatin derivatives [51]. Molecules in development include combretastatin-1 phosphate also in preclinical trial at OxyGene, and AVE8062 (**39**), a prodrug with a serine residue which is currently in Phase I clinical

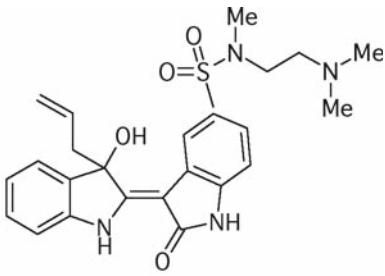
trials at Sanofi-Aventis. It is noteworthy to mention that the potential of combretastatin is not confined to oncology: combretastatin 4A phosphate has recently entered Phase I clinical trials for the treatment of wet age-related macular degeneration [49, 52].

3.7 Indirubins

Indirubins are a relatively new but highly attractive class of lead compounds showing strong antiproliferative activity. Indirubin (**40**) is the red component of indigo dyes which have been used since the Bronze Age for textile coloring. Indirubin and its blue isomer indigo are formed from precursor compounds during the drying or processing of various plants including *Polygonum tinctorium* (Polygonaceae), *Isatis tinctoria* (Brassicaceae) and *Indigofera tinctoria* (Fabaceae), as well as some marine molluscs of the genus *Murex* [53]. The significance of natural sources for the production of indigo dye vanished with the chemical synthesis of indigo and indirubin by Bayer in the late 19th Century. At the same time, several indigo and indirubin producing plants have a long history as medicinal plants and are still used in particular to treat inflammatory and malignant diseases. In 1979, indirubin was reported to be the active principle of Danggui Longhui Wan, a mixture of plants used in traditional Chinese medicine to treat various diseases including chronic myelocytic leukaemia [54, 55]. Indirubin effectively inhibits the growth of various human tumor cell lines in the low micromolar range and also shows activity in human xenograft models. The mechanism of action was revealed – at least in part – only in the late 1990s: indirubins are potent ATP-competitive inhibitors of cyclin dependent kinases, a group of enzymes involved in the cell cycle machinery. Indirubins induce cell cycle arrest in G2 and/or G2/M phase leading to apoptosis [54, 56]. In addition, indirubins inhibit glycogen synthase kinase-3b (GSK-3), c-Src kinase, and activate the aryl hydrocarbon receptor. Very recently, indirubin has also been shown to inhibit NF-κB activation and expression of NF-κB-regulated gene products, contributing to enhancement of apoptosis and suppression of invasion [57]. X-ray data obtained from co-crystal structures of various indirubins with GSK3, CDK2 and CDK5 and modelling studies were applied to prepare large series of new indirubin derivatives with different selectivity profiles [55, 58]. Some



40 indirubin

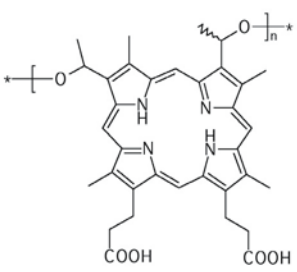


41

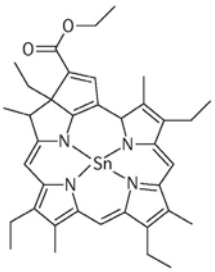
analogs (e.g., **41**) obtained at Schering are readily soluble, almost colorless and selectively inhibit CDK2, thus providing good premise for preclinical studies.

3.8 Photosensitizers

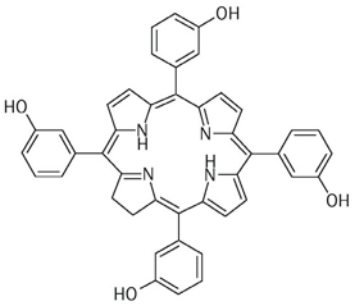
Photodynamic therapy (PDT) is a promising new modality for cancer therapy. While the use of PDT in cancer therapy was reported for the first time in 1903 by von Tappeiner and Jesionek [59], it was only at the end of the 20th Century that the potential of PDT was really exploited. PDT consists of the systemic administration of a photosensitizing agent which is preferentially taken up and retained by tumor cells, followed by a targeted irradiation of the tumor with visible light. Photosensitizing agents have a low dark toxicity but, upon irradiation in the presence of oxygen, produce highly cytotoxic species such as singlet oxygen and superoxide radical. Since the aforementioned oxygen reactive species have a very short lifetime, their action is restricted to the immediate area containing the photosensitizing drug. Clearly, PDT is particularly well suited to the treatment of skin cancers which are easy to irradiate. However, the development of powerful lasers combined with optical fibers enables its clinical use also in cancers endoscopically accessible, such as oesophageal, lung, bladder, ovarian, head and neck cancers [60]. Porphyrins are probably the most extensively explored group of sensitizer/tumor-selective of compounds [61, 62]. Naturally occurring porphyrins, e.g., chlorophyll, do not induce photochemical reactions or are rapidly quenched. Tetrapyrroles become efficient photosensitizers by removal of the coordinated metal ion [63]. Photofrin® (**42**)



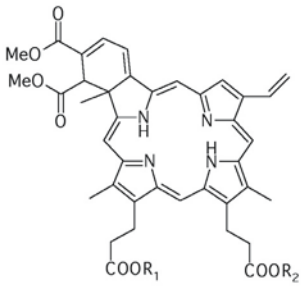
42 photofrin



43 SnEt2

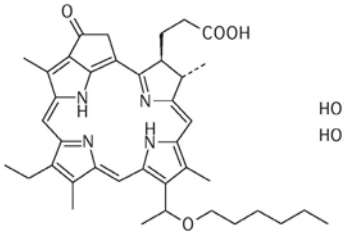


44 mTHPC

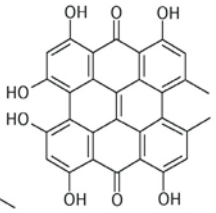


$R_1 = CH_3$, $R_2 = H$ or $R_1 = H$, $R_2 = CH_3$

45 BPD-MA



46 HPPH



47 hypericin

was the first product to be approved. It consists of a porphyrin oligomer mixture obtained from Hematoporphyrin-IX by acid-catalyzed oligomerization [62]. It has been widely applied for more than two decades in the treatment of various cancers, including bladder and palliative oesophageal and non-small cell lung cancers, and is still the most broadly used photosensitizer in the clinic [63]. Other compounds approved for clinical application include tin etiopurpurin (**43**, SnEt2), tetra (meso-hydroxy) phenyl chlorine (**44**, mTHPC) and benzoporphirin derivative monoacid ring A (**45**, BPD-MA) [63]. Finally, several porphyrin derivatives currently underlie extensive investigation, e.g., HPPH (**46**, 2-[1-hexyloxyethyl]-2-devinyl pyropheophorbide-a), a pheophorbide-a analog, derived from chlorophyll-a which was found to be extremely effective and shows limited toxicity. The

compound is currently in Phase II clinical trials for various indications at the Roswell Park Cancer Institute (Buffalo, NY, USA) [64].

In recent years, considerable interest has also been focussed on hypericin (**47**) as a powerful naturally occurring photosensitizer. Hypericin is a naphthodianthrone which occurs together with its congener pseudohypericin in many *Hypericum* species (Hypericaceae), in particular in St. John's Wort (*H. perforatum*). Hypericin shows a particularly low intrinsic dark toxicity as confirmed by the wide use of St. John's Wort extracts in the treatment of mild to moderate depression. Hypericin exhibits a high affinity for tumors and strong antineoplastic activity both *in vitro* and *in vivo* in the presence of oxygen. Hypericin has been shown to trigger necrosis or apoptosis, depending on the level of oxidative stress produced. The apoptotic pathway is mediated by mitochondrial release of pro-apoptotic signals, in particular cytochrome c, which in turn activates the caspase cascade. In addition, hypericin has been shown to inhibit protein kinase C and to activate the JNK and p38 MAPK signalling pathways [60, 65]. Recent clinical trials have confirmed the potential of hypericin-PDT in the therapy of recurrent mesothelioma and of basal and squamous carcinoma. The compound also holds much promise for the diagnosis and treatment of bladder cancer.

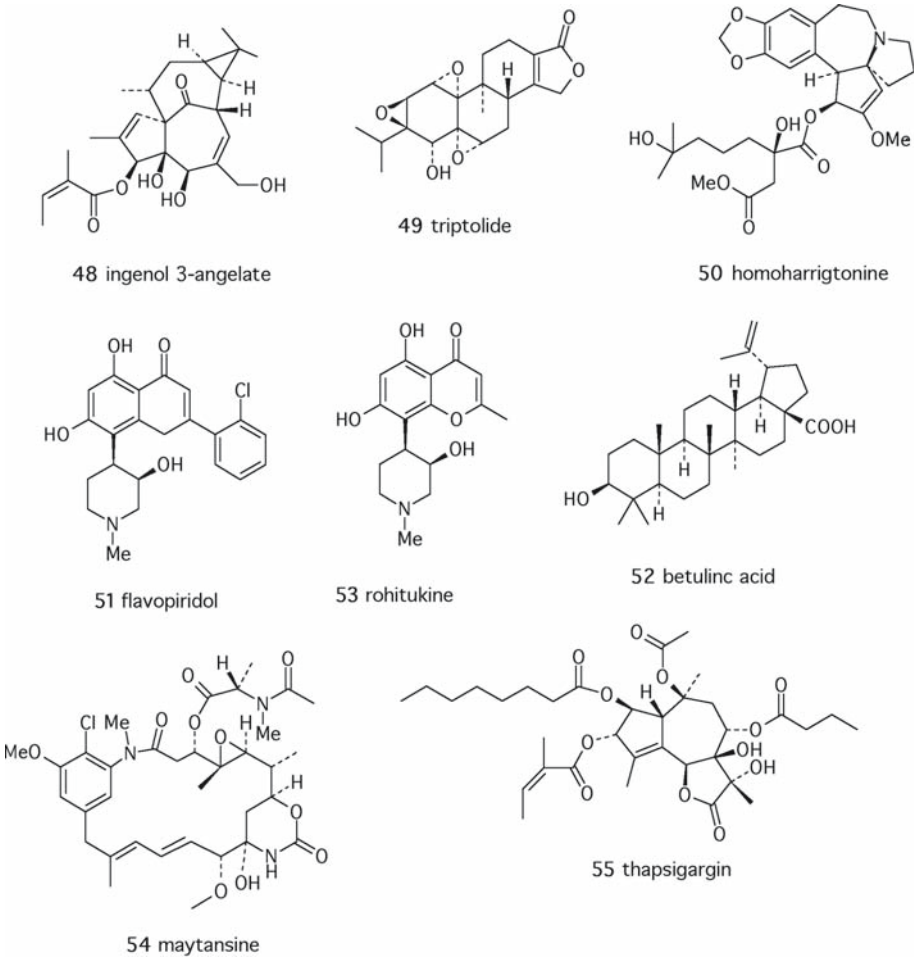
3.9 Ingenol 3-angelate

The discovery of ingenol 3-angelate (**48**) is a remarkable example of taking advantage of information available from the popular medicine. The latex of the common weed *Euphorbia peplus* (Euphorbiaceae) has been used as a home remedy for various skin conditions including warts and certain cancers. Clinical studies in the 1970s provided first evidences of efficacy. Particularly impressive was a case of remission of a basal cell carcinoma after home treatment with crude *E. peplus* sap [66]. The active principle of *E. peplus* juice is a tetracyclic diterpene ester, ingenol 3-angelate (PEP005) which has demonstrated topical activity against human cancer cell lines grown as subcutaneous tumors in mice [66, 67]. The compound also shows potent antiproliferative activity against a number of leukemia cell lines and blasts isolated from patients diagnosed with acute myeloid leukemia. The antitumor activity of ingenol 3-angelate is mediated through

activation of protein kinase C (PKC), a family of signalling isoenzymes involved in cell proliferation, differentiation and apoptosis. In particular, the antileukemic effects of PEP005 were shown to be PKC-δ dependent with sensitivity of cell lines correlating with expression of the enzyme. Topical PEP005 is developed by the Australian company Peplin and is currently in Phase II clinical trials for the treatment of actinic keratoses and basal cell carcinoma. Phase I trials are planned for an iv formulation in the treatment of acute myeloid leukemia and intracavitary wash in superficial bladder cancer.

3.10 Miscellaneous

This survey of plant-derived compounds in more or less advanced development is by no means exhaustive. Additional metabolites are in development or are the subject of in-depth investigations. Of particular interest is triptolide (**49**), a diterpenoid extracted from the Chinese perennial vine *Triptergium wilfordi*. Triptolide shows potent anticancer activity, in addition to remarkable efficacy in rheumatoid arthritis and strong immunossupressive properties. The mechanism of action, however, is not fully understood and toxicity largely hampered so far its clinical use [68]. A succinyl derivative of triptolide, TriptoSar (PG490-88Na), is in Phase I clinical development by Pharmagenesis, in collaboration with Pierre Fabre [23]. Another remarkable compound which has quite a long record of clinical investigation is homoharringtonine, an alkaloid found in the bark of coniferous evergreen trees of the genus *Cephalotaxus* (Cephalotaxaceae). It is a dose- and time-dependent inhibitor of protein synthesis [69]. Homoharringtonine (**50**) has already been in clinical trials in China and in the USA, and has been used in China since 1974 for the treatment of leukemia [21] [69]. However, a new set of Phase II clinical trials are now being conducted at ChemGenex (Australia/USA), focussing on the treatment of chronic myeloid leukemia and myelodysplastic syndrome [23]. Further promising compounds include flavopiridol (**51**) and betulinic acid (**52**). Flavopiridol is a synthetic compound, the structure of which was derived from the chromone alkaloid rohitukine (**53**) from *Dysoxylum malabaricum* (Meliaceae). The compound is currently in clinical development at the NCI in collaboration with Aventis Sanofi. The lupane triterpene betulinic acid



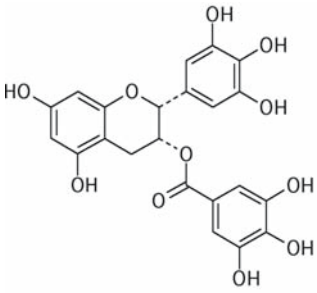
(**52**) exhibits a variety of biological activities, the most interesting being its potent antiviral properties (see below). In addition, it shows good efficacy in human melanoma xenograft models. The exact mechanism of action is still a matter of debate [70] but its activity appears to be pH-related. It is non-toxic to normal cells at pH > 7 but exerts cytotoxic effect on cancer cells at reduced pH, typically pH 6.8. In particular its lack of toxicity makes it an attractive clinical candidate in the treatment of melanoma [71].

A recurrent problem with natural products in the field of chemotherapy is the poor water solubility and high toxicity which often preclude their

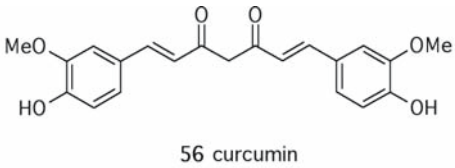
clinical use. To circumvent these difficulties, an alternative approach has been to attach the drug to monoclonal antibodies specifically targeting tumor epitopes. Two particularly promising molecules for such applications are maytansine (**54**) from the Ethiopian plant *Maytenus serrata* (Celastraceae) and thapsigargin (**55**) isolated from the Mediterranean plant *Thapsus garganica* (Apiaceae) [21]. Cantuzumab mertansine, a derivative of maytansine conjugated to a monoclonal antibody directed against the muc1 epitope expressed in a series of human cancers is undergoing Phase I clinical trials at ImmunoGen Inc in the USA [23].

3.11 Chemoprevention

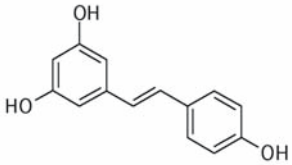
Increasing efforts in the field of chemoprevention have paralleled the search for new anticancer drugs. Prevention remains in fact the most successful approach in the fight against cancer. The interest for chemopreventive agents has been greatly stimulated by the observation that several plant secondary metabolites contained in vegetable and fruits have the potential to lower the incidence of numerous cancers [72–75]. Among the most promising compounds are curcumin (**56**), resveratrol (**57**), indol-3-carbinol (**58**) and epigallocatechin gallate (**59**). Besides their chemotherapeutic potential, these compounds have been also shown to increase the sensitivity of cancer cells to chemotherapeutic agents and could be useful in the treatment of malignant diseases as well [76]. Curcumin, a major constituent of turmeric (*Curcuma longa*, Zingiberaceae), shows chemopreventive activity in various animal models. Curcumin inhibits the initiation and promotion of tumors induced by various carcinogens. It modulates several targets and, in particular, inhibits NF-κB activation and STAT3 phosphorylation as well as AP-1 activity. Curcumin induces also apoptosis and displays antiangiogenic activity [77, 78]. Resveratrol is a stilbene found in grape skin. It is also found at high concentrations in the roots of the knotweed *Polygonum cuspidatum* (Polygonaceae). The compound inhibits carcinogenesis initiation, promotion and progression in animal models. Resveratrol possesses antioxidative properties and interferes with numerous cellular mechanisms, including the inhibition of Phase I enzymes, the induction of Phase II enzymes and the induction of cell cycle arrest and apoptosis. In addition, resveratrol is a selective oestrogen receptor modula-



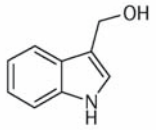
59 epigallocatechin gallate



56 curcumin



57 resveratrol



58 indol-3-carbinol

tor (SERM) [79]. Indol-3 carbinol (I3C) is a degradation product of glucobrassicin, a glucosinolate occurring in cabbage. It inhibits tumorigenesis in various animal models. Several mechanisms appear to be involved, and investigation is complicated by the fact that I3C undergoes condensation to form multiple complexes in the acidic conditions of the stomach. While the compound appears particularly promising, there are also concerns, since tumor promotion has been reported under certain conditions, in particular if I3C is given in the postinitiation stage of carcinogenesis [80]. Epigallocatechin gallate (EGCG) is the main catechin of green tea. It possesses antimutagenic properties owing to its strong antioxidative activity. EGCG inhibits protein kinase C and controls cell division. However, epidemiological studies yield contradictory results with regard to the protective effect of green tea on various cancers, and further investigations are needed to establish the real potential of green tea polyphenols [81].

Curcumin, resveratrol, indol-3-carbinol and green tea polyphenols, together with further dietary phytochemicals, are currently investigated in NCI-sponsored clinical trials for the prevention of various cancers, including prostate, breast and colon cancers [82–84].

4 Antimalarial agents

Malaria is by far the most important tropical disease with 300 to 500 million clinical cases and about 2 million deaths every year, mainly infants

and young children. The disease is caused by parasitic protoza of the genus *Plasmodium*. While different *Plasmodium* species can cause malaria, the majority of deadly cases result in cerebral malaria and other complications occurring after infection with *P. falciparum* [85]. The disease is transmitted to man exclusively by the bite of an infected female *Anopheles* mosquito. After inoculation, the sporozoites enter the human blood circulatory system and develop in the liver into merozoites. These subsequently colonize erythrocytes, where asexual reproduction occurs. The symptomatology of malaria is due to the erythrocytic phase of the parasite [7].

Quinine, extracted from Cinchona bark, was for a long time one of the only antiparasitic agents effective against malaria. Intense efforts were made during the World War II to find more accessible antimalarial agents and resulted in the development of chloroquine as a very potent drug. However, the emergence of chloroquine resistant *P. falciparum* strains led to a steady deterioration of the situation with the malaria mortality being estimated to have doubled in the last 20 years [85–87]. Consequently, intense efforts have been undertaken to find substitute drugs for controlling malaria.

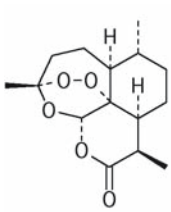
4.1 Artemisinin

The discovery of artemisinin (**60**) by Chinese scientists in 1972 was the starting point of one of the most remarkable advances in the chemotherapy of malaria. Artemisinin, originally named qinghaosu, was isolated from the herb qinghao (*Artemisia annua*, Asteraceae) in the course of a systematic investigation of plant remedies used in China. *A. annua* had been employed over 2,000 years for the treatment of fever conditions and malaria [88]. Artemisinin is a sesquiterpene lactone with a highly unusual endoperoxide group in a 1,2,4 trioxane ring that is crucial for the activity. Artemisinin is highly active and displays only low toxicity. However, its high lipophilicity led to problems with its administration as a drug. A series of hemisynthetic derivatives have been prepared by reduction and functionalization of the lactone moiety. Ethers, aliphatic and aromatic esters, and carbonates have all been synthesized. Among the most active compounds are artemether (**61**), arteether (**62**) and sodium artesuanate (**63**), the latter being considered as the current artemisinin ‘gold standard’.

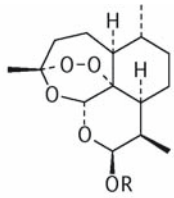
These compounds have been used extensively for more than 20 years in Asia without any serious side effect or clinically relevant resistance up to now.

The mode of action of artemisinin and its analogs involves the homolytic cleavage of the weak peroxide bond mediated by heme Fe(II) or other sources of ferrous iron within the parasite. This generates an alkoxy radical quickly rearranging to carbon-centered radical species [89, 90]. While it was initially proposed that the activity was merely due to the oxidative stress resulting of the generated radicals, in particular the C4 alkyl radical, it is now believed that the parasite death involves more specific processes and targets [89]. Both *in vitro* and *in vivo* experiments have shown that the C4 alkyl radical efficiently alkylates the four meso carbons of the heme group in an intramolecular reaction. The selective toxicity of artemisinin to malarial parasites is probably due to the accumulation of the drug into the parasite within infected erythrocytes. While the key role played by heme alkylation in the activity of artemisinin is well established, the role of this event in the parasiticidal activity remains incompletely understood [89]. A proposed mechanism involves the interaction of heme-artemisinin adducts with *P. falciparum* histidine-rich protein (HRP-II) that promotes aggregation of heme to its detoxification product hemozoin. This may hinder sequestration of toxic heme as hemozoin, thus poisoning the parasite with its own waste [89]. In addition, there is evidence that heme is not the single target. Selective alkylation of a small number of 'target' parasite proteins in particular the *P. falciparum* translationally controlled tumor protein (TCTP) [89] and PfATPase6, the sarco/endoplasmic reticulum Ca^{2+}-ATPase of the parasite, has been observed and could contribute to the activity of artemisinin [87].

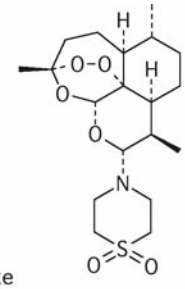
The first generation of artemisinin derivatives suffers from unfavorable pharmacokinetic properties. Artemether and arteether have a short half time as a consequence of cytochrome P450 transformation to dehydroartemisinin which is then cleared *via* glucuronidation. Artesunate is chemically unstable and hydrolyzes rapidly to dihydroartemisinin [87, 90]. Attempts have been made to develop new derivatives with improved properties, in particular by synthesising C-10 carbon-substituted analogs. A representative compound, artemisone (**64**) [91], is currently evaluated in Phase I clinical trials by Bayer AG and Medicines for Malaria Venture (MMV). Artemisone has a significantly enhanced bioavailability and a



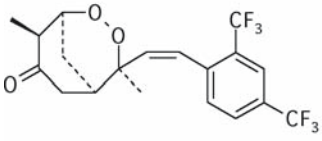
60 artemisinin



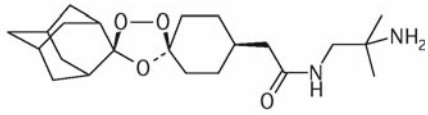
61 R = Me artemether
62 R = Et arteether
63 R = $OCOCH_2COONa$ sodium artesuanate



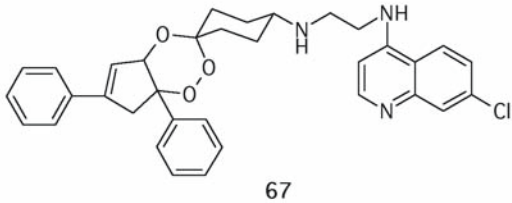
64 artemisone



65 arteflene



66 OZ 277



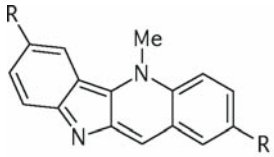
67

higher and more sustained activity compared to artesuanate. A common downside of the hemisynthetic derivatives is that they require artemisinin as starting material. Artemisinin is extracted from the plant in low yield (0.01–0.8%) [87]; this could become a supply and cost issue for a drug to be used on a large scale, in particular in developing countries. Therefore, several groups have attempted to produce fully synthetic peroxide analogs, some of which show remarkable activity. They include trioxane, tetraoxane and endoperoxide analogs such as arteflene (**65**) [87, 92]. In addition, easily synthesized 1,2,4-trioxolanes substituted with an adamentane ring were recently shown to be active in the nanomolar range, chemically stable and orally active in mice [87]. Ranbaxy Laboratories started, in 2004, with Phase I clinical trials on OZ-277 (**66**) [23]. Finally, quinoline-peroxide hybrids known as trioxaquines (e.g., compound **67**) have also been prepared [87].

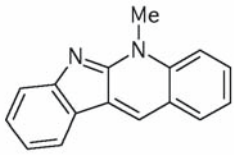
The use of artemisinin derivatives in monotherapy is associated with significant rates of recrudescence, as a consequence of their short half-life in the body [88]. For this reason, and also to minimize the development of artemisinin-resistant parasites, it is recommended that therapy with artemisinin is accompanied or followed by a second antimalarial. Therefore, a major line of development has recently been in the combination of an artemisinin analog with a second, long acting drug such as mefloquine or lumefantrine [93, 89]. Coartem®, a combination of artemether and lumefantrine developed by Novartis, received marketing approval in 1998 for use as an antimalarial in children [88]. The drug now constitutes about 70% of all clinically used artemisinin combination therapies [94] and is included in the Model List of Essential Medicines of the World Health Organization (WHO).

4.2 Cryptolepine

The remarkable activity of quinine and related drugs and the success of artemisinin stimulated the search for new plant-derived antimalarials. A large number of plants, and in particular species used in traditional medicine, have been screened for antiplasmodial activity. Various antiprotozoal compounds have been isolated but only a few were extensively evaluated for their potential as lead compounds. The quassinoids are a class of degraded triterpenes found in various species of Simaroubaceae. These compounds however display general cytotoxicity due to inhibition of protein synthesis. Attempts to find derivatives with improved selectivity against *P. falciparum* have failed [88]. A more promising lead candidate is the indoloquinoline alkaloid cryptolepine (**68**). This compound is present in relatively large amounts in the roots of *Cryptolepis sanguinolenta* (Periplocaceae), a West African climbing shrub which has been traditionally used for the treatment of malaria. The compound is highly potent *in vitro*, but originally failed to cure malaria in mice when administered orally and was toxic when given intraperitoneally. Moreover, cryptolepine exhibits moderate cytotoxicity on account of DNA-intercalation and topoisomerase II inhibition. These properties might have precluded its consideration as an antimalarial lead compound, but the finding that the mechanism of the antiplasmodial action was different from that of the cytotoxic proper-



68R = H cryptolepine
69R = Br



70 neocryptolepine

ties has stimulated the search of more selective derivatives. The synthesis of cryptolepine is straightforward and a large number of analogs can be readily prepared. The most promising compound, 2,7-dibromocryptolepine (**69**) is approximately nine-fold more potent that cryptolepine and shows no apparent toxicity in mice [85]. Recent studies suggest that other mechanisms, in addition to hemozoin formation, could be involved in the activity of this compound and other cryptolepine analogs. Particular interest is also paid currently to the isomer neocryptolepine (**70**), a minor constituent of *C. sanguinolenta* [88].

5 Antiviral agents

5.1 HIV

Acquired immunodeficiency syndrome (AIDS) is a devastating disease caused by the human immunodeficiency virus (HIV) and has led to more than 25 million deaths worldwide since the first cases were reported in 1981. HIV is a retrovirus which infects and destroys human CD4 T lymphocytes, thereby weakening the immune system and rendering the patient vulnerable to various opportunistic infections and other illnesses ranging from pneumonia to cancer. There are two forms of the virus, HIV-1 and HIV-2; HIV-1 is the cause of the global AIDS pandemic. Despite continuous advances in antiretroviral chemotherapy, the epidemic is still expanding and more than 60 million people have been infected. AIDS has become the leading cause of death in Sub-Saharian Africa and the fourth worldwide [95].

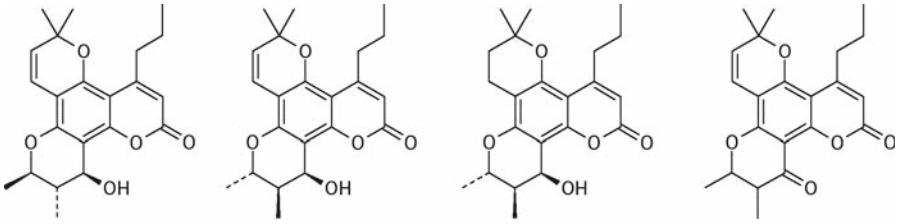
Four classes of drugs are currently in use: nucleoside analog reverse transcriptase inhibitors, protease inhibitors, non-nucleoside reverse transcrip-

tase inhibitors [96] and one fusion inhibitor [97]. A major breakthrough in AIDS therapy was the introduction, in the late 1990s, of drug combinations from two or more classes of inhibitors, a strategy termed highly active antiretroviral therapy (HAART). While this approach enables a successful suppression of HIV plasma levels in patients over many years, latent reservoirs of the virus persist which prevent an eradication of the virus. Further challenges faced by HAART are toxicity and the emergence of drug resistance as a consequence of incomplete suppression of viral replication. In view of the global threat of HIV/AIDS, there is an urgent need for novel drugs targeting different stages of virus replication.

During the last two decades, large collections of plants constituents and extracts have been assayed for anti-HIV activity. In one of the largest screening programs, the NCI tested from 1987 to 1996 over 30,000 plant extracts in an *in vitro* cell-based anti-HIV assay [98]. These screening efforts resulted in the discovery of a considerable number of compounds exhibiting *in vitro* anti-HIV activity. While there is still no anti-HIV drug of plant origin on the market, several plant metabolites proved to be valuable lead compounds. Some of these compounds or derivatives have advanced to clinical or preclinical development.

5.2 Calanolides

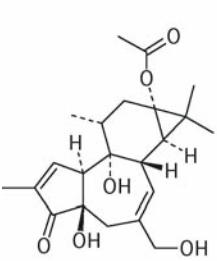
Calanolides were discovered during the anti-HIV screening program of the National Cancer Institute (NCI). An extract of leaves and twigs of the tree *Calophyllum lanigerum* (Clusiaceae) collected in Sarawak, Malaysia, in 1987 yielded the coumarin calanolide A (**71**) which exhibited potent antiviral activity [99]. Unfortunately, attempts to relocate the original tree failed, and investigation of other specimen of *C. lanigerum* gave only traces of calanolide A [98]. At the same time, a large survey of *C. lanigerum* and related species led to the discovery of a diastereomer, calanolide B (**72**), from the latex of *C. teysmanii*. Calanolide B is slightly less active than calanolide A, but has the advantage to be readily obtainable in 35–40% yields from a sustainable source. The calanolides were licensed by the NCI to Medichem Research Inc and are now being developed by Sarawak Medichem Pharmaceuticals, a joint venture company between the Sarawak State Government and Medichem Research Inc. Calanolides act as non nucleoside reverse



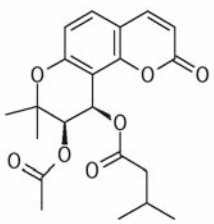
transcriptase inhibitors. The clinical development of calanolide A started in 1997. The compound is now obtained by synthesis and is currently in Phase II trials focussed on evaluation of its long-term activity in combination with other anti-HIV agents. Calanolide B (costatolide), dihydrocalanolide B (**73**) and oxocalanolide (**74**) are in preclinical development [95]. Calanolides A and B were also found to be active against *Mycobacterium tuberculosis* [100]. This property is unique among antiviral agents and could be of great value for the treatment of patients infected with both HIV and tuberculosis [23]. Tuberculosis has become the first opportunistic infection affecting HIV-infected people in developing countries.

5.3 Prostratin

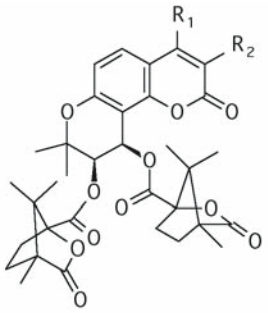
The phorbol ester prostratin (**75**) was isolated at the National Cancer Institute in 1992 from the wood of *Homalanthus nutans* (Euphorbiaceae), a tree used by traditional healers in Western Samoa for the treatment of viral hepatitis [101]. Prostratin is an activator of protein kinase C, but devoid of the usual tumor promoting properties of phorbol esters. This unique feature is due to the lack of a hydroxyl group at C-12. Similar to other phorbol esters, prostratin inhibits *de novo* infection and viral spread at the entry/fusion step of viral cycle probably through interaction with a cellular target such as an HIV receptor and co-receptors [102]. In addition, prostratin stimulates HIV replication in latently infected cells [102] through activation of the NF-κB signalling pathway *via* stimulation of one or several members of the PKC family [103]. This feature could be exploited to eliminate latent viral reservoirs which persist under HAART antiviral



75 prostratin



76 suksdorfin



77 R_1=R_2 = H
78 R_1 = CH_3 R_2 = CH_2OH

therapy by rendering latently infected cells more susceptible to targeted destruction by the immune system or by other therapeutic agents. The further development of prostratin is conducted by the AIDS ReSearch Alliance of America (ALA) which signed an agreement with the government of Samoa. Preclinical studies are ongoing and, should the compound prove to be safe, prostratin could become in the future a highly promising clinical candidate for adjunctive use in AIDS therapy.

5.4 Khellactones

The pyranocoumarin suksdorfin (**76**) was isolated in 1994 as the active principle of the fruit of *Lomatium suksdorfii* (Apiaceae). The plant which is known as Suksdorf's Desert parsley grows on the West Coast of the USA [104]. While the activity of suksdorfin was moderate, modifications at 3',4' positions yielded 3'-R, 4'-R-diO-(-)-camphanoyl-(+)-cis-khellactone (**77**, DCK) which showed improved activity. Further optimization led to 4-Me DCK [105] and the preclinical candidate 3-hydromethy-4-methyl DCK (**78**), which is a nanomolar inhibitor orally bioavailable in rats and dogs [95, 106]. DCK is a unique HIV reverse transcriptase inhibitor that inhibits the DNA-dependent DNA polymerase activity without affecting the RNA-dependent DNA polymerase activity [107]. The most promising derivative, 3-hydromethyl-4-methyl DCK, is currently in preclinical development at the American company Panacos Pharmaceuticals Inc.

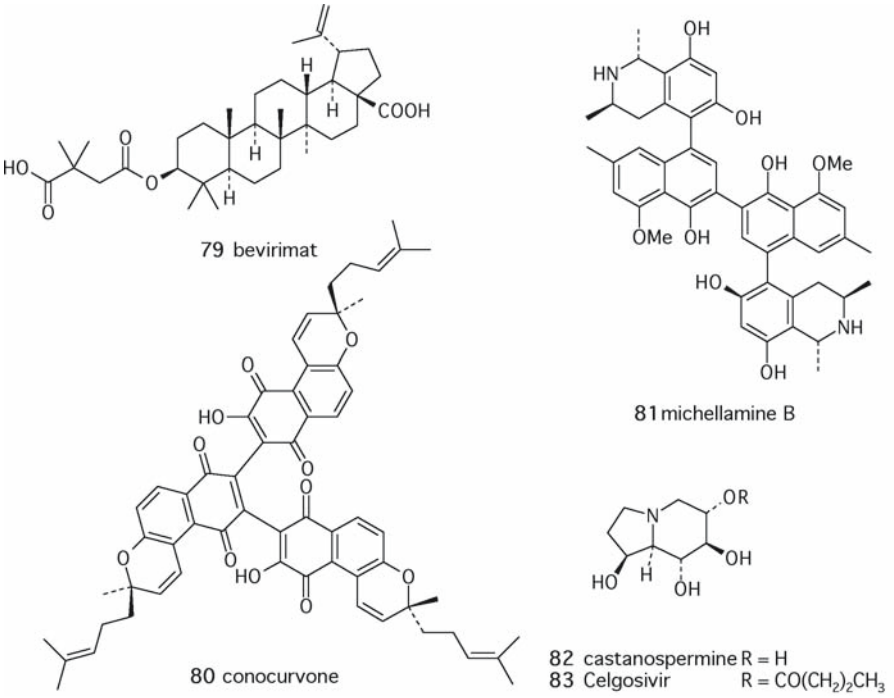
5.5 Betulinic acid

The lupane triterpene betulinic acid (**52**) was originally identified as a weak inhibitor of HIV replication [108]. However, a structural optimization program at the University of North Carolina was successful in affording more active derivatives. In particular, esterification at C-3 resulted in promising compounds with strongly improved activity and toxicity index values. Among these, 3-O-(3,3'dimethylsuccinyl)betulinic acid (Bevirimat, PA-457) (**79**) has been licensed by Panacos Pharmaceuticals Inc. Phase IIb clinical trials have been initiated in June 2006. PA-457 is one of the most promising compounds being evaluated for the treatment of HIV infection. It is the first-in-class of a new type of HIV drug called maturation inhibitors. It disrupts core condensation by targeting a late step in the processing of the viral protein known as Gag, resulting in the formation of defective, non-infectious virus particles [109, 110].

The activity of betulinic acid was first discovered when it was isolated together with platonic acid as the active principle of the Taiwanese herb *Syzigium claviflorum* (Myrtaceae) [108]. However, betulinic acid can be more readily extracted from ubiquitous plane and birch trees.

5.6 Miscellaneous

In addition to the drug candidates presented above, some further promising leads were in early development but have been abandoned. Conocurvone (**80**), a trimeric naphthoquinone isolated from the Australian shrub *Conospermum incurvum* (Proteaceae) [111] has been under joint development by the Australian company AMRA and the NCI, but its investigation has been discontinued [98]. The naphthylisoquinoline alkaloid, michellamine B (**81**), from the African liana *Ancistrocladus korupensis* (Ancistrocladaceae) [112] was very potent but proved to be too toxic at doses required for antiviral therapy. Studies aimed at clinical development were recently terminated at the NCI [98]. Finally, a semi-synthetic derivative of castanospermine (**82**), a α-glucosidase inhibitor from the Australian tree *Castanospermum australe* (Fabaceae) has been evaluated by Aventis for the treatment of HIV. However, problems were encountered with achieving therapeutic serum concentrations of the drug. Castano-



spermine now appears more promising as a drug against hepatitits C virus HCV (see below).

5.7 HCV

Hepatitis C virus infection (HCV) is as major global health concern. It is estimated than about 2.2% of the world population, corresponding to 170 million people, are infected with HCV. In the United States alone, over 4 million people are infected. After an acute phase, up to 80% infections may develop into chronic hepatitis which leads to cirrhosis in 10–20% of patients, and to hepatocellular carcinoma in 1–2% of cases. Current therapy is based on α-interferon, with or without ribavarin, but the response rate of only 40–50% of chronic patients remains unsatisfactory. Moreover,

the treatment is very expensive and associated with numerous side effects [113]. A generally useful vaccine appears to be a distant prospect due to the high genetic variability of the virus. Hence, there is an acute need for more effective drugs.

5.8 Castanospermine

Castanospermine (**82**) is an iminosugar derivative with a tetrahydroxyindolizidine structure occurring in the toxic, chestnut-like seeds of the evergreen Australian tree *Castanospermum australe* (Fabaceae), commonly called Moreton Bay Chestnut. Castanospermine inhibits glucosidase I, an enzyme involved in post-translational glycoprotein maturation which holds promise as an antiviral target. This host enzyme is implicated in the processing of oligosaccharide chains during the biosynthesis of viral envelope glycoproteins. These are essential for virion assembly and secretion and/or infectivity [114]. Castanospermine exhibits *in vitro* and *in vivo* activity against several viruses including HIV, herpes simplex virus (HSV), HCV and bovine viral diarrhea virus (BVDV), often used as a surrogate for HCV. One problem with castanospermine is that it also inhibits intestinal sucrases and causes osmotic diarrhea. In contrast, Celgosivir (**83**), a semisynthetic 6-butanoyl derivative, is relatively inactive against sucrases. Celgosivir, a prodrug, is more efficiently absorbed than castanospermine and is rapidly and extensively converted in intestinal epithelial cells into the parent compound castanospermine. Celgosivir has been evaluated in clinical trials against HIV-1 and HCV. Celgosivir has been licensed to the Canadian Company MIGENIX from Virogen which in turn had licensed it from Aventis. The agent has been selected for further development as a treatment for HCV-infection and is currently in Phase II clinical trials to evaluate its activity in patients with chronic hepatitis [23]. Phase IIb clinical trials of Celgosivir in combination with interferon and Ribavirin have been announced for the second half of 2006. A potential asset of Celgosivir is that it inhibits a viral target which may be less susceptible to the development of drug resistance. On the other hand, potential effects on host cell glycoproteins need to be carefully evaluated.

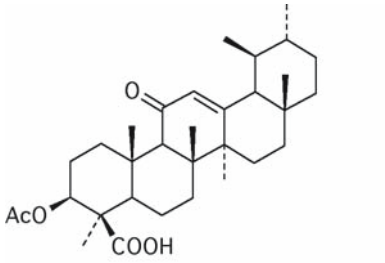
6 Inflammation, nociception, immune system

6.1 Boswellic acids

Remedies containing francincense or olibanum, the gum resin of *Boswellia* species (Burseraceae), belong to the oldest pharmaceutical preparations with written records. Francincense (from *B. carterii*) was used in ancient Egypt, Greece and Rome, whereas the resin of the Indian *B. serrata* figured in the Ayurvedic text books as salai guggal and was used for treatment of respiratory and gastrointestinal problems as well as diseases of the central nervous system [115]. Olibanum was an officinal herbal drug in European countries, but disappeared from the Western drug inventory in the early 20th Century. In 1986, Indian scientists described the anti-inflammatory properties of olibanum in animal models [116]. The inhibitory effect of the resin on leukotriene synthesis was reported shortly after, a discovery that received much attention and revived the interest in this old drug.

The oleogum resin is a complex mixture of volatile and non-volatile isoprenoids and polysaccharides. Overall, more than 200 different compounds have been identified in the oleoresins of different *Boswellia* species. Among these, mainly the pentacyclic triterpenoids are pharmacologically relevant [115]. First experimental evidence for the anti-inflammatory properties of *Boswellia* was obtained with an ethanolic extract of the resin. Oral administration inhibited the carageenan-induced odema in in rats and mice, and dextran-induced odema in rats [116]. Shortly after, boswellic acids were found to be specific, non-redox inhibitors of 5-lipoxygenase (5-LO) [117]. They decreased the formation of LTB4 in calcium stimulated polymorphnuclear neutrophils. Acetyl-11-keto-β-boswellic acid (**84**) (AKBA) was the most effective, with an IC_{50} of 1.5 μM. AKBA acts directly on the 5-LO enzyme at a site selective for pentacyclic triterpenes which is different from the arachidonate substrate binding site [118]. AKBA also inhibits the activity of human leucocytic elastase, reduced the oxidative burst in stimulator-sensitive human blood PMN preparations, and inhibited the generation of TNF-α in stimulated monocytes [119].

Several clinical studies have been conducted in severe chronic inflammatory diseases related to autoimmune disorders, but given that they mostly had a pilot character, the results have to be considered with some caution. Certainly, more extensive confirmatory trials are needed. Indica-



84 acetyl-11-keto-β-Boswellic acid

tions studied included chronic rheumatoid arthritis, osteoarthritis, chronic inflammatory bowel disease, ulcerative colitis, Crohn's disease, and bronchial asthma. These trials which were typically conducted against placebo or standard therapy mostly provided indication of efficacy. A detailed overview on clinical trials is given by Ammon [115]. Although there is no rigorous analysis of side effects, the frequency and severity of adverse reactions with *Boswellia* extracts appears to be low [118]. An interesting anecdotal use of *Boswellia* has been in astrocytoma, where extremely high doses have been taken over extended periods without serious adverse effects [120, 121]. In summary, *Boswellia* resin extracts and, in particular, AKBA as the major active principle, possess a pharmacological profile which may be of interest in a range of inflammatory diseases. However, clinical studies were performed only with resin extracts in a galenical formulation with poor bioavailability. Enriched extracts [122], pure AKBA, and formulations with optimized bioavailability are needed to develop this ancient but highly interesting Ayurvedic drug into a modern pharmaceutical.

6.2 Curcumin

Curcumin (**56**) is a major compound in *Curcuma longa* and *C. xanthorrhiza* (Zingiberaceae). Both species play an important role in Asia as spices and herbal drugs. Biogenetically, curcumin belongs to the diarylheptanoids which are derived from two phenylpropanoid moieties and an additional C-atom. Due to the extensive conjugation, the compound is an intensely yellow pigment which confers the characteristic color to the rhizome. Uses

in traditional Indian medicine have been, among others, in biliary and hepatic disorders, rheumatism, diabetic wounds, and anorexia.

First experimental data on the anti-inflammatory properties of curcumin were reported in the early 1970s. The compound was active in the carrageenan-induced rat paw odema, and in various animal models for arthritis and chronic inflammation [123]. The mechanisms involved in the anti-inflammatory activity of curcumin are complex and include downregulation of transcription factors, and of the expression of cyclooxygenase-2, lipoxygenase, inducible NO synthase, TNF, matrix metalloproteinase-9, and cell surface adhesion molecules [124]. The compound inhibits TNF-induced IκBα kinase complex and Akt activation, leading to suppression of events required for NF-κB gene expression [125]. Part of the pharmacological profile is also due to the antioxidant properties of curcumin which has been extensively studied *in vitro* and *ex vivo* [126].

Clinical studies with curcumin had mostly a pilot character. Early studies were conducted in India and are not readily accessible (for an overview, see [123]). More recently, pilot studies with indications of beneficial effects were carried out in patients with inflammatory bowel disease [127] and with hepatic steatosis [126].

In rodents and humans, curcumin has low oral bioavailability and may undergo intestinal metabolism. Absorbed compound undergoes rapid metabolization to hexahydrocurcumin and to Phase II conjugates which are mostly excreted in the bile [124]. Curcumin appears to be a safe compound. Among others, a preclinical toxicity study conducted at the NCI with doses up to 3.5 g/kg body weight over 3 months in rats, monkeys and dogs did not reveal toxic effects [124]. Safety assessment in clinical studies with daily doses between 2 and 8 g/day for up to 4 months did not show any discernible toxicity.

In summary, curcumin is a readily available natural product which has some therapeutic potential in inflammatory intestinal and liver diseases. However, larger controlled clinical trials are needed to substantiate its use as a drug substance in these indications.

6.3 Capsaicinoids

The vanilloid receptors belong to the transient receptor potential (TRP) superfamily which consists of non-selective cation channels that are

involved in a wide range of physiological functions [128]. TRP channels transduce chemical and physical stimuli into neuronal activity and play a major role in chemically and physically evoked sensations such as heat, cold, olfaction, mechanosensation, and nociception. Six vanilloid receptors have been identified so far. The vanilloid receptor type 1 (TRPV1), a sensor for hot temperatures, has attracted most interest. It is expressed on unmyelinated C-fibers, in neurons of the central nervous system, and, among others, in the dermis and epidermis, the gastrointestinal tract, the lung, and in the urinary bladder. Altogether, the TRPV1 receptor seems involved in diverse physiological functions, but nociceptor sensitization by inflammatory agents appears the most obvious and therapeutically attractive. TRPV1 activation results in a hyperalgesic condition and triggers the release of proinflammatory mediators [129]. The expression is upregulated in inflamed tissue. In addition to its contribution to neurogenic inflammation, the TRPV is of interest for the treatment of neuropathic, post-operative and chronic pain. The receptor is a tetrameric membrane protein, each subunit consisting of six putative transmembranar segments and intracellular N- and C-domains. Several drug binding sites have been described for the TRPV1 receptor.

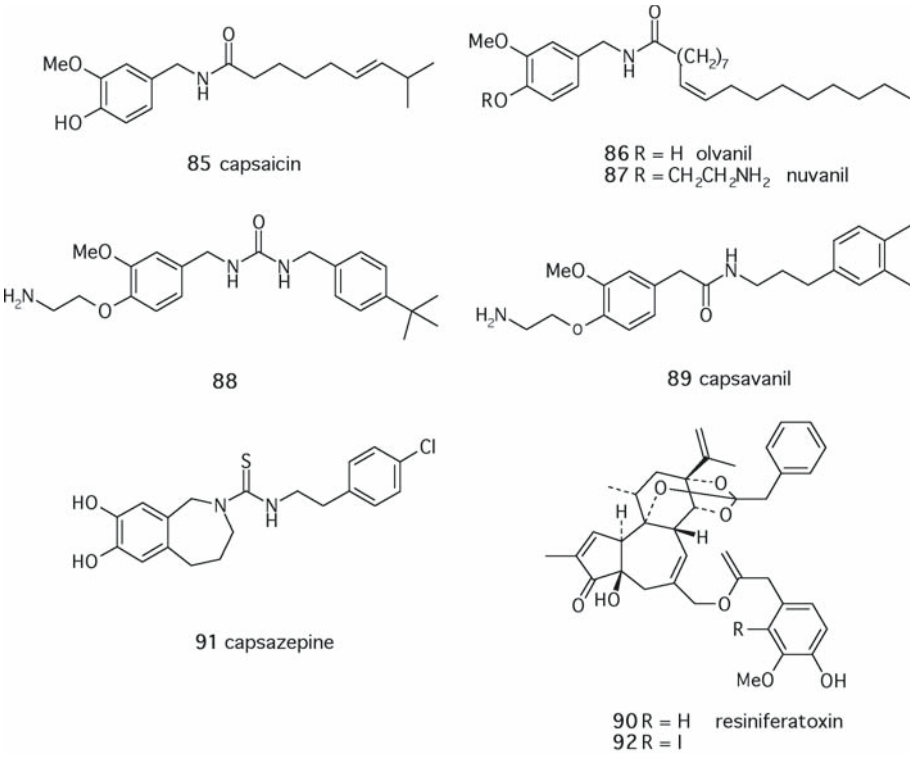
Numerous natural products bind to vanilloid receptors [129]. The prototypical agonist of TRPV1 is capsaicin (**85**). The major pungent principle in *Capsicum* was isolated in 1846 and *Capsicum* extracts were proposed as instant relief for toothache in the mid of the 19th Century. The Ca^{2+} influx caused by capsaicin releases neuropetides such as substance P and calcitonin gene-related peptides (CGRP) which are thought to be mainly involved in the resulting neurogenic inflammation. On the other hand, capsaicin produces degeneration of a large portion of C-fibers, leading to a prolonged analgesic effect [130]. Capsaicin is used as topical ointment formulation for neuropathic and osteoarthritic pain. The drawback of such formulations is in the difficulty of containment of capsaicin action. NeurogesX is currently developing a dermal patch system called NGX-4010. Efficacy has been studied in postherpetic neuralgia and in HIV sensory neuropathy and is currently in Phase III evaluation [131]. ALGRX-4975 is a high-concentration injectable formulation of capsaicin which is being developed by AlgoRX Pharmaceuticals. The intended uses are in management of pain associated with osteoarthritis, tendonitis and postsurgical conditions. Phase II clinical trials in osteoarthritis were underway in 2004.

The formulation is injected directly into the site of the pain or instilled into the wound during surgery. It is thought that a single administration provides analgesia for extended periods of up to several months [131]. *Cis*-capsaicin, the unnatural isomer of capsaicin is in clinical development as an intranasal spray known as WL-1001. Phase II trials with a 0.025% spray formulation in migraine and cluster headache apparently were encouraging and further Phase II and III studies were planned in 2004. A topical formulation listed as WL-1002 was in Phase III clinical trials for the management of ostheoarthritic pain.

Capsaicin has poor oral bioavailability, is highly pungent, and relatively toxic upon systemic administration. Therefore, there has been considerable interest in the synthesis of improved analogs. Modification of the acyl moiety led to a first generation of analogs. Olivanil (**86**) and nuvanil (**87**) were synthesized by Procter & Gamble in the 1980s as TRPV1 agonists with low pungency and were aimed as a systemic analgesic. However, Phase II trials were abandoned in the mid-1990s [131]. SDZ-249-665 (**88**) is another low pungency analog which was synthesized by Sandoz and was further pursued by Novartis [132]. Capsavanil (**89**) an amide of homovanillic acid, is in Phase II clinical trials as topical analgesic [133].

The most potent natural TRPV1 agonist (IC_{50} 10 pM) known so far is resiniferatoxin (RTX) (**90**), a diterpenoid ester from the latex of *Euphorbia resinifera* (Euphorbiaceae). Interestingly, there are historical accounts on the use of dry latex for local analgesia in toothache [134]. The homovanillic moiety, the C-3 keto group and the orthoesterphenyl moieties seem essential structural elements for activity. RTX is being developed as a desensitizing agent for diabetic neuropathy and the treatment of urinary incontinence [128]. A major difficulty in the development of RTX is in the supply situation. The minor compound has to be purified from the natural source.

Competitive TRPV1 antagonists have been synthesized by modification of the agonist templates. Capsazepine (**91**) was the first antagonist identified. The compound attenuated hyperalgesia in guinea pigs but performed poorly in rodents and was not carried to clinical development [135]. Introduction of an iodine atom at C-5 of the vanilloyl moiety of RTX leads to a potent agonist (IC_{50} 3.9 nM). 5-iodiRTX (**92**) is currently in clinical trials as an analgesic [128].

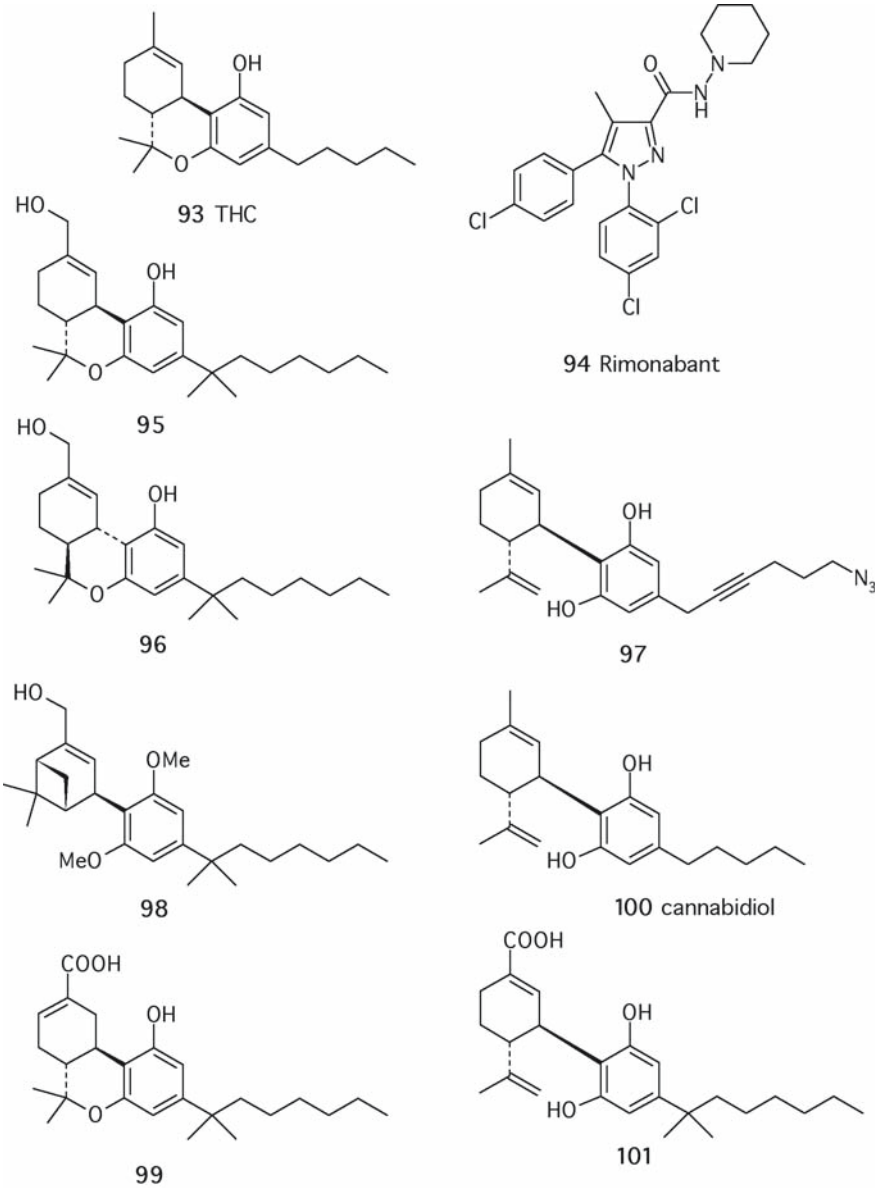


6.4 Cannabinoids

Cannabis sativa (Cannabaceae) is one of the oldest cultured plants as can be ascertained by archeological remains of the Neolithic period. First records of medicinal use were by the Chinese emperor Huang Ti who advised taking *Cannabis* for the relief of cramps and pain. However, in 1964 only the psychotropic substance in *Cannabis* was identified as Δ9-tetrahydrocannabinol (Δ9-THC) (93) [136]. First indication of the existence of specific cannabinoid receptors were reported in 1988 [137], and the CB_1 receptor was cloned in 1990 in a screening of orphan G-protein-coupled receptors. The CB_2 receptor was identified shortly after [138]. Endogenous ligands of CB receptors are derivatives of polyunsaturated long-chain fatty acids. Both cannabinoid receptors are G-protein coupled and modulate adenyl-

ate cyclases (mostly inhibitory), and mitogen-activated kinases (mostly activation). CB_1 receptors inhibit voltage-gated Ca^{2+} channels, and stimulate inwardly rectifying K^+ channels [139]. CB_1 receptors are most abundant in the CNS and also occur in peripheral nervous system, whereas CB_2 receptors are mostly restricted to immune tissues and cells [138]. Their physiological roles in cellular and humoral immune response show that CB_1 agonists could have possible clinical use in chronic, inflammatory and neuropathic pain, whereas CB_1-receptor antagonists/inverse agonists are of interest for the treatment of obesity and tobacco dependence, and possibly in palliative care of Parkinson's and Alzheimer's disease [140]. CB_2 agonists may be useful in inflammatory and neuropathic pain. Cannabinoid receptor inactive cannabinoids have shown promise as analgesic and anti-inflammatory drugs. Rimonabant (**94**), a selective CB_1 antagonist with application in obesity treatment, is the first approved drug targeting the endocannabinoid system, although it is not derived from a cannabinoid template.

Δ9-THC is a non-specific agonist at CB_1 and CB_2 receptors. Medicinal chemistry programs have aimed at conferring specificity and eliminating the psychotropic effects. HU-210 (**95**) is an ultrapotent CB_1 and CB_2 receptor agonist, whereas HU-211 (**96**) is CB receptor inactive and currently in Phase III clinical trial as a neuroprotective compound. O-2654 (**97**) has been reported as a CB_1-receptor selective antagonist. HU308 (**98**) is a selective CB_2 agonist [138]. Ajulemic acid (**99**) is a THC-11-oic acid analog with potent analgesic and anti-inflammatory activity which, however, still shows some effects on CNS mediated functions due to its ability to bind to both CB receptors, in particular to CB_1 [141, 142]. (–)-Cannabidiol (CBD) (**100**) interestingly possesses strong anti-inflammatory activity but no affinity to CB receptors. The CBD analog HU320 (**101**) was synthesized in an effort to produce cannabinoid derived molecules with increased anti-inflammatory and low CNS-related side effects [143]. HU-320 shows very low affinity to CB_1 and CB_2 receptors (K_i 1.9 and 5.0 μM, respectively), in contrast to the (+)-enantiomer (K_i 6 and 156 nM) [144]. The compound was evaluated in a number of *in vitro* and *in vivo* models. It showed anti-arthritic properties in collagen-induced arthritis in mice, and downregulated the release of important mediators such as TNF and IFN-γ, and suppressed proliferative response of lymph node cells. The compound also shows some antioxidative activity [142]. Given that cannabidiol has



extremely low toxicity *in vivo*, its metabolites and CBD acid derivatives are thought to be non toxic. The molecular mechanisms of CBD and HU-320 are still an enigma.

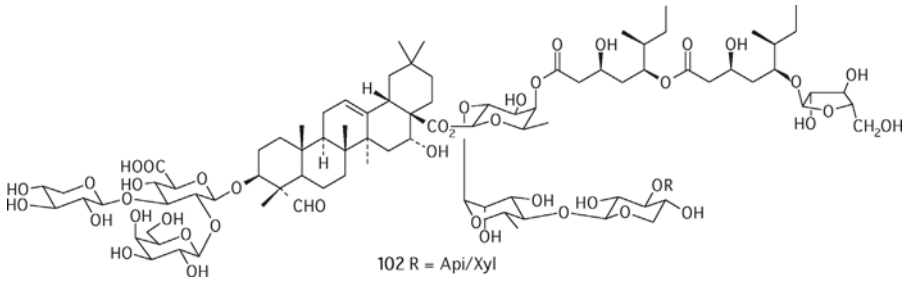
6.5 *Quillaja* saponins

Saponins are glycosides containing a triterpene or steroidal aglycone. Due to their amphiphilic properties, saponins are natural detergents and foaming agents with a number of applications. Among the most important sources for industrial saponins is *Quillaja saponaria* (Rosaceae), a tree which is indigenous to Chile. Its bark contains approximately 5% of a complex mixture of saponins. Crude *Quillaja* saponin is used as a technical surfactant, but also as a foaming agent in beverages and natural emulsifier in foods. *Quillaja* extracts are approved for human consumption in the USA, EC and Japan [145].

Quillaja saponins are bidesmosidic triterpenoids bearing sugar moieties at C-3 and C-28. The major aglycone is quillaic acid. A particular feature of *Quillaja* saponins are the aldehyde group attached to C-4, and a structurally elaborate branched acyl moiety with a terminal L-arabinose [146].

The adjuvant properties of saponins were observed as early as in the 1930s, and use of saponins as adjuvants in animal vaccines was first tested in 1951 [145]. Later, it was shown that only *Quillaja* saponins were effective adjuvants, and a fraction with reduced toxicity was purified and commercialized as Quil-A for the use in animal vaccines. Quil-A could be further separated by HPLC into four major fractions with varying degrees of toxicity. A fraction termed QS-21A$_{api/xyl}$ (**102**) emerged as being among the most promising new adjuvants, given its high level of potency and favourable toxicity profile [147].

QS-21A elicits both Th1- and Th2-type cytokines and amplifies T-cell and B-cell mediated immune responses. It is an immune-stimulatory adjuvant which possibly facilitates the uptake of the antigen into antigen-presenting cells [148]. Following complete elucidation of the complex structure of QS21A saponin, structure-function studies were carried out to define features which were essential for immunogenicity. Modification at the carboxylic group of the glucuronyl moiety led to a moderate decrease in antibody titers, but the aldehyde function in the aglycone proved to be essential [149]. Further studies with deacylated QS-21A showed a decrease in IgG1 response, and a loss of IgG2 and cytotoxic T lymphocytes (CTL) response [150], whereas derivatives of the deacylated saponin bearing an aliphatic amide at the carboxyl moiety stimulated IgG2 and CTL [151]. QS-21 is in clinical trials with a number



of vaccines and has been tested so far in over 60 trials with a total of more than 3,000 patients [148, 150].

Quillaja saponins are able to form so-called immune-stimulating complexes (ISCOMS). These are cage-like, stable complexes of cholesterol, phospholipid and saponin which can serve as carriers for amphipathic protein antigens such as viral envelope proteins [152]. ISCOMS with *Quillaja* saponins have been shown to stimulate both humoral and cellular immune response, similar to the effects of the saponins. However, the adjuvant effect was greater than for the micellar form. An interesting field of possible application of ISCOMS is in the development of mucosal vaccines [153]. In view of possible pharmaceutical application in humans, the stability of these bidesmosidic *Quillaja* saponins has been investigated. At a pH of 5.5–6, stability was satisfactory over 2 years. Currently, economically feasible supply of highly purified QS-21 seems to be a major obstacle in the commercial development of this promising adjuvant.

7 Cardiovascular and metabolic disorders

Atherosclerosis is the singly most important pathological process in the development of coronary heart disease, which today is the most common cause of morbidity and death in developed nations. Major atherosclerotic risk factors are diabetes, smoking, hypertension and hyperlipidemia. Metabolic syndrome is defined as a cluster of abnormalities including obesity, impaired glucose tolerance and type 2 diabetes, atherogenic dyslipidemia, hypertension and coagulopathy. All of the components have been shown to increase the risk of cardiovascular disease. Metabolic syndrome is highly

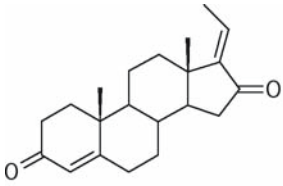
prevalent in industrialized countries; it affects approximately 24% of adults in the USA. While single drugs such as lipid lowering statins reduce the risk of fatal CHD, it is believed that improved and possibly combination therapies will be needed in the future [154].

7.1 Guggul sterols

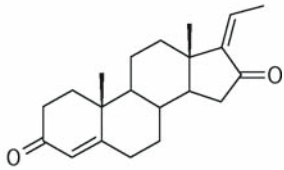
In Ayurveda, the traditional Indian system of medicine, the gum resin obtained from the tree *Commiphora mukul* (Burseraceae) is used in the treatment of inflammatory diseases and disorders of lipid metabolism. *Commiphora mukul* is a thorny and bushy tree growing in the arid zones of the Indian subcontinent. The resin is obtained *via* incisions in the bark and is collected after drying. Pharmacological studies in the 1960s corroborated the hypolipidemic properties in hyperlipidemic rabbits and were followed by a series of animal and clinical studies. An ethyl acetate extract, termed gugulipid, of the resin has been marketed in India for 20 years as a hypolipidemic drug [155].

Subsequent fractionation led to the identification of *E*- and *Z*-guggulsterone (**103**, **104**) as the hypolipidemic constituents of the resin. The concentration of the isomer mixture in the crude resin is about 2%, and approximately 5% in gugulipid. The hypolipidemic activity of the resin and of gugulsterone has been corroborated by several animal models with different species [156]. Decrease in serum cholesterol and serum triglycerides was observed. A number of clinical studies have been conducted, with gum resin as well as with gugulipid, over the past 20 years, mostly in India [155, 157]. These studies generally reported a decrease in cholesterol and triglyceride levels by 10 to 30%. However, while the decreases were significant for the study, the effect was not seen in all individuals. Even though these studies were often conducted with comparably small patient numbers and often did not meet GCP trial standards, they provide good support for clinical efficacy. A recent study, however, did not reveal a statistically significant hypolipidemic effect of gugulipid [157]. Further clinical trials are certainly warranted.

The molecular mechanism of gugulsterone action remained elusive until recently when the compounds were reported as potent antagonists at the farnesoid X receptor (FXR) [158]. FXR is a nuclear receptor critically



103 E-guggulsterone



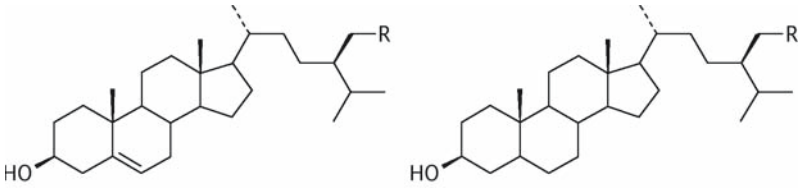
104 Z-guggulsterone

involved in the regulation of cholesterol and bile acids. Upon activation by bile acids, the natural ligands, FXR regulates the expression of genes involved in cholesterol/bile acid homeostasis and thus appears to play a key role in steroid homeostasis. Gugulsterones inhibited the activation of FXR target genes in response to bile acids and a synthetic agonist, and decreased hepatic cholesterol levels in wild-type mice fed with a high cholesterol diet, but not in FXR null mice. Despite the steroidal skeleton of the compounds, they are reportedly devoid of estrogenic, antiestrogenic and progesterone-like activity, and also do not activate or inhibit transactivation of other nuclear receptors involved in lipid metabolisms [158]). However, weak activation of the pregnane X receptor was found which possibly explains the induction of certain CYP isoenzymes.

Although no comprehensive regulatory toxicity data have been published for gugulipids, the toxicity seems low. In general, adverse effects reported for products were mild and mostly related to gastrointestinal problems. Findings concerning a possible stimulation of thyroid function are controversial so far [157].

7.2 Plant sterols, stanols and stanol esters

Phytosterols are non-nutritive plant-derived compounds with a structure similar to that of cholesterol, the difference being in a modified side chain in the plant sterols. β-Sitosterol (**105**) and campesterol (**106**) are the most abundant phytosterols. They occur in all plant membranes, and the content is particularly elevated in fat-rich vegetables and derived products such as nuts, fruits and vegetable oils. Stanols (**107**, **108**) occur as minor metabolites in plants, but are readily prepared by hydrogenation of corresponding plant sterols.



107 R = Me β-sitostanol
108 R = H campestanol

Plant sterols have been long known to reduce plasma LDL-cholesterol levels, by competing with dietary and biliary cholesterol absorption in the intestine. After a number of animal studies in the early 1950s, β-sitosterol was introduced in 1957 by Eli Lilly as a drug for the treatment of hypercholesterolemia [159]. Because of its poor solubility and bioavailability, doses of up to 18 g/day had to be given. Phytosterols were soon abandoned, but renewed interest arose more recently with the possible use of plant sterols and sterol derivatives as nutraceutical agents. Phytosterols and phytostanols are esterified with fatty acids, and these higly lipophilic esters can be readily incorporated into the lipid phase of foods such as margarines.

Metaanalysis of more than 40 clinical intervention studies with such products showed that a daily intake of 2 g of plant phytostanols lowers cholesterol levels by 10%, and by 15% with an intake of 3 to 4.2 g [159].

Sterols, phytosterols and cholesterol alike, are absorbed as micelles. Phytosterols displace cholesterol from these micelles, making it less available for reabsorption. However, bioavailability of plant sterols and stanols is low, due to poor intestinal absorption and a high biliary elimination rate. Cholesterol and plant sterols are absorbed into enterocytes *via* ABC G5 and G8 transporters, which also function as efflux transporters [160]. The uptake of plant sterols is low, and the outward transport higher than with cholesterol, resulting in an overall low resorption rate of 1.5 to 5% for sterols, and 0.05 to 0.15% for stanols [161].

The toxicity of sterols and stanols has been extensively studied. Animal studies showed no evidence of toxicity, including a lack of reproductive and genotoxicity. Sterols have been clinically used for five decades without any signs of toxicity, and the FDA and EU Scientific Committee has sanc-

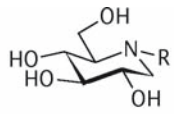
tioned their use in foods. There has been some concern about a lowered resorption of lipophilic vitamins, but reductions by 10 to 15% have been considered as within the range of individual variation [161].

7.3 Glucosidase inhibitors

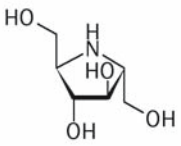
Glycosidases are enzymes that catalyze the cleavage of glycosidic bonds in oligosaccharides or glycoconjugates. The activity of glucosidases is fundamental to biological processes such as degradation of dietary polysaccharides, lysosomal glycoconjugate catabolism and glycoprotein processing, and biosynthesis of oligosaccharide units in glycoproteins and glycolipids [162]. Glycoprocessing enzyme inhibitors have also been explored in the treatment of cancers and as antiviral drugs [163] as reviewed earlier in this chapter.

Currently three compounds are used as drugs, namely acarbose, a microbial metabolite, miglitol (**109**) and N-butyl-1-deoxynojirimycin (**110**). The latter two are derivatives of iminosugars or polyhydroxyalkaloids which bear structural resemblance with monosaccharides. These compounds inhibit glucosidases because they are able to mimic the conformation and charge of the oxycarbenium ion intermediate generated in the transition state during the cleavage of the glycosidic bond. Nojirimycin was discovered in 1966 as the first iminosugar. In the meantime, numerous mono- and bicyclic iminosugars have been identified, bearing piperidine, pyrrolidine, pyrrolizidine, indolizidine or nor-tropane skeletons [164]. Representative examples include 1-deoxynojirimycin (**111**) from *Morus alba* (Moraceae) 2,5-dideoxy-2,5-imino-D-mannitol (**112**) from *Derris elliptica* (Fabaceae) mimicking D-frucofuranose, castanospermine (**82**) from *Castanospermum australe* (Fabaceae), alexine (**113**), a pyrrolizidine from the *Alexa leiopetala* (Fabaceae), and calystegine A3 (**114**) from *Calystegia sepium* (Convulvulaceae).

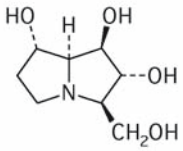
Isolation of 1-deoxynojirimycin followed after the observation that extracts of mulberry leaves were able to suppress the postprandial glucose peak. The discovery that the alkaloid inhibited mammalian α-glucosidase activity prompted development towards a drug candidate. However, the compound was only moderately active *in vivo*. Numerous derivatives were synthesized with the aim to increase potency, and miglitol (**109**)



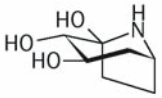
109 R = C_2H_4OH miglitol
110 R = n-Bu N-butyl-1-deoxynojirimycin
111 R = H 1-deoxynojirimycin



112



113 alexine



114 calystegine A

was finally developed to a commercial drug. N-butyl-1-deoxynojirimycin (Zavesca) (**110**) was initially developed as an anti-HIV agent. The compound inhibits processing α-glucosidase I and to impair viral entry at the level of post-CD-4 binding, probably due to an effect on components of the viral envelope, and a structural change in the V1/V2 loop region of gp-120 [164]. However, difficulties to achieve therapeutic concentrations limited the practical use as an antiviral drug. The compound was subsequently found to be a specific inhibitor of the glucosyltransferase-catalyzed biosynthesis of glucosylceramide. Deficits in glucocerebrosidase occur in type I Gaucher's disease, a hereditary lysosomal storage disorder leading to neuropathology. Zavesca and structurally related iminosugar-derivatives are being developed for substrate reduction treatment or chaperone-mediated therapy for a number of other lysosomal storage disorders [165].

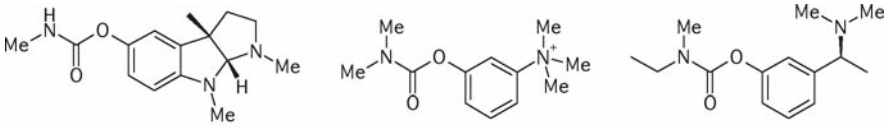
7.4 Other antidiabetic agents

Traditional Chinese medicine has a rich collection of plants that have documented effects in treating blood glucose imbalances and diabetes. Animal pharmacology and clinical studies show that these single herbs

or herbal mixtures show diverse effects, such as protection of pancreas islets and beta cells, reduction of insulin resistance, and improvement of microcirculation. However, the molecular mechanisms and compounds responsible for the activity are mostly unknown [166]. Also the traditional Ayurvedic medicine has a rich collection of plants with antidiabetic properties, some of which are corroborated by animal pharmacology and clinical data [167]. A systematic review of data available on herbs used in glycemic control of diabetes showed that there is reasonably good evidence for a number of plants, but clinical studies typically had a preliminary character [168]. The antidiabetic potential of purified plant metabolites has been reviewed but, with exception of the iminosugars, none of the plant derived compounds have been developed to commercial drugs [169].

8 Alzheimer's disease

Chronic neurodegenerative diseases are caused by a combination of events that gradually impair normal neuronal function. Among these, Alzheimer's disease (AD) is the most commonly occurring disorder. In the USA alone, 4.5 million patients suffer from AD [170]. Given that age is the most important risk factor, the prevalence will substantially increase in industrialized countries due to an increase in life expectancy [171]. The pathology of Alzheimer's disease is characterized by two major lesions, namely extracellular amyloid plaques and intracellular neurofibrillary tangles of hyperphosphorylated tau protein. As either a cause or a consequence of these pathologies, neuronal cells undergo apoptosis due to excitotoxicity, impaired energy metabolism and mitochondrial dysfunction, and oxidative stress [172]. Apoptosis of acetylcholine-containing neurons leads to characteristic deficits in cholinergic neurotransmission. Currently approved drug treatments include acetylcholin esterase (AChE) inhibitors and NMDA antagonists. The neurotransmitter acetylcholin is rapidly hydrolyzed after release at the synapse by AChE. Inhibitors increase the concentration of neurotransmitter and thus compensate, at least in part, for the deficit in cholinergic neurons. The rationale for using cholinergic drugs is that they are agonists at nicotinic cholinoreceptors and compensate for low neurotransmitter levels.



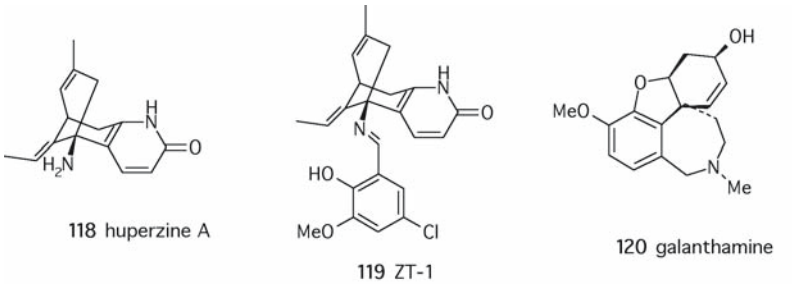
115 physostigmine 116 neostigmine 117 rivastigmine

8.1 Physostigmine

The first and prototypical AChE inhibitor was physostigmine (**115**), an alkaloid from the calabar bean (*Physostigma venenosum*, Fabaceae). Physostigmine was isolated in 1864 as the toxic principle of an African ordeal poison and was subsequently introduced as a glaucoma treatment. Recent clinical studies with the compound showed cognitive benefits in normal and AD patients, but a short half-life prevented further development. The carbamate moiety of the compound is essential for its activity. It interacts with a serine in AChE which is normally involved in the cleavage of ACh by transient formation of a covalent bond with the acetyl moiety. *Via* a nucleophilic attack of the serine, the carbamoyl group of physostigmine is covalently bound to the catalytic site of the enzyme. Subsequent hydrolysis is slow, and the alkaloid thus acts as a quasi-irreversible inhibitor. Modification of the side chain and introduction of quaternary ammonium derivatives led to peripherally active AChE inhibitors such as neostigmine (**116**) which has been used for the treatment of myasthenia gravis. Development of lipophilic analogs, able to cross the blood-brain barrier, led to rivastigmine (**117**) which was introduced in Alzheimer therapy in 2000 [173].

8.2 Huperzine A

Huperzine A (**118**) is an alkaloid isolated from the clubmoss *Huperzia serrata*. This herbal drug has been used in traditional Chinese medicine to treat various diseases. Huperzine A belongs to the class of *Lycopodium* alkaloids which are unique to clubmosses. The compound is a potent, reversible and selective inhibitor of AChE, with a 1,000-fold selectivity over



butylcholinesterase [174]. The compound has no affinity for cholinergic receptors [175]. Interestingly, the alkaloid possesses some neuroprotective properties which may be due to its ability to block excitatory action at NMDA-type glutamate receptors. Bioavailability and penetration through the blood-brain barrier are high. Several clinical studies have been conducted with huperzine A, and significant improvement over placebo was reported [174]. The template of huperzine A has been the starting point for numerous synthetic efforts to find compounds with an improved activity and safety profile. ZT-1 (**119**), a Schiff base prepared by condensation of huperzine A with 5-chloro-o-vanillin, is reportedly more selective as an AChE inhibitor and shows less toxicity than the parent compound by retaining potency, bioavailability, and duration of action. ZT-I has successfully passed Phase I clinical trials and is reportedly in a large multicenter Phase II study [176].

8.3 Galanthamine

Galanthamine (**120**) was first isolated from the Caucasian snowdrop (*Galanthus woronowii*, Amaryllidaceae) by Bulgarian chemists in 1952, and the AChE inhibitory properties of the alkaloid were discovered in 1960 [177]. Galanthamine belongs to the class of Amaryllidaceae alkaloids which are unique to this plant family. The pharmacological profile of the compound differs from that of other AChE inhibitors. In various models, the alkaloid was shown to be only a moderate inhibitor of the enzyme. It is a good ligand for nicotinic ACh receptors and appears to act as an allosteric potentiating ligand. Galanthamine also enhances GABA and glutamate

release in hippocampal slices, and dopamine in the striatum. There is also evidence for neuroprotecant properties in models using β-amyloid or thapsigargin as toxic stimuli.

Galanthamine has been clinically tested in mild to moderate AD patients. A Cochrane review concluded that the drug showed efficacy on global ratings, cognitive tests, and assessments of daily living and behavior. Functional long-term clinical studies indicate a substantial delay in cognitive decline in the longer term, up to 4 years. This observation is consistent with a neuroprotective effect that goes beyond the symptomatic effects attributed to pure AChE inhibitors.

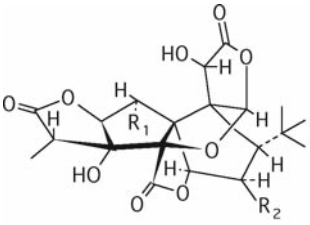
There have been limited efforts to develop galanthamin derivatives. Ring-D open analogs had lower AChE inhibitory activity than the parent compound. Based on X-ray structural analysis of the (-)-galanthamine AChE complex, some bis-ligands with significantly higher potency were designed [177]. None of these compounds seems to be tested in animal models.

8.4 Other AChE inhibitors

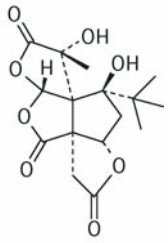
If one consults the natural products literature, numerous reports on plant derived AChE inhibitors can be found. They encompass a wide range of structural classes. One may speculate that this is the result of converging evolution, given that AChE inhibition is an effective defense strategy of a sessile organism against predators. Most compounds, however, are only moderately potent and thus not suitable starting points for drug development. Reviews on natural AChE inhibitors have been published [173, 178].

8.5 Ginkgo biloba

The special extract EGb761 is obtained from the leaves of *Ginkgo biloba* (Ginkgoaceae) by a complex extraction and fractionation process which leads to an enrichment of flavonoid glycosides (24%) and the terpenoid lactones, ginkgolides A–C (**121–123**) and bilobalide (**124**) (6%). The extract was initially commercialized as a phytomedicine to treat peripheral and



121 $R_1 = R_2 = H$ ginkgolide A
122 $R_1 = OH$, $R_2 = H$ ginkgolide B
123 $R_1 = R_2 = OH$ ginkgolide C



124 bilobalide

central vascular disorders. Recent clinical studies provided evidence of clinical benefits in AD patients, and EGb761 is approved in certain countries as drug for this indication.

Given the complex composition of EGb761 its pharmacological properties are equally complex and likely a result of synergistic effects of the constituents. The extract has antioxidant properties, inhibits synthesis and release of various mediators of inflammation, protects against neuronal death induced by various neurotoxic agents, and reduces mitochondria-initiated apoptosis [179]. EGb761 increases α-secretase activity *in vitro* and *in vivo*, and reduces glutamate-associated excitotoxicity by a non-competitive antagonism. Bilobalide appears to be responsible for this activity. The antioxidant properties have been attributed to the flavonoid fraction, whereas ginkgolides are PAF antagonists, and antagonists at glycine and $GABA_A$ receptors. Neuroprotective and cognitive enhancing properties of EGb761 have been demonstrated in various animal models [180].

Over the past decade, EGb761 has been submitted to clinical trials in AD patients. The extract had reproducible effects on cognitive function and delayed progression of the disease [181]. A large study conducted in France over 7 years suggested a preventive effect in an aged population cohort [182]. On the basis of this study, two large interventional studies have been initiated in the USA and in France [180].

Regulatory toxicology showed a remarkable safety of *Ginkgo* extract. The LD_{50} in mice, for example, is >9,600 mg/kg [179]. EGb761 had no significant adverse effects in formal clinical trials. In doses up to 240 mg/day, the drug was usually well tolerated and the significance of adverse events was similar to the placebo groups [181]. The outcome of the ongoing trials

will be critical for the future role of *Ginkgo* extracts in pharmacotherapy of AD, and in particular the possibility for preventive administration in the elderly without a clinical diagnosis of the disease.

8.6 Curcumin

The anti-inflammatory, antioxidant and cancer chemopreventive properties of curcumin (**56**) are well known and have been reviewed earlier in this chapter. There is some epidemiological evidence associating curry consumption with cognitive function and lower incidence of AD [183, 184]. Furthermore, epidemiological studies indicate a reduced AD risk associated with long-term use of non-steroidal anti-inflammatory drugs. Extending on these lines, curcumin was tested in an Alzheimer transgenic mouse model and shown to lower oxidized protein and IL-1β in the brain of these animals. Insoluble and soluble β-amyloid, and plaques were significantly reduced. Curcumin inhibits aggregation of the amyloidogenic abeta-42 oligomer and reduced amyloid levels and plaque burden *in vivo* [185, 186]. A placebo controlled Phase II clinical study sponsored by the US National Institute on Aging was initiated in 2003 to test the safety and tolerability of two doses of curcumin in patients with mild to moderate AD. Considering the safety record of turmeric as one of the most widely consumed spices, evidence of a clinically beneficial effect in AD patients might open the perspective of a preventive use of curcumin in elderly.

9 Industrial scale production of plant derived drugs – supply and conservation issues

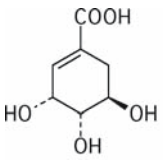
Sustainable, ecologically responsible and economically viable production of plant material is important for industrial exploitation of plant derived metabolites to be used as such or as starting material for synthesis. Typically, plant biomass is harvested from the wild or is cultivated. In both cases, the production is dependent on a number of factors that are difficult or impossible to control. Harvests may be dramatically affected by weather, pests and plant diseases. There may be considerable difficulties to cope with fluctuating needs for supply. Given the particularities of any

given plant/compound combination, the solutions that have been found to ensure supply differ considerably. A selected number of examples should highlight this.

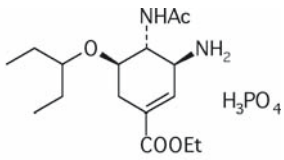
The anticancer bisindole alkaloids vinblastine (**9**) and vincristine (**10**) occur in the Madagascan periwinkle, *Catharanthus roseus*, only in minute amounts. One ton of the plant material yields about 1 g of vinblastine and 20 mg of vincristine. Cultivation of this perennial herb is not difficult, and commercial cultures were initially established in Australia, Africa and India. However, a practical synthetic route for these bisindoles, *via* oxidative coupling of the much more abundant 'monomeric' precursors vindoline and catharanthin, was developed [187]. The starting compounds for synthesis are extracted from leaves or roots, or produced in plant cell culture [188].

When the demand for *Gingko* special extract EGb761 started to grow in the early 1980s, the Dr. Willmar Schwabe group of pharmaceutical companies had to find a sustainable and secure supply for the leaf material, to replace local manual collection in East Asia. The French partner Beaufour-IPSEN then established *Ginkgo biloba* plantations in the USA and France. Within a few years, cultures of several millions trees came into operation. The plantations were optimized for mechanized cultivation and harvest. Leaves are stripped off in late summer using modified cotton pickers, immediately dried in industrial heat dryers and packed into bales. Quality management is secured by applying Good Agricultural Practice (GAP) regulations all along the production process [189].

In the development of paclitaxel (Taxol®) (**14**), supply of the drug for clinical trials already was a major challenge. Paclitaxel is only found in the thin bark of *Taxus* species, and the concentration of the compound is low. Harvest of the bark leads to destruction of the plant which belongs to the slowest growing trees, and the bark of three mature trees provides only 1 g of paclitaxel. When the first clinical trials required some 13,000 trees of Pacific yew to be felled in the unique temperate rainforests of northwestern USA, there were significant environmental concerns. It was calculated that commercial sourcing from the bark would lead to a global eradication of all yew species within few years. Ensuring a sustainable supply was quintessential for a successful commercial development. Fortunately, a suitable precursor, 10-deacetylbacchatin III was found to occur in rather high concentration in the renewable leaves, and practical routes for semisynthesis



125 shikimic acid



126 Tamiflu

of paclitaxel were found [190]. Thus, a sustainable industrial production from yew clippings could be established. Attempts to produce paclitaxel in plant cell culture started in the late 1980s. A first patent filed in 1991 described yields of 1–3 mg/l in cell suspension cultures. Commercial cell culture-based paclitaxel production was subsequently established by Samyang Genex (Korea) and Bristol-Myers Squibb/IVAX (USA). A production of 140–295 mg/l is reportedly achieved in high-density culture condition [191, 192]. Paclitaxel can be considered as the first economical success of a cell culture-based production of a pharmaceutical.

The most advanced stage for production of critical plant-derived natural products is by recombinant microbial biocatalysis. The first compound to be produced in this manner is shikimic acid (**125**) which serves as starting material for the synthesis of the neuraminidase inhibitor oseltamvir (Tamiflu®) (**126**) [193]. Shikimic acid is an intermediate in the biosynthesis of essential amino acids and phenylpropanoids, and occurs in high concentration in the fruit of star anise (*Illicium verum*, Illicaceae). The dramatically increasing demand for Tamiflu led to a supply shortage for shikimic acid. Fermentation of a genetically modified *E. coli* currently supplies approximately 30% of the demand for the USA market.

However, there are cases of successful drugs for which the supply issue has not yet been resolved. A case in point is podophyllotoxin (**21**), which is extracted from the rhizome of *Podophyllum hexandrum*. The compound has limited use as such but is the starting material for synthesis of the anticancer drugs etoposide (**22**) and teniposide (**23**). *P. hexandrum* grows naturally in most Himalayan states of India where it is collected from the wild. The plant has a slow regeneration in nature, and the species has now become endangered due to uncontrolled collection. Work on mass propagation and attempts for plant cell culture production of podophyllotoxin have been reported but not developed to industrial scale [194].

The North American species *P. peltatum* is being evaluated as an alternative source, and podophyllotoxin can be obtained from podophyllotoxin-4-*O*-β-D-glucoside contained in the leaves *via* controlled enzymatic hydrolysis [195]. However, even though high yielding accessions have been selected, economically viable cultivation is not yet achieved. There is currently no economical synthetic route to podophyllotoxin.

The case of *P. peltatum* highlights the problem of insufficient or lacking control mechanisms for endangered species that most developing countries are facing. Not only the production of prescription drugs, but also an increasing demand for certain herbal medicines and food supplements can lead to uncontrolled collection from the wild which imperils natural plant populations. Three selected examples, among others, from the African continent include *Hoodia gordonii* (Asclepiadaceae), claimed to be an appetite supressant, *Harpagophytum procumbens* (Pedaliaceae), popular as an antirheumatic phytomedicine, or *Drosera madagascariensis* (Droseraceae). The latter species is collected as a replacement for the European *Drosera rotundifolia* which is strictly protected. On the other hand, there are examples of successful implementation of sustainable harvesting procedure for sourcing from natural habitats. One case in point is *Quillaja saponaria* (Rosaceae), a tree that grows in a limited area in Chile and is harvested as a raw material for industrial extraction of *Quillaja* saponins. Originally, mature trees were felled to obtain the saponin-containing bark. Instead, harvesting of shoots and branches is sustainable and has been implemented successfully [145].

10 Plant-derived drug discovery after the convention on biological diversity

In June 1992, 150 government leaders at the Rio Earth Summit signed the Convention on Biological Diversity (CBD), which has been ratified since by a large majority of the United Nations Member States. The Convention was dedicated to promoting sustainable development and was conceived as a practical tool for translating the principles of Agenda 21 into reality.

Three objectives should be reached with the aid of the CBD: the conservation of biological diversity, sustainable development of biological

resources, and equitable sharing of resulting benefits. In accord with the principles of the CBD, numerous countries have put in place various forms of regulations and legislation which control the acquisition of biological materials and indigenous knowledge.

For academic and industrial research groups, the implementation of these regulations had profound impact on the way natural product research is carried out. Agreements to access biological materials now require lengthy negotiations. With plants, the intellectual property issues can be daunting if one considers that plant distribution does not follow national boundaries and that numerous plants have been naturalized in other countries or are exported, e.g., as ornamental plants. Possible indigenous knowledge on plants may render the intellectual property issues and benefit sharing even more complex. The hope that the CBD would lead to a higher level of international collaboration has not materialized; rather the opposite has happened. During the 1990s and after 2000, several pharmaceutical companies renounced to natural product screening, the complications of the negotiation process usually could not meet the timelines of the industry. On the other hand, there have been a number of successful cooperative agreements between academic groups, pharmaceutical industries and governmental institutions in biodiversity-rich countries [196]. Countries such as China have successfully negotiated large contracts with the pharmaceutical industry.

A major obstacle in reaching pragmatic solutions has often been the discussion about the value of biological resources and the issue of just compensation [197]. Highly divergent opinions have been voiced since the Rio Earth Summit. On the one hand, there have been accusations of 'biopiracy' [198], and numerical simulations suggesting that the bioprospecting value of certain genetic resources could be large enough to support market-based conservation of biodiversity [199]. On the other hand, there are the outcomes 15 years after the Rio Summit. The example of INBio (National Biodiversity Institute) of Costa Rica, one of the flagships of biodiversity conservation, shows that the money for bioprospection still largely comes from public research agencies and that no major royalty generating patent has resulted so far from their past and current cooperation with several leading pharmaceutical and biotechnology companies [200]. Thus, source countries should rather emphasize knowledge transfer and capacity building [201].

11 Conclusions

If one compares the compounds and projects described in this chapter with those reported in earlier reviews on the subject [9, 202–204], the continuous influx of new plant-derived molecules into the drug development pipeline is obvious. Equally important has been the progress in our understanding of the molecular targets for these compounds.

Compared to other biomes, higher plants are taxonomically well characterized and their secondary metabolism has been widely explored. Thus, the majority of secondary metabolites known today originate from plants. Should one, therefore, refrain from plant-based lead discovery and favor lesser explored sources? While there is certainly no reason not to study lesser known groups of organisms, higher plants are far from exhausted as sources for new drug leads. Indeed, several new lead templates have been discovered from plants since 1990 with compounds currently in clinical trials [23]. Also, it should not be forgotten that the vast majority of phytochemicals have been isolated in an academic setting by scientists who were primarily interested in their structural features and biosynthesis. Most compounds have never been tested for any bioactivity, and if they underwent some sort of testing, the bioassays were usually rather simple and narrow in their scope. The largest screening effort for plants and plant-derived compound has been by the National Cancer Institute (NCI) of the United States, in the fields of anticancer and antiviral drug discovery. However, all the plants and pure substances which were screened have to be considered as 'uninvestigated' with respect to any other pharmacological activity. Therefore, the continuing success of plant-derived drug discovery and development will largely depend on the quality of the research being carried out, and on the innovativeness of the scientists attempting to unravel the secrets of natural products.

References

1 Huffman MA (1997) Current evidence for self-medication in primates: A multidisciplinary perspective. *Yearbook Phys Anthropol* 40: 171–200

2 Krief S, Huffman MA, Sevenet T, Hladik CM, Grellier P, Loiseau PM, Wrangham RW (2006) Bioactive properties of plants ingested by chimpanzees (*Pan troglodytes schweinfurthii*) in the Kibale National Park, Uganda. *Am J Primat* 68: 51–71

3 Sneader W (2005) *Drug discovery – a history*. Wiley, Chichester
4 Hesse M (2000) *Alkaloide – Fluch oder Segen der Natur?* Verlag Helvetica Chimica Acta, Zürich, 322
5 Harvey AL (2002) Natural products for high-throughput screening. In: MM Iwu, JC Wootton (eds): *Ethnomedicine and drug discovery*, Elsevier, Amsterdam, 39–44
6 Sneader W (1996) *Drug prototypes and their exploitation*. Wiley, Chichester, 3–10
7 Hamburger M, Marston A, Hostettmann K (1991) Search for new drugs of plant origin. *Advances in Drug Research* vol. 20, Academic Press Ltd, 167–215
8 Farnsworth NR, Kaas CJ (1981) An approach utilizing information from traditional medicine to identify tumor-inhibiting plants. *J Ethnopharmacol* 3: 85–99
9 Hamburger M, Hostettmann K (1991) Bioactivity in plants: The link between phytochemistry and medicine. *Phytochemistry* 30: 3864–3874
10 Wolfender JL, Terreaux C, Hostettmann K (2000) The importance of LC-MS and LC-NMR in the discovery of new lead compounds from plants. *Pharm Biol* 38 (Suppl): 41–54
11 Wolfender JL, Ndjoko K, Hostettmann K (2003) Liquid chromatography with ultraviolet absorbance-mass spectrometric detection and with nuclear magnetic resonance spectrometry: A powerful combination for the on-line structural investigation of plant metabolites. *J Chromatogr A* 1000: 437–455
12 Bringmann G, Messer K, Wohlart M, Kraus J, Dumbuya K, Rückert M (1999) HPLC-CD on-line coupling in combination with HPLC-NMR and HPLC-MS/MS for the determination of the full absolute stereostructure of new metabolites in plant extracts. *Anal Chem* 71: 2678–2686
13 Niessen WMA, Lin J, Bondoux JC (2002) Developing strategies for isolation of minor impurities with mass spectrometry-directed fractionation. *J Chromatogr A* 970: 131–140
14 Reynolds WF, Enriquez RG (2002) Chosing the best pulse sequences, acquisition parameters, postacquisition processing strategies, and probes for natural product structure elucidation by NMR spectroscopy. *J Nat Prod* 65: 221–244
15 Potterat O, Hamburger M (2006) Natural products in drug discovery – concepts and approaches for tracking bioactivity. *Curr Org Chem* 10: 899–920
16 VanMiddlesworth F, Cannell RJP (1998) Dereplication and partial identification of natural products, In: RJP Cannell (ed.) In: *Methods in Biotechnology, vol. 4, Natural Product Isolation*, Human Press Inc, Totowa, New Jersey, 279–327
17 Bindseil K, Jakupovic J, Wolf D, Lavayre J, Leboul J, vand der Pyl D (2001) Pure compound libraries; a new perspective for natural product based drug discovery. *Drug Discov Today* 6: 840–847
18 Abel U, Koch C, Speitling M, Hanske FG (2002) Modern methods to produce natural-product libraries. *Curr Opin Chem Biol* 6: 453–458
19 Jia Q (2003) Generating and screening a natural product library for cyclooxygenase and lipoxygenase dual inhibitors. *Stud Nat Prod Chem* 29 (Bioactive Natural Products (Part J)): 643–718
20 Eldridge GR, Vervoort, HC, Lee CM, Cremin PA, Williams CT, Hart SM, Goering MG, O'Neil-Johnson M, Zeng L (2002) High-throughput method for the production and analysis of large natural product libraries for drug discovery. *Anal Chem* 74: 3963–3971
21 Cragg GM, Newman DJ (2005) Plants as a source of anti-cancer agents. *J Ethnoparmacol* 100: 72–79

22 Newmann DJ, Cragg GM, Snader KM (2003) Natural products as sources of new drugs over the period 1981–2002. *J Nat Prod* 66: 1022–1037
23 Butler MS (2005) Natural products to drugs: Natural product derived compounds in clinical trials. *Nat Prod Rep* 22: 162–195
24 Bennouna J, Campone M, Delord JP, Pinel MC (2005) Vinflunine: A novel antitubulin agent in solid malignancies. *Expert Opin Inv Drug* 14: 1259–1267
25 McIntyre JA, Castaner J (2004) Vinflunine: Antimitotic vinca alkaloid. *Drug Future* 29: 574–580
26 Wani MC, Taylor HL, Wall ME, Coggin P, McPhail AT (1971) Plant antitumor agents: VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J Am Chem Soc* 93: 2325–2327
27 Horwitz SB, Cohen D, Rao S, Ringel I, Shen HJ, Yang CP (1993) Taxol; mechanisms of action and resistance. *J Natl Cancer I Monogr* 15: 55–61
28 Barboni L, Ballini R, Giarlo G, Appendino G, Fontana G, Bombardelli E (2005) Synthesis and biological evaluation of methoxylated analogs of the newer generation taxoids IDN5109 and IDN5390. *Bioorg Med Chem* Lett 15: 5182–5186
29 Beckers T, Mahboobi S (2003) Natural, semisynthetic and synthetic microtubule inhibitors for cancer therapy. *Drug Future* 28: 767–785
30 Dubois J, Guénard D, Guéritte F (2003) Recent development in antitumour taxoids. *Expert Opin Ther Patents* 13: 1809–1823
31 Ojima I, Geney R, Ungureanu IM, Li D (2002) Medicinal chemistry and chemical biology of new generation taxcane antitumor agents. *Life* 53: 269–274
32 Geney R, Chen J, Ojima I (2005) Recent advances in the new generation taxane anticancer agents. *Med Chem* 1: 125–139
33 Petrangolini G, Cassinelli G, Pratesi G, Tortoreto M, Favini E, Supino R, Lanzi C, Belluco S, Zunino F (2004) Antitumour and antiangiogenic effects of IDN 5390, a novel C-seco taxane, in a paclitaxel-resistant human ovarian tumour xenograft. *Brit J Cancer* 90: 1464–1468
34 Pratesi G, Laccabue D, Lanzi C, Cassinelli G, Supino R, Zucchetti M, Frapolli R, M D'Incalci E, Bombardelli E, Morazzoni P et al (2003) IDN 5390: An oral taxane candidate for protracted treatment schedules. *Brit J Cancer* 88: 965–972
35 Jordan MA, Ojima I, Rosas F, Distefano M, Wilson L, Scambia G, Ferlini C (2002) Effects of novel taxanes SB-T-2113 and IDN5109 on tubulin polymerisation and mitosis. *Chem Biol* 9: 93–101
36 Bradley MO, Swindell CS, Anthony FH, Witman PA, Devanesan P, Webb NL, Baker SD, Wolff AC, Donehower RC (2001) Tumor targeting by conjugation of DHA to paclitaxel. *J Contr Rel* 74: 233–236
37 Kuznetsova L, Chen J, Sun L, Wu X, Pepe A, Veith JM, Pera P, Bernacki R, Ojima I (2006) Syntheses and evaluation of novel fatty acid-second-generation taxoid conjugates as promising anticancer agents. *Bioorg Med Chem* Lett 16: 974–977
38 Stierle, A, Strobel G, Stierle D (1993) Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. *Science* 260: 214–216
39 Auclair C (1987) Multimodal action of antitumor agents on DNA: The ellipticine series. *Arch Biochem Biophys* 259: 1–14
40 Lee K-H, Xiao Z (2003) Lignan in treatment of cancer and other diseases. *Phytochem Rev* 2: 341–362
41 Kluza J, Mazinghien R, Irwin H, Hartley JA, Bailly C (2006) Relation between DNA

strand breakage and apopoptic progression upon treatment of HL-60 leukemia cells with tafluposide or etoposide. *Anti-Cancer Drug* 17: 155–164

42 Srivastava V, Negi AS, Kumar JK, Gupta MM, Khanuja SPS (2005) Plant-based anticancer molecules: A chemical and biological profile of some important leads. *Bioorg Med Chem* 13: 5891–5908

43 Hsiang YH, Hertzberg R, Hecht S, Liu LF (1985) Camptothecin induces protein-linked DNA breaks *via* mammalian DNA topoisomerase I. *J Biol Chem* 260: 14873–14878

44 Demarqua D, Huchet M, Coulomb H, Lesueur-Ginot L, Lavergne O, Camara J, Kasprzyk PG, Prevost G, Bigg DCH (2004) BN80927: A novel homocamptothecin that inhibits proliferation of human tumor cells *in vitro* and *in vivo*. *Cancer Res* 64: 4942–4949

45 Wink M, Alfermann AW, Franke R, Wetterauer B, Distl M, Windhoevel J, Krohn O, Fuss E, Garden H, Mohagheghzadeh A et al (2005) Sustainable bioproduction of phytochemicals by plant *in vitro* cultures: Anticancer agents. *Plant Genet Resour* 3: 90–100

46 Asano T, Watase I, Sudo H, Kitajima M, Takayama H, Aimi N, Yamazaki M, Saito K (2004) Camptothecin production by *in vitro* cultures of *Ophiorrhiza liukiuensis* and *O. kuroiwai*. *Plant Biotechnol* 21: 275–281

47 Puri SC, Verma V, Amna T, Qazi GN, Spiteller M (2005) An endophytic fungus from *Nothapodytes foetida* that produces camptothecin. *J Nat Prod* 68: 1717–1719

48 Hsieh HP, Liou JP, Mahindroo N (2005) Pharmaceutical design of antimitotic agents based on combrestratins. *Curr Pharm Design* 11: 1655–1677

49 Young SL, Chaplin DJ (2004) Combretastatin A4 phosphate: Background and current clinical status. *Exp Opin Investig Drugs* 13: 1171–1182

50 Tozer GM, Kanthou C, Baguley BC (2005) Disrupting tumour blood vessels. *Nat Rev Canc* 5: 423–435

51 Tron GC, Pirali T, Sorba G, Pagliai F, Busacca S, Genazzani AA (2006). Medicinal chemistry of combretastatin A4: Present and future directions. *J Med Chem* 49: 3033–3044

52 Eichler W, Yafai Y, Wiedermann P, Fengler D (2006) Antineovascular agents in the treatment of eye diseases. *Curr Pharm Design* 12: 2645–2660

53 Hurry JB (1930) *The woad plant and its dye*. Oxford University Press, London

54 Hoessel R, Leclerc S, Endicott JA, Nobel MEM, Lawrie A, Tunnah P, Leost M, Damiens E, Marie D, Marko D et al (1999) Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. *Nat Cell Biol* 1: 60–67

55 Jautelat R, Brumby T, Schäfer M, Briem H, Eisenbrand G, Schwahn S, Krüger M, Lücking U, Prien O, Siemeister G (2005) From the insoluble dye indirubin towards highly active soluble CDK-2 inhibitors. *ChemBioChem* 6: 531–540

56 Merz KH, Schwahn S, Hippe F, Mühlbeyer S, Jakobs S, Eisenbrand G (2004) Novel indirubin derivatives, promising anti-tumor agents inhibiting cyclin-dependent kinases. *Int J Clin Pharm Th* 42: 656–658

57 Sethi G, Ahn KS, Sandur SK, Lin X, Chaturvedi MM, Aggarwal BB (2006) Indirubin enhances tumor necrosis factor-induced apoptosis through modulation of nuclear factor-kB signaling pathways. *J Biol Chem* 281: 23425–23435

58 Polychronopoulos P, Magiatis P, Skaltsounis AL, Myrianthopoulos V, Mikros E, Tarricone A, Musacchio A, Roe SM, Pearl L, Leost M et al (2004) Structural basis for the synthesis of indirubins as potent and selective inhibitors of glycogen synthase kinase-3 and cyclin-dependent kinases. *J Med Chem* 47: 935–946

59 Von Tappeiner H, Jesionek A (1903) Therapeutische Versuche mit fluorescierenden Stoffen. *Münch Med Wochenschr* 47: 2042–2044

60 Agostinis P, Vantieghem A, Merlevede W, de Witte PAM (2002) Hypericin in cancer treatment: More light on the way. *Int J Biochem Cell Biol* 34: 221–241
61 Ion RM (2000) Porphyrin for tumor destruction in photodynamic therapy. *Curr Top Biophys* 24: 21–34
62 Vincente MGH (2001) Porphyrin-based sensitizers in the detection and treatment of cancer: Recent progress. *Curr Med Chem–Anti-Cancer Agents* 1: 175–194
63 Berg K, Selbo PK, Weyergang A, Dietze A, Prasmickaite L, Bonsted A, Enesaeter BØ, Angell-Petersen E, Warloe T, Frandsen N et al (2005) Porphyrin-related photosensitizers for cancer imaging and therapeutic applications. *J Microsc* 218: 133–147
64 Pandey S, Suresh K, Dubey S, Zheng X, Chen Y, Butt CA, Morgan J, Ciesielski M, Sajjad M, Nabi HA et al (2006) PDT efficiency of phephorbide a analogs upon conjugation with cyclo(RGDfK). Abstract of Papers, 232nd ACS National Meeting, San Francisco CA, USA, 10–14 September 2006
65 Kiesslich T, Krammer B, Plaetzer K (2006) Cellular mechanisms and prospective applications of hypericin in photodynamic therapy. *Curr Med Chem* 13: 2189–2204
66 Hampson P, Chahal H, Khanim F, Hayden R, Mulder A, Assi LK, Bunce CM, Lord JM (2005) PEP005, a selective small-molecule activator of protein kinase C, has potent antileukemic activity mediated *via* the delta isoform of PKC. *Blood* 106: 1362–1368
67 Hampson P, Wang K, Lord JM (2005) PEP-005. *Drug Future* 30: 1003–1005
68 Wang X, Matta R, Shen G, Nelin LD, Liu Y (2006) Mechanism of triptolide-induced apoptosis: Effect on caspase activation and Bid cleavage and essentiality of the hydroxyl group of triptolide. *J Mol Med* 84: 405–415
69 Itokawa H, Wang X, Lee KH (2005) Homoharringtonine and related compounds. *Anti-Canc Agents Nat Prod* 47–70
70 Cichewicz RH, Kouzi SA (2004) Chemistry, biological activity, and chemotherapeutical potential of betulinic acid for the prevention and treatment of cancer and HIV infection. *Med Res Rev* 24: 90–114
71 Patocka J (2003) Biologically active pentacyclic triterpenes and their current medicine signification. *J Appl Biomed* 1: 7–12
72 Fresco P, Borges F, Diniz C, Marques MPM (2006) New insights on the anticancer properties of dietary polyphenols. *Med Res Rev* 26: 747–766
73 Shukla Y, Pal SK (2004) Dietary cancer chemoprevention: an overview. *Int J Hum Genet* 5: 265–276
74 Harris DM, Go VLW (2006) How dietary components protect from cancer. In: AB Awad, PG Bradford (eds.) In: *Nutrition and cancer prevention*, CRC Press, Boca Raton, Florida, 27–58
75 Aggarwal BB, Shishodia S (2006) Molecular targets of dietary agents for prevention and therapy of cancer. *Biochem Pharmacol* 71: 1397–1421
76 D'Incalci M, Steward WP, Gescher AJ (2005) Use of chemopreventive phytochemicals as antineoplastic agents. *Lancet Oncol* 6: 899–904
77 Singh S, Khar A (2006) Biological effect of curcumin and its role in cancer chemoprevention and therapy. *Anti-Cancer Agents Med Chem* 6: 259–270
78 Maheshwari RK, Singh AK, Gaddipati J, Srimal RC (2006) Multiple biological activities of curcumin: A short review. *Life Sci* 78: 2081–2087
79 Signorelli P, Ghidoni R (2005) Resveratrol as an anticancer nutrient: Molecular basis, open questions and promises. *J Nutr Biochem* 16: 449–466
80 Jeffery EH, Jarrell J (2001) Cruciferous vegetables and cancer prevention. In: REC Wild-

man (ed.) In: *Handbook of nutraceuticals and functional foods*. CRC Press, Boca Raton, Florida, 169–191

81 Park EJ, Pezzuto JM (2002) Botanicals in cancer chemoprevention. *Cancer Metast Rev* 21: 231–255

82 Greenwald P, Milner JA, Anderson DA, McDonald SS (2002) Micronutrients in cancer chemoprevention. *Cancer Metast Rev* 21: 217–230

83 Greenwald P (2004) Clinical trials in cancer prevention: Current results and perspectives for the future. *J Nutr* 134: 3507S–3512S

84 Bemis DL, Katz AE, Buttyan R (2006) Clinical trials of natural products as chemopreventive agents for prostate cancer *Expert Opin Inv Drug* 15: 1191–1200

85 Wright CW (2005) Plant-derived antimalarial agents: New leads and challenges. *Phytochem Rev* 4: 55–61

86 Trape JF, Pison G, Spiegel A, Enel C, Rogier C (2002) Combating malaria in Africa. *Trends Parasitol* 18: 224–230

87 Bray PG, Ward SA, O'Neil PM (2005) Quinolines and artemisinin: Chemistry, biology and history. *Curr Top Microbiol* 295: 3–38

88 Wright CW (2005) Traditional antimalarials and the development of novel antimalarial drugs. *J Ethnoparmacol* 100: 67–71

89 Robert A, Benoit-Vical F, Meunier B (2005) The key role of heme to trigger the antimalarial activity of trioxanes. *Coordin Chem Rev* 249: 1927–1936

90 Posner GH, O'Neil PM (2004) Knowledge of the proposed chemical mechanism of action and cytochrome P450 metabolism of antimalarial trioxane like artemisinin allows rational design of new antimalarial peroxides. *Acc Chem Res* 37: 397–404

91 Haynes RK, Fugmann B, Stetter J, Rieckmann K, Heilmann H-D, Chan H-W, Cheung M-K, Lam W-L, Wong H-N, Croft SL et al (2006) Artemisone – a highly active antimalarial drug of the artemisinin class. *Angew Chem Int Ed* 45: 2082–2088

92 Hofheinz W, Burgin H, Gocke E, Jaquet C, Masciadri R, Schmid G, Stohler H, Urwyler H (1994) Ro 42-1611 (arteflene), a new effective antimalarial: Chemical structure and biological activity. *Trop Med Parasitol* 45: 261–265

93 Ashley EA, White NJ (2005) Artemisinin-based combinations. *Curr Opin Infect Dis* 18: 531–536

94 Haynes RK (2006) From artemisinin to new antimalarials: Biosynthesis, extraction, old and new derivatives, stereochemistry and medicinal chemistry requirements. *Curr Top Med Chem* 6: 509–537

95 Singh IP, Bharate SB, Bhutani KK (2005) Anti-HIV natural products. *Curr Sci* 89: 269–290

96 Joshi SP (2002) Plant products as anti-HIV agents. *J Med Arom Plant Sci* 24: 1006–1023

97 Meadows DC, Gervay-Hague J (2006) Current developments in HIV chemotherapy. *ChemMedChem* 1: 16–29

98 Cragg GM, Newman DJ (2003) Plants as a source of anti-cancer and anti-HIV agents. *Ann Appl Biol* 143: 127–133

99 Kashman Y, Gustafson KR, Fuller RW, Cardellina II JH, McMahon JB, Buckheit RW, Hughes SH, Cragg GM, Boyd MR (1992) The calanolides, a novel HIV-inhibitory class of coumarin derivatives from the tropical rainforest tree, *Calophyllum lanigerum*. *J Med Chem* 35: 2735–2743

100 Xu Z-Q, Barrow WW, Suling WJ, Westbrook L, Barrow E, Lin Y-M, Flavin MT (2004)

Anti-HIV natural product (+)-calanolide A is active against both drug-susceptible and drug-resistant strains of *Mycobacterium tuberculosis*. *Bioorg Med Chem*. 12: 1199–1207

101 Gustafson KR, Cardellina II JH, McMahon JB, Gulakowski RJ, Ishitoya J, Szallasi Z, Lewin NE, Blumberg PM, Weislow OS, Beutler JA et al (1992) A non-promoting phorbol from the Samoan medicinal plant *Homalanthus nutans* inhibits cell killing by HIV-1. *J Med Chem* 35: 1978–1986

102 Hezareh M (2005) Prostratin as a new therapeutic agent targeting HIV viral reservoirs. *Drug News Perspect* 18: 496–500

103 Williams S, Chen LF, Kwon, H, Fenard D, Bisgrove, D, Verdin, E Greene WC (2004) Prostratin antagonizes HIV latency by activating NF-κB. *J Biol Chem* 279: 42008–42017

104 Huang L, Kashiwada Y, Cosentino LM, Fan S. Lee K-H (1994) 3′4′-Di-O-(–)-camphanoyl-(+)-cis-khellactone and related compounds: A new class of potent anti-HIV agents. *Bioorg Med Chem* Lett 4: 593–598

105 Zhang Q, Chen Y, Xia P, Xia Y, Xia Y, Yang ZY, Donglei Y, Morris-Natschke SL, Lee KH (2004) Anti-AIDS agents. Part 62: Anti-HIV activity of 2′substituted 4-methyl-3′4′-di-O-(–)-camphanoyl-(+)-cis-khellactone (4-methyl DCK) analogs. *Bioorg Med Chem Lett* 14: 5855–5877

106 Yu D, Suzuki M, Xie L, Morris-Natschke SL, Lee KH (2003) Recent progress in the development of coumarin derivatives as potent anti-HIV agents. *Med Res Rev* 23: 322–345

107 Huang L, Yuan X, Yu D, Lee KH, Chen CH (2005) Mechanism of action and resistant profile of anti-HIV-1 coumarin derivatives. *Virology* 332: 623–628

108 Fujioka T, Kashiwada Y, Kilkuskie RE, Cosentino LM, Ballas LM, Jiang JB, Janzen WP, Chen IS, Lee K-H (1994) Anti-aids agents, 11. Betulinic acid and platonic acid as anti-HIV principles from *Syzygium claviflorum*, and the anti-HIV activity of structurally related triterpenoids. *J Nat Prod* 57: 243–247

109 Whelan J (2004) Promising Phase I results against new HIV target. *Drug Discov Today* 9: 823

110 Li F, Goila-Gaur R, Salzwedel K, Kilgore NR, Reddick M, Matallana C, Castillo A, Zoumplis D, Martin DE, Orenstein JM et al (2003) PA-457: A potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *Proc Natl Acad Sci USA* 100: 13555–13560

111 Decosterd LA, Parson IC, Gustafson KR, Cardellina II JH, McMahon JB, Cragg GM, Murata Y, Pannell LK, Steiner JR, Clardy J et al (1993) Structure, absolute stereochemistry, and synthesis of conocurvone, a potent, novel HIV-inhibitory naphthoquinone trimer from *Conospermum* sp. *J Am Chem Soc* 115: 6673–6679

112 Boyd MR, Hallock YF, Cardellina II JH, Manfredi KP, Blunt JW, McMahon JB, Buckheit Jr RW, Bringmann G, Schaeffer M, Cragg GM et al (1994) Anti-HIV michellamines from *Ancistrocladus korupensis*. *J Med Chem* 37: 1740–1745

113 Sorbera, LA, Castaner J, Garcia-Capdevilla L (2005) Celgosivir. *Drug Future* 30: 545–552

114 Asano N (2003) Glycosidase inhibitors: Update and perspectives on practical use. *Glycobiology* 13: 93R–104R

115 Ammon HPT (2006) Boswellic acids in chronic inflammatory diseases. *Planta Med* 72: 1100–1116

116 Singh GB, Atal CK (1986) Pharmacology of an extract of salai guggal ex – *Boswellia serrata*, a new non-steroidal anti-inflammatory agent. *Agents Actions* 18: 407–412

117 Safayhi H, Mack T, Sabieraj J, Anazodo MI, Subramanian LR, Ammon HPT (1992)

Boswellic acids: Novel, specific nonredox inhibitors of 5-lipoxygenase. *J Pharmacol Exp Ther* 261: 1143–1146

118 Safayhi H, Sailer ER, Ammon HPT (1995) Mechanism of 5-lipoxygenase inhibition by acetyl-11-keto-β-boswellic acid. *Mol Pharmacol* 47: 1212–1216

119 Singh GB, Bani S, Singh S (1996) Toxicity and safety of boswellic acids. *Phytomedicine* 3: 87–90

120 Glaser T, Winter S, Groscurth P, Safayhi H, Sailer ER, Ammon HPT, Schabet M, Weller M (1999) Boswellic acids and malignant glioma: Induction of apoptosis but no modulation of drug sensitivity. *Brit J Cancer* 80: 756–765

121 Winking M, Sarikaya S, Rahmanian A, Jödicke A, Böker DK (2000) Boswellic acids inhibit glioma growth: A new treatment option? *J Neuro-Oncol* 46: 97–103

122 Lalithakumari K, Krishnaraju AV, Sengupta K, Subbaraju GV (2006) Safety and toxicological evaluation of a novel, standardized 3-O-acetyl-11-keto-β-boswellic acid (AKBA)-enriched *Boswellia serrata* extract (5-Loxin R). *Toxicol Mech Method* 16: 199–226

123 Ammon HPT, Wahl MA (1991) Pharmacology of *Curcuma longa*. *Planta Med* 57: 1–7

124 Sharma RA, Gescher AJ, Steward WP (2005) Curcumin: The story so far. *Eur J Cancer* 41: 1955–1968

125 Aggarwal S, Ichikawa H, Takada Y, Sandur SK, Shishodia S, Aggarwal BB (2006) Curcumine down-regulates expression of cell proliferation and anti-apoptotic and metastatic gene products through suppression of IKBα kinase and Akt activation. *Mol Pharmacol* 69: 195–206

126 Shapiro H, Bruck R (2005) Therapeutic potential of curcumin in non-alcoholic steatohepatitis. *Nutr Res Rev* 18: 212–221

127 Holt PR, Katz S, Kirshoff R (2005) Curcumin therapy in inflammatory bowel disease: A pilot study. *Digest Dis Sci* 50: 2191–2193

128 Messeguer A, Planells-Cases R, Ferrer-Montiel A (2006) Physiology and pharmacology of the vanilloid receptor. *Curr Neuropharmacol* 4: 1–15

129 Appendino G, Munoz E, Fiebich BL (2003) TRPV1 (vanilloid receptor, capsaicin receptor) agonists and antagonists. *Expert Opin Ther Patents* 13: 1825–1837

130 Calixto JB, Kassuya CAL, André E, Ferreira J (2005) Contribution of natural products to the discovery of the transient receptor potential (TRP) channels family and their functions. *Pharmacol Therapeut* 106: 179–208

131 Bley KR (2004) Recent developments in transient receptor potential vanilloid receptor 1 agonist-based therapies. *Expert Opin Inv Drug* 13: 1445–1456

132 Urban L, Campbell EA, Panesar M, Patel S, Chaudhry N, Kane S, Buchheit KH, Sandells B, James IF (2000) *In vivo* pharmacology of SDZ-249665, a novel, non-pungent capsaicin analogue. *Pain* 89: 65–74

133 Park NS, Seong CM, Jung YS, Kim WB, Kim SH (2000) DA-5018 (Capsavanil, KR-25018). *Drug Future* 25: 1131–1137

134 Appendino G, Szallasi A (1997) Euphorbium: Modern research on its active principle resiniferatoxin, revives an ancient medicine. *Life Sci* 60: 681–696

135 Walker KM, Urban L, Medhurst SJ, Patel S, Panesar M, Fox AJ, McIntyre P (2003) The VR1 antagonist capsazepine reverses mechanical hyperalgesia in models of inflammatory and neuropathic pain. *J Pharmacol Exp Ther* 304: 56–62

136 Gaoni Y, Mechoulam R (1964) Isolation, structure, and partial synthesis of an active constituent of hashish. *J Am Chem Soc* 86: 1646–1647

137 Devane WA, Dysarz FA, Johnson MR, Melvin LS, Howlett AC (1988) Determination and characterisation of a cannabinoid receptor in rat brain. *Mol Pharmacol* 34: 605–613
138 Di Marzo V, Bifulco M, De Petrocelli L (2004) The endocannabinoid system and its therapeutic exploitation. *Nat Rev Drug Discov* 3: 771–784
139 Lambert DM, Fowler CJ (2005) The endocannabinoid system: Drug targets, lead compounds, and potential therapeutic applications. *J Med Chem* 48: 5059–5087
140 Klein TW (2005) cannabinoid-based drugs as anti-inflammatory therapeutics. *Nat Rev Immunol* 5: 400–411
141 Burstein S (2001) Ajulemic acid. Antiinflammatory, oncolytic. *Drug Future* 26: 324–345
142 Mechoulam R, Sumariwall PF, Feldmann M, Gallily R (2005) Cannabinoids in models of chronic inflammatory conditions. *Phytochem Rev* 4: 11–18
143 Burstein SH, Audette CA, Breuer A, Devane WA, Colodner S, Doyle SA, Mechoulam R (1992) Synthetic nonpsychotropic cannabinoids with potent antiinflammtoary, analgesic and leucocyte antiadhesion action. *J Med Chem* 35: 3135–3141
144 Fride E, Ponde D, Breuer A, Hanus L (2005) Peripheral, but not central effects of cannabidiol derivatives: mediation by CB1 and unidentified receptors. *Neuropharmacology* 48: 1117–1129
145 San Martin R, Briones R (1999) Industrial uses and sustainable supply of *Quillaja saponaria* (Rosaceae) saponins. *Econ Bot* 53: 302–311
146 Kim YJ, Wang P, Navarro-Villalobos M, Rohde BD, Derryberry JM, Gin DY (2006) Synthesis of complex immunostimulants from *Quillaja saponaria*: Synthesis of the potent clinical immunoadjuvant QS-21Aapi. *J Am Chem Soc* 128: 11906–11915
147 Kensil CR, Patel U, Lennick M, Marciani D (1991) Separation and characterization of saponins with adjuvant activity from *Quillaja saponaria* Molina cortex. *J Immunol* 146: 431–437
148 Kensil CR (1996) Saponins as vaccine adjuvants. *Crit Rev Ther Drug* 13: 1–55
149 Soltysik S, Wu JY, Recchia J, Wheeler DA, Newmann MJ, Coughlin RT, Kensil CR (1995) Structure/function studies of QS-21 adjuvant: Assessment of triterpene aldehyde and glucuronic acid roles in adjuvant function. *Vaccine* 13: 1403–1410
150 Liu G, Anderson C, Scaltreto H, Barbon J, Kensil CR (2002) QS-21 structure/function studies: Effect of acylation on adjuvant activity. *Vaccine* 20: 2808–2815
151 Marciani DJ, Reynolds RC, Pathak AK, Finley-Woodman K, May RD (2003) Fractionation, structural studies, and immunological characterization of the semi-synthetic *Quillaja* saponins derivative GPI-0100. *Vaccine* 21: 3961–3971
152 Kersten GFA, Spiekstra A, Beuvery EC, Crommelin DJA (1991) On the structure of immune-stimulating saponin-lipid complexes (iscoms). *Biochim Biophys Acta* 1062: 165–171
153 Kersten GFA, Crommelin DJA (1995) Liposomes and ISCOMS as vaccine formulations. *Biochim Biophys Acta* 1241: 117–138
154 Bays H, Stein EA (2003) Pharmacotherapy for dyslipidemia – current therapies and future agents. *Expert Opin Pharmaco* 4: 1901–1918
155 Urizar NL, Moore DD (2003) Gugulipid: A natural cholesterol-lowering agent. *Annu Rev Nutr* 23: 303–313
156 Satayavati GV (1988) Gum guggul (*Commiphora mukul*) – the success story of an ancient insight leading to a modern discovery. *Indian J Med Res* 87: 327–335
157 Ulbricht C, Basch E, Szapary P, Hammerness P, Axentsev S, Boon H, Kroll D, Garraway

L, Vora M, Woods J (2005) Guggul for hyperlipidemia: A review by the Natural Standard Research Collaboration. *Compl Ther Med* 13: 279–290

158 Urizar NL, Liverman AB, Dodds DT, Silva FV, Ordentlich P, Yan Y, Gonzalez FJ, Heyman RA, Mangelsdorf DJ, Moore DD (2002) A natural product that lowers cholesterol as an antagonist ligand for FXR. *Science* 296: 1703–1706

159 Patel MD, Thompson PD (2006) Phytosterols and vascular disease. *Atherosclerosis* 186: 12–19

160 Sudhop T, Lütjohann D, von Bergmann K (2005) Sterol transporters: targets of natural sterols and new lipid lowering drugs. *Pharmacol Ther* 105: 333–341

161 Kritchevsky D, Chen SC (2005) Phytosterols – health benefits and potential concerns: A review. *Nutr Res* 25: 413–428

162 De Melo EB, da Silveira Gomes A, Carvalho I (2006) α- and β-glucosidase inhibitors: Chemical structure and biological activity. *Tetrahedron* 62: 10277–10302

163 Zou W (2005) C-glycosides and aza-C-glycosides as potential glycosidase and glycosyltransferase inhibitors. *Curr Top Med Chem* 5: 1363–1391

164 Asano N, Nash RJ, Molyneux RJ, Fleet GWJ (2000) Sugar-mimic glycosidase inhibitors: Natural occurrence, biological activity and prospects for therapeutic application. *Tetrahedron-Asymmetr* 11: 1645–1680

165 Butters TD, Dwek RA, Platt FM (2005) Imino sugars inhibitors for treating the lysosomal glycosphingolipidoses. *Glycobiology* 15: 43R–52R

166 Jia W, Gao W, Tang L (2003) Antidiabetic herbal drugs officially approved in China. *Phytother Res* 17: 1127–1134

167 Grover JK, Yadav S, Vats V (2002) Medicinal plants of India with antidiabetic potential. *J Ethnopharmacol* 81: 81–100

168 Yeh GY, Eisenberg DM, Kaptschuk TJ, Phillips RS (2003) Systematic review of herbs and dietary supplements for glycemic control in diabetes. *Diabetes Care* 26: 1277–1294

169 Jung M, Park M, Lee HC, Kang YH, Kang ES, Kim SK (2006) Antidiabetic agents from medicinal plants. *Curr Med Chem* 13: 1203–1218

170 Bossy-Wetzel E, Schwarzenbacher R, Lipton SA (2004) Molecular pathways to neurodegeneration. *Nat Med* 10: S2–S9

171 Huber A, Stuchbury G, Bürkle A, Burnell J, Münch G (2006) Neuroprotective therapies for Alzheimer's disease. *Curr Pharm Design* 12: 705–717

172 Francis PT, Nordberg A, Arnold SE (2005) A preclinical view of cholinesterase inhibitors in neuroprotection: Do they provide more than symptomatic benefits in Alzheimer's disease? *Trends Pharmacol Sci* 26: 104–111

173 Houghton PJ, Ren Y, Howes MJ (2006). Acetylcholinesterase inhibitors from plants and fungi. *Nat Prod Rep* 23: 181–199

174 Wang R, Yan H, Tang XC (2006) Progress in studies of huperzine A, a natural cholinesterase inhibitor from Chinese herbal medicine. *Acta Pharmacol Sin* 27: 1–26

175 Kozikowski AP, Tückmantel W (1999) Chemstry, pharmacology, and clinical efficacy of the Chinese nootropic agent huperzine A. *Acc Chem Res* 32: 641–650

176 Ma X, Gang DR (2004) The *Lycopodium* alkaloids. *Nat Prod Rep* 21: 752–772

177 Marco-Contelles J, do Carmo Carreiras M, Rodriguez C, Villaroya M, Garcia AG (2006) Synthesis and pharmacology of galanthamine. *Chem Rev* 106: 116–133

178 Mukherjee PK, Kumar V, Mal M, Houghton PJ (2007) Acetylcholinesterase inhibitors from plants. *Phytomedicine* 14: 289–300

179 Zimmermann M, Colchiaghi F, Cattabeni F, Di Luca M (2002) *Ginkgo biloba* extract:

From molecular mechanisms to the treatment of Alzheimer's disease. *Cell Mol Biol* 48: 613–623

180 Christen Y (2004) *Ginkgo biloba* and neurodegenerative disorders. *Frontiers Biosci* 9: 3091–3104

181 Gertz HJ, Kiefer M (2004) Review about *Ginkgo biloba* special extract EGb 761 (Ginkgo). *Curr Pharm Design* 10: 261–264

182 Andrieu S, Gilette S, Amouyal K, Nourashemi F, Reynish W, Ousset PJ, Albarede JL, Vellas B, Grandjean H (2003) Association of Alzheimer's disease onset with *Ginkgo biloba* and other symptomatic cognitive treatments in a population of women aged 75 years and older from the EPIDOS study. *J Gerontol A Biol Sci Med Sci* 58: 372–377

183 NG TP, Chiam PC, Lee T, Chua HC, Lim L, Kua EH (2006) Curry consumption and cognition function in the elderly. *Am J Epidemiol* 164: 898–906

184 Mancuso C, Scapagnini G, Curro D, Stella AMG, De marco C, Butterfield DA, Calabrese V (2007). Mitochondrial dysfunction, free radical generation and cellular stress response in neurodegenerative disorders. *Frontiers Biosci* 12: 1107–1123

185 Lim GP, Chu T, Yang FS, Beech W, Frautschy SA, Cole GM (2001) The curry spice curcumin reduces oxidative damage and amyloid pathology in an Alzheimer transgenic mouse. *J Neurosci* 21: 8370–8377

186 Yang FS, Lim GP, Begum AN, Ubeda OJ, Simmons MR, Ambegaokar SS, Chen PP, Kayed R, Glabe CG, Frautschy SA, Cole GM (2006) Curcumin inhibits formation of amyloid beta oligomers and fibrils, binds plaques, and reduces amyloid *in vivo*. *J Biol Chem* 280: 5892–5901

187 Langlois N, Gueritte F, Langlois Y, Potier P (1976) Application of a modification of Polonovski reaction to synthesis of vinblastine-type alkaloids. *J Am Chem Soc* 98: 7017–7024

188 Mishra P, Uniyal GC, Sharma S, Kumar S (2001) Pattern of diversity of morphological and alkaloid yield related traits among periwinkle *Catharanthus roseus* accessions collected from and around Indian subcontinent. *Genet Resour Crop Ev* 48: 273–286

189 Schmid W, Balz JP (2005) Cultivation of *Ginkgo biloba* L. on three continents. *Acta Horticult* 676: 177–180

190 Suffness M (1993) Taxol – from discovery to therapeutic use. *Annu Rep Med Chem* 28: 305–314

191 Zhong JJ (2002) Plant cell culture for production of paclitaxel and other taxanes. *J Biosci Bioeng* 94: 591–599

192 Tabata H (2006) Production of paclitaxel and the related taxanes by cell suspension cultures of Taxus species. *Curr Drug Target* 7: 453–461

193 Farina V, Brown JD (2006) Tamiflu: the supply problem. *Angew Chem Int Ed* 45: 2–7

194 Farkya S, Bisaria VS, Srivastava AK (2004) Biotechnological aspects of the production of the anticancer drug podophyllotoxin. *Appl Microbiol Biot* 65: 504–519

195 Bedir E, Tellez M, Lata H, Khan I, Cushman KE, Moraes RM (2006) Post-harvest and scale up extraction of American mayapple leaves for podophyllotoxin production. *Ind Crop Prod* 24: 3–7

196 Soeharto DD, Gyllenhaal C, Fong HHS, Xuan LT, Hiep NT, Hung NV, Bich TQ, Southavong B, Sydara K, Pezzuto JM (2004) The UIC ICBG (University of Illinois at Chicago International Cooperative Biodiversity Group) memorandum of agreement: A model of benefit-sharing arrangement in natural products drug discovery and development. *J Nat Prod* 67: 294–299

197 Iwu MM (1996) Implementing the biodiversity treaty: How to make international co-operative agreements work. *Trends Biotechnol* 14: 78–83
198 Anon. (1994) 'Bio-piracy' costs developing world 5 billion a year. *Scrip* 1980: 19
199 Rausser GC, Small AA (2000) Valuing research leads: Bioprospecting and the conservation of genetic resources. *J Polit Econ* 108: 173–206
200 Dalton R (2006) Cashing in on the rich coast. *Nature* 442: 567–569
201 Coley PD, Heller MV, Aizprua R, Arauz B, Flores N, Correa M, Gupta M, Solis PN, Ortega-Barria E, Romero LI et al (2003) Using ecological criteria to design plant collection strategies for drug discovery. *Front Ecol Environ* 1: 421–428
202 Balandrin MF, Kinghorn AD, Farnsworth NR (1993) Plant-derived natural products in drug discovery and development – an overview. *ACS Symp Series* 534: 2–12
203 Kinghorn AD (2001) Pharmacognosy in the 21st century. *J Pharm Pharmacol* 53: 135–148
204 Shu YZ (1998) Recent natural products based drug development: A pharmaceutical industry perspective. *J Nat Prod* 61: 1053–1071

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Evolutionary mechanisms underlying secondary metabolite diversity

By Holger Jenke-Kodama[1], Rolf Müller[2] and Elke Dittmann[1]

[1]Humboldt University Berlin,
Institute of Biology,
Department of Molecular Ecology,
Chausseestr. 117,
10115 Berlin, Germany
<elke.dittmann@rz.hu-berlin.de>

[2]Saarland University,Department of Pharmaceutical Biotechnology,
PO Box 151150,
66041 Saarbrücken, Germany

Abstract

The enormous chemical diversity and the broad range of biological activities of secondary metabolites raise many questions about their role in nature and the specific traits leading to their evolution. The answers to these questions will not only be of fundamental interest but may also provide lessons that could help to improve the screening protocols of pharmaceutical companies and strategies for rational secondary metabolite engineering. In this review, we try to dissect evolutionary principles leading to the emergence, distribution, diversification and selection of genes involved in secondary metabolite biosyntheses. We give an overview about recent insights into the evolution of the different types of polyketide synthases (PKS) in microorganisms and plants and highlight unique mechanisms underlying polyketide diversity. Although phylogenetic and experimental data have significantly increased our knowledge about the role and evolution of secondary metabolites in the last decades there is still much dissent about the impact of natural selection. In order to understand the evolution towards metabolic diversity we therefore need more thorough investigations of the ecological role of secondary metabolites in the future.

1 Introduction

Secondary metabolites are defined as a group of natural compounds that, in contrast to primary metabolites, are not directly involved in growth, development or reproduction of organisms. Typically, secondary metabolites belong to diverse chemical substance classes and are regarded as characteristic for individual species (chemotaxonomy). The term 'secondary' in the context of metabolism was introduced in 1891 by the German plant physiologist Kossel [1]. Facing a steadily growing conglomeration of newly identified substances produced by living organisms, Kossel suggested to concentrate the research endeavours on the 'essential' molecules, i.e., the ubiquitous ones, which he called 'primary'. Thus, this discrimination was not based on function, but rather it was purely phenomenological and resulted in an *ad hoc* definition. Although the terms 'primary metabolism' and 'secondary metabolism' are commonly used, there are a number of substances that are not easy to classify. Lipids or polysaccharides, for example, play essential roles for each organism but some compounds of these classes may provide individual characteristic features for their producing organisms [2]. The basis of the specific functions of secondary metabolites is their prevalent potent biomolecular activity that assures a specific interaction with their targets. In that way, secondary metabolites frequently act as biological warfare agents against predators, parasites and

diseases or may fulfil intrinsic physiological functions for their producing organisms [3].

But how have secondary metabolites evolved and which evolutionary mechanisms have led to their striking diversity and their species–specific occurrence? The answers to these questions are not only of a fundamental nature. They may also provide lessons that could help to improve the screening protocols of pharmaceutical companies and strategies for rational secondary metabolite engineering. In this review we attempt to give an overview about existing hypotheses about the evolution and diversification of biosynthetic enzymes involved in secondary metabolite production. Secondary metabolites include diverse 'small molecule' classes, such as alkaloids, terpenoids, glycosides or peptides. We will, however, mainly focus on the diverse types of polyketides produced by microorganisms and plants. Polyketides from bacteria and plants include some of the most important drug leads such as erythromycin, tetracycline and epothilone [4, 5]. The different types of synthases involved in polyketide production exemplify very different evolutionary strategies of organisms to generate secondary metabolite diversity. Apart from discussing an individual selection of examples we will try to draw some general conclusions about the evolutionary strategies that have led to the plethora of secondary metabolites in individual organisms.

2 Evolution of secondary metabolite biosynthesis enzymes from primary metabolism

It is generally assumed that pathways of secondary metabolism originate from precursors in the primary metabolic network. Individual secondary pathways may have started to evolve as the result of incidental duplication and subsequent mutation of one of the gene copies [6]. In many cases, ancestry of such enzymes from primary metabolism can be easily anticipated. Many enzymes from secondary metabolite pathways use universally present precursors and show sequence similarity to their primary counterparts. With the increasing amount of sequence information in databases robust phylogenetic analyses can be applied to test the evolutionary connection between the different enzyme types. Polyketide synthases (PKS), as an example, use the same type of enzymatic reactions like essential fatty

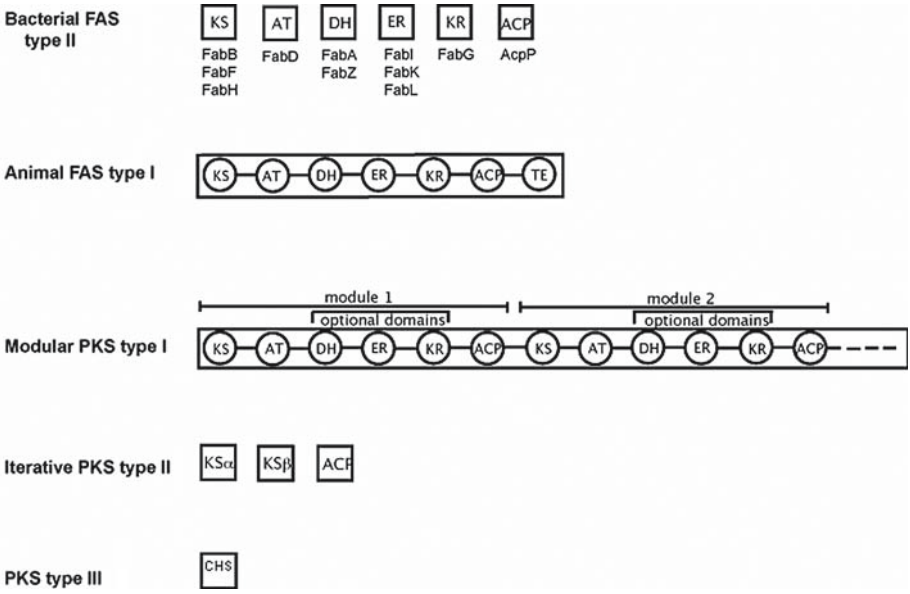


Figure 1
Schematic representation of fatty acid synthases (FAS) and polyketide synthases (PKS). Enzyme systems carrying the catalytic sites on separate proteins are classified as type II FAS and PKS, respectively. Enzyme systems possessing a multidomain architecture are classified as type I FAS and PKS, respectively. PKS III are characterised by a single domain architecture. Distinct proteins are shown as squares and domains integrated within proteins as circles, respectively. Optional domains of PKS I are indicated.
Abbreviations: KS, ketosynthase; KSα and KSβ, subtypes of KS in PKS II systems; AT, acyltransferase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; ACP, acyl carrier protein; CHS, chalcone synthase

acid synthases (FAS). The multidomain type I PKS in bacteria and fungi resemble type I FAS from animals; whereas type II PKS are similar to type II FAS from bacteria and plants that carry their catalytic sites on distinct proteins ([7], Fig. 1). The PKS superfamily of enzymes further includes the chalcone and stilbene synthases that were designated as type III PKS (Fig. 1). These homodimeric single-domain proteins represent the simplest type of PKS and were originally considered as plant specific; however more recently the presence of PKS III encoding enzymes was repeatedly reported for different groups of bacteria [8–10]. A first step towards an understanding of the evolutionary history of the different types of PKS is to evaluate

their relationship with FAS from primary metabolism. This question has been addressed in independent studies for PKS I, PKS II and for PKS III, respectively. An analysis of ketosynthase (KS) domains and proteins of FAS (I, II) and PKS (I, II) has revealed that PKS and FAS have passed through a long joint evolution process. The phylogenetic analyses suggest that first the more ancient type II FAS and PKS evolved from their ancestor proteins before the first 'fusion-type' multidomain FAS emerged in some actinobacteria and fungi. These multidomain FAS differ in their domain architecture from PKS I and from FAS I in animals. Presumably, a second multidomain enzyme evolved independent from the fungal type FAS and served as ancestor for the iterative and modular types of PKS I in bacteria and fungi and for the FAS of animals (Fig. 2). These striking correlations also explain the astonishing fact that animal FAS are much more similar to bacterial and fungal PKS I than to FAS from fungi. FAS and PKS therefore represent an example where secondary metabolite biosynthesis enzymes not only descend from primary metabolism enzymes but have never followed a separate path in evolution. Such a close connection of primary and secondary metabolism is not always common. An example that might be more characteristic is the evolution of PKS III. The most similar type of enzyme in the database is the FabH type ketosynthase that belongs to type II FAS systems. The phylogenetic analysis of FabH enzymes and PKS III from bacteria and plants has suggested that eubacterial PKS III derive from a eubacterial FabH ancestor protein. However, the evolutionary distance between the two enzyme types is remarkable. It can be speculated that the development of a PKS III type enzyme from a FabH type enzyme occurred only once during evolution. Following this scenario, eubacterial PKS III can be predicted to be the ancestors of the diverse plant chalcone and stilbene synthases and that type III PKS are not an original invention of plants (Fig. 2). The examples of the diverse types of PKS may represent the different routes that have led to the emergence of secondary metabolite biosyntheses and their products. It may not always be easy to recognise the primary progenitor proteins for each secondary metabolite pathway and in some cases missing evolutionary links could circumvent the uncovering of specific relationships. Nevertheless, the first step towards the development of secondary metabolism follows common principles. It seems much more difficult to understand the striking distribution of the secondary metabolites in different organisms.

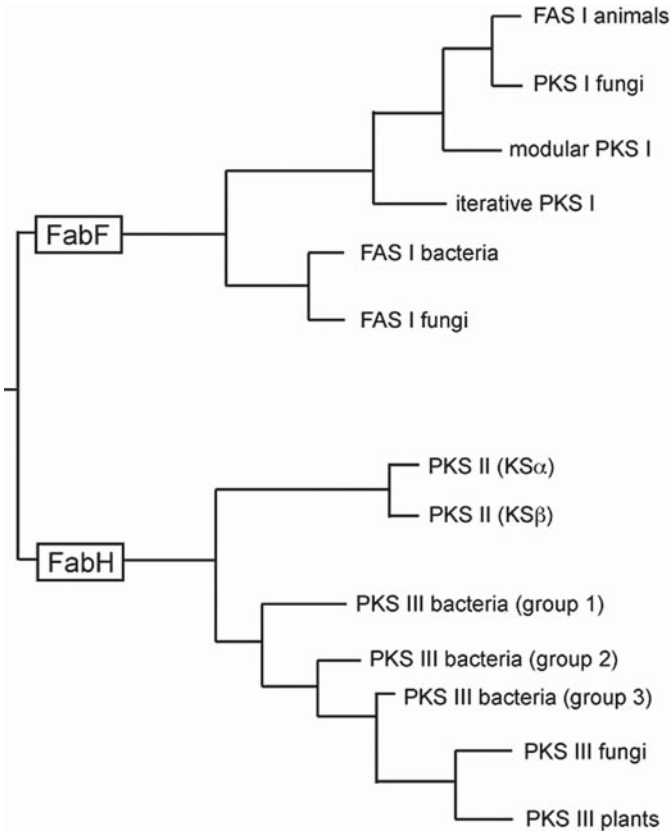


Figure 2
Schematic representation of the phylogenetic relationships between the different types of FAS and PKS based on amino acid sequences of KS domains. KS domains of diverse types of PKS in bacteria and plants descend from FabF and FabH type KS that are part of type II FAS systems in eubacteria. The tree combines the results from Jenke-Kodama et al. (2005) and Gross et al. (2006) and unpublished data. The branch lengths are arbitrary. Clade credibility values are given in the original publications. Schematic representations of PKS types are shown in Figure 1.

3 Factors influencing the distribution of biosynthetic enzymes among species

Compared to the primary metabolic network, secondary metabolites and their biosynthetic pathways show many more variations with respect to the organisms they occur in and show an often patchy distribution even

among closely related strains of the same species. This irregular distribution has most likely resulted from a sum of evolutionary processes and even a thorough analysis is prone to failures as the different factors may interfere. At first, enzymes involved in the biosynthesis of secondary metabolites underlie general speciation processes. Enzymes have evolved independently in different species and their evolutionary distance usually reflects the time that has passed since the last common ancestor of those species existed. However, enzymes involved in secondary metabolite pathways may be subject to different diversification and selection processes in the species (see following chapters) and are not necessarily as well conserved as most enzymes of the primary metabolism. A second factor that may influence the distribution of these enzymes is the so-called 'birth and death' evolution that is often characteristic of multigene families [11]. Many enzymes of secondary metabolite pathways are derived from duplication events and have established such enzyme families. According to the 'birth and death' model some of the duplicated genes are maintained for a long time, whereas others are deleted or become non-functional through deleterious mutations [11]. This model applies to families such as the immunoglobulins, the major histocompatibility complex (MHC) and disease resistance genes [11]. The 'birth and death' theory also provides a possible explanation for a species–specific loss (or retention) of secondary metabolite enzymes. Indications for an involvement of 'birth and death' evolution on the distribution of PKS I genes were found in a recent analysis of the completely sequenced genome of *Streptomyces avermitilis*. In addition to three giant modular PKS gene clusters encoding pathways for avermectin, oligomycin and an uncharacterised polyene, the genome contains a number of PKS genes that are probably nonfunctional due to substantial mutations and appear to be fragmented remnants of once functional clusters (ORFs) [12]. These ORFs match the characteristics of typical bacterial pseudogenes like truncation or disruption over a sizeable portion [13]. For instance, the protein encoded by ORF pks8-2 consists of a single KS domain, which is flanked at the C-terminus by remnants of a KS-AT interdomain region. The KS domain itself is shortened dramatically in length from typically around 420 residues to only 265 residues. This is due to a number of large deletions that are distributed over the whole domain with some of them affecting strictly conserved regions. Similar observations were made for the pks1 and pks9 ORFs in the *S. avermitilis* genome.

A third factor that has an impact on the distribution of secondary metabolites, in particular in bacteria, is the horizontal transfer of the responsive genes. Whereas horizontal gene transfer (HGT) between bacteria and higher eukaryotes can be considered as a rare event in evolution it generally plays a major role for the distribution of certain gene families in bacteria [14].

One aspect that is limiting the success of HGT events of secondary metabolite gene clusters is the toxicity of the products for the producing organisms. Many antibiotics affect prokaryotic DNA, protein translation or cell wall synthesis [15]. The producing bacteria have to cope with these bioactivities either by sequestering the metabolites or by modifying them into noneffective variants. *Streptomyces verticillus*, the producer of the bleomycin family of antitumor antibiotics, has evolved at least two independent strategies to avoid suicide. The BlmA protein shows a very high affinity to the glycopeptides thus keeping the metabolite apart from the target DNA [16]. A second protein, BlmB acts as an N-acetyltransferase capable of inactivating bleomycins and related antibiotics [17]. In many other cases, sequestering of antibiotics is achieved by specific efflux pumps. Examples for polyketide antibiotics that are accompanied by this type of resistance strategy include leinamycins of *Streptomyces atroolivaceus* [18] and pyoluteorins of *Pseudomonas fluorescens* [19]. Yet another resistance strategy is used by the pactamycin producer *Streptomyces pactamus* and the tylosin producer *Streptomyces fradiae*. Here, resistance against the antibiotics involves methylation of the 16S RNA and 23S RNA targets, respectively [20, 21]. Taking the individuality of these resistance strategies into advance it is not surprising that genes conferring resistance against specific antibiotics are frequently clustered with the corresponding biosynthetic genes. Although phylogenetic evidence is missing it can be speculated that biosynthetic genes and resistance genes went through a concerted evolution process. A successful HGT event of antibiotic gene clusters would accordingly have to include the transfer of resistance genes. As both types of genes often form part of a genomic island, an important precondition for a joint HGT event is fulfilled. Two examples of secondary metabolite gene clusters that are located on mobile genomic elements are the yersiniabactin biosynthetic gene cluster that is part of a high pathogenicity island (HPI) in E.coli [22] and the coronatin biosynthesis cluster that is frequently located on plasmids in phytopathogenic *Pseudomonas* strains [23]. However, in most cases

direct evidence for the transfer of genomic islands is missing. Nevertheless it can be speculated that HGT is one of the major reasons for the patchy distribution of secondary metabolite clusters in bacteria.

How difficult it is to assess the influence of the different factors on the distribution of secondary metabolite pathways can be exemplified for PKS and their genes. In bacteria, PKS I and II genes are confined to specific groups or genera, whereas most bacteria seem to have lost the ability to produce polyketides or have never evolved this feature. Partly, this was apparent from screening surveys that have revealed actinobacteria, myxobacteria, pseudomonads, bacilli and cyanobacteria as predominant producers of such compounds. In the postgenomic era this observation is now largely being confirmed by the absence of PKS genes from most bacterial genomes present in the database [7]. The impact of HGT for the distribution of PKS I seems to differ between bacterial groups. Whereas in actinobacteria and cyanobacteria most PKS I genes may have evolved from common ancestors, HGT may have played a major role for the evolution of PKS I in proteobacteria. In the first case, individual differences between strains could result from speciation, diversification and 'birth and death' evolution. Unlike in actinobacteria and cyanobacteria, PKS genes in γ-proteobacteria are frequently located on mobile genomic islands that differ from the 'core' genome in their GC content [24]. There are, however, other bacteria (in particular myxobacteria) where the evolutionary history of PKS remains obscure. Whereas the distribution of the sequences in phylogenetic trees clearly indicate HGT events (followed by frequent duplication events in myxobacteria) no further support for this theory comes from the analysis of the GC content of these genes [7]. A GC content analysis would also fail to detect HGT between closely related species. The impact of HGT on PKS evolution can therefore not be assessed without an experimental verification.

In contrast to bacteria, for the evolution of secondary metabolite enzymes in plants usually a vertical descent can be anticipated. Nevertheless, an irregular distribution of secondary metabolite genes is also characteristic for these higher eukaryotes. As explained above, these patterns could be explained by 'birth and death' evolution processes and by a strain- or genus specific diversification of the enzymes. Such an evolutionary scenario can thus also be presumed for the family of chalcone and stilbene synthases (PKS III) in plants. A phylogenetic comparison of PKS III

sequences from bacteria, fungi and plants clearly rules out the possibility of a HGT between these large organismic groups [8].

In order to explain the inconsistent distribution pattern of secondary metabolites therefore only rarely HGT must be invoked. It seems more important to understand the mechanisms for the diversification of secondary metabolite biosynthesis enzymes and the underlying selection pressure.

4 Mechanisms leading to the diversification of enzymes and their products

In order to interpret the impact of diversification processes on the evolution of secondary metabolites, an understanding of the enzymatic basis of their diversity and of the capability of the respective biosynthetic enzymes to produce high amounts of different compounds is needed. Nature has found different solutions to fulfil this task. There are special biosynthetic systems that are found exclusively in the context of secondary metabolism and, on the other hand, there are typical enzyme pathways as found in primary metabolism. Although their general organisation is very different from each other, they seem to serve the same aim, namely to provide the potential to produce many different structures by means of rather limited genetic resources.

A unique diversification principle is found in modular PKS I. Although PKS I is highly similar to FAS the modular organisation facilitates the production of manifold compounds. The basis of the product diversity is the construction kit-like principle underlying a typical PKS I assembly line. Each module catalyses one cycle of elongation thereby prolonging the polyketide chain by one building block. It depends on the module type which kind of building block is incorporated and to which reduction level it is modified after formation of the new C-C bond. A similar principle can be anticipated for modular non-ribosomal peptide synthetases (NRPS). The existence of modules with different substrate specificities and variable domain compositions in conjunction with the possibility of combining them in a permutational manner results in a vast diversity of polyketide or peptide structures. One can calculate that a PKS system of six elongation modules is theoretically capable to produce more than

100,000 different structures [25]. Indeed, experimental approaches were used to combine different sets of modules in the laboratory for many years [26, 18] before the growing number of complete bacterial genome sequences provided insight into nature's way of handling biosynthetic modularity. In a recent study the actinobacterium *Streptomyces avermitilis* served as the model organism as its genome encodes the largest number of PKS modules of all currently available bacterial genomes and most of these modules can be assigned to the biosynthesis of already characterised polyketide compounds. Using phylogenetic reconstruction methods it could be demonstrated that about two thirds of the modules cannot be explained as the outcome of a mere duplication process. Rather, they show extensive exchange of domains ([8]; Fig. 3). Furthermore, the putative recombination sites were mainly identified within interdomain regions and the recombination-based reprogramming of modules was confined to domains which account for the structural diversity of the polyketide backbone. Modularisation of PKS I and reprogramming of modules by recombination can likely occur in a relatively short period of time. Support for this theory comes from an analysis of the mycolactone biosynthetic gene cluster that is implicated in the Buruli ulcer desease [27]. The KS domains of this cluster that catalise the C-C bond forming steps show 97% identity among each other and most probably result from recent duplication events. The evolution of this toxin has therefore with all probability occurred after speciation in the specific host-pathogen environment. Secondary metabolites that are highly specific for their ecological niches could have evolved by fast diversification and selection processes. Taken together, modular PKSs can be interpreted as molecular machines invented by nature to create chemical diversity in a 'gene-saving' way. To produce relatively small molecules by means of giant enzyme complexes only makes sense because of the inbuilt flexibility of the assembly line. Taking this into account, it also becomes clear why these biosynthesis systems occur only in the production of secondary metabolites. They produce a certain compound and concurrently bear the potential to change it easily by recombination processes without the necessity to mutate amino acid residues in the enzyme's active centre or in other functionally important regions. One can speculate that the same principles and conclusions apply to the likewise modularly organised NRPSs, albeit an analogous study analysing these systems is still missing.

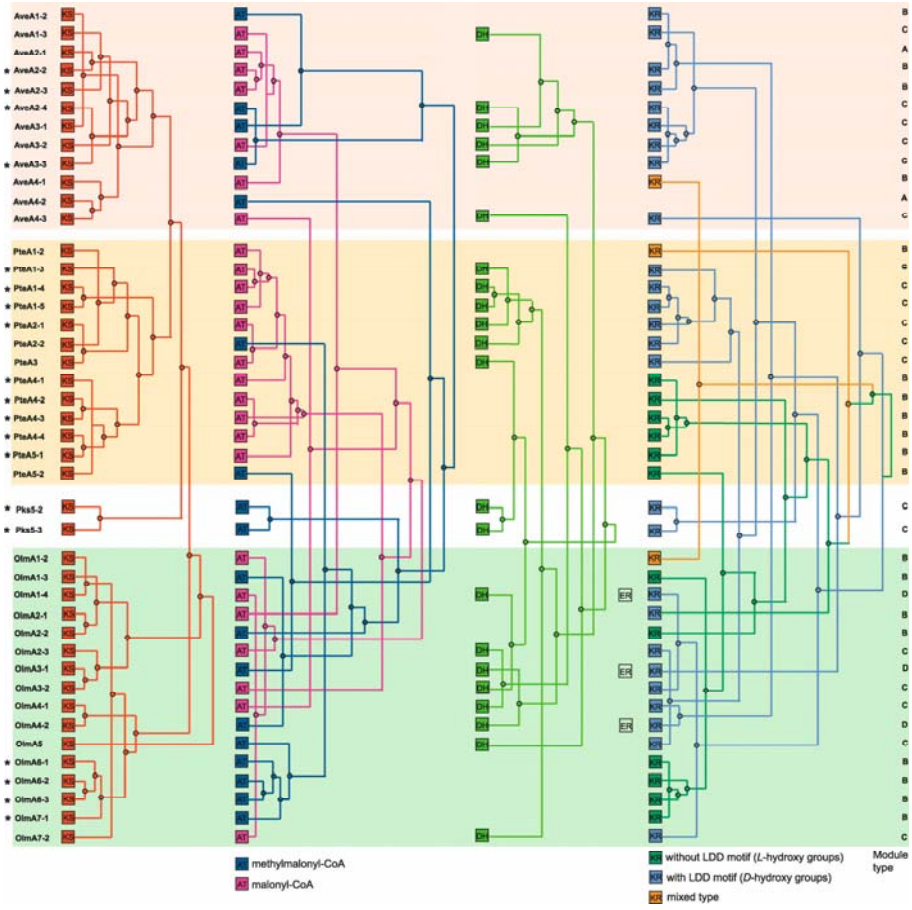


Figure 3
The modular architecture of multimodular PKS I facilitates a frequent exchange of gene segments by recombination. Shown here are the phylogenetic relationships of different domain types that are part of multimodular PKS I encoded by the actinobacterium *Streptomyces avermitilis*. The strain encodes three major PKS I complexes that can be assigned to the synthesis of the avermectin (Ave), a polyene macrolide (Pte) and oligomycin (Olm). In a first step, all three PKS gene cluster have presumably evolved by duplication from a single pathway specific ancestor module. This evolutionary history is reflected by KS domain sequences that go back to a cluster specific ancestor in all Ave, Pte and Olm PKS modules, respectively. In a second step, recombination has occurred between different module types leading to an exchange, loss or gain of AT, DH and KR domains. The different subtypes of AT and KR domains are shown in different colours. Only those modules that show similar phylogenetic distances of all their domains to the respective domains of another module can be explained by whole module duplications (indicated by asterics). 65% of the modules show phylogenetic incongruities, indicating recombination events following duplication. These data suggest that recombination is a major driving force for the diversification of multimodular PKS I. Figure taken from [12].

Besides these special secondary metabolism systems, there are many classical pathway-type reactions for the biosynthesis of natural products as in the case of terpenoids, phenylpropanoid-derived compounds and alkaloids. Nevertheless, enzymes involved in the respective pathways can be expected to have some specific features compared to their 'primary counterparts'. It has been discussed that biosynthetic enzymes involved in secondary metabolite production exhibit a broader substrate specificity and that they are typically organised in branched pathways and matrix pathways [2]. These attributes were predicted as prerequisites for the generation of a large variety of structures with a limited number of enzymes. In order to support their so-called Screening Hypothesis, Firn and Jones searched the literature for studies that provided evidence for their predictions and compiled dozens of impressive examples [28]. For instance, it could be demonstrated that in Scotch spearmint a mutation in a single enzyme resulted in the production of a series of new p-menthane monoterpenes. The mutation caused an oxygenation at C3 instead of the typical C6 position and gave rise not only to one new compound but to a whole series as the enzymes following the oxygenation step are capable of converting the new substrate due to their capability to process a variety of substrates [29]. An ostensive case of a grid-like pathway organisation was identified for astaxanthin biosynthesis in *Agrobacterium aurantiacum* where β-carotene can be converted into astaxanthin *via* eight different intermediates by the action of only two enzymes, a hydroxylase and a ketolase [30]. For PKS III of plants, it was shown that their evolution follows the gene duplication model with point mutations in the active centre leading to new products by divergent cyclisation mechanisms. In that way, stilbene synthases have evolved several times independently from chalcone synthases [31–33]. Taken together, different lines of evidence suggest that enzymes involved in secondary metabolism are to some extent organised or equipped to be capable of producing a maximum of chemical variety. However, not every single step of a secondary metabolism pathway is necessarily arranged in this way. Many enzymes involved in the production of natural products catalyse very specific reactions, e.g., in the biosynthesis of the phenylpropanoids and alkaloids [34]. The evolution of chemical diversity and the development of a specific bioactivity are only two aspects of one concept.

5 Hypotheses about functional roles of secondary metabolites

Natural selection of organisms will favour advantageous traits that either increase general fitness or defence capabilities. Selection is therefore directly connected to the functional roles of secondary metabolites. The enormous diversity of secondary metabolites from plants and microorganisms has provoked many speculations and debates about their biological significance. A widespread hypothesis, particularly in the older literature, is the assumption that secondary metabolites are waste or detoxification products of organisms [35]. Another prevalent opinion is that secondary metabolites primarily represent an arsenal of natural chemical warfare [36]. Opponents of this theory have suggested that secondary metabolites form a pool of nonfunctional variety and function as 'inventive' metabolites [37]. Yet other authors favoured the opinion that these metabolites are remnants of former functional metabolites [38]. These historical examples do not present a complete enumeration, but they give an overview about the most interesting and influential suggestions in the last century. Some of them are obviously not suited to explain the whole range of diversity; others are clearly disproved by experimental data.

The view on secondary metabolites as simple waste products of primary metabolism reflects the lack of knowledge about biological functions of secondary metabolites in the first half of the last century. A second problem of the waste product theory is that it cannot provide an explanation for the high number and diversity of secondary metabolites. It is also in contradiction with the fact that plants and microbes, the main producers of secondary metabolites, do not need such a detoxification system. Plants are able to control their metabolic inputs and microbes would be expected to export waste compounds to their surroundings instead of developing a myriad of energy consuming reactions to transform them. If detoxification pathways were essential earlier in evolution, their components would have evolved a long time ago and would most likely have resulted in more or less conserved reaction sequences in different groups of organisms. Furthermore, there are many examples of compounds for which it is a well-established fact that they play an important role in ecological interactions [39, 40]. We cannot exclude that *some* compounds may serve or have served a role as waste or detoxification products. However, this idea is far from giving a general assertion.

In the last few years another perception of secondary metabolites in microorganisms obtained larger acceptance: the role of secondary metabolites as communication signals in complex communities [41]. It is being argued that the high concentrations required for antibiotic activities of compounds are rarely achieved under true physiological conditions whereas subinhibitory concentrations can still lead to global transcriptional changes in the producing microorganisms or their neighbours in the habitat. Indeed, a number of recent studies support a role of different types of secondary metabolites in signalling and cellular interactions. As an example, it could be shown that antibiotics such as erythromycin and rifampicin, at low concentrations, alter global bacterial transcription patterns as measured by the stimulation or inhibition of a variety of promoter–*lux* reporter constructs in a *Salmonella typhimurium* library [42]. In *Myxococcus xanthus*, DKxanthenes, a unique class of pigments, were shown to have a strong influence on fruiting body formation and viability of spores [43]. Another example for an impact of secondary metabolites on cellular interactions is provided by the potent cyanobacterial toxin microcystin that was shown to influence expression and oligomerisation of a surface exposed lectin and hence cell–cell recognition and cell–cell attachment in the colonial cyanobacteria [44].

In the field of plant secondary metabolism, the idea of chemical warfare and of important relations between phytochemical diversity on the one hand and ecological interactions on the other hand became more and more influential from the 1960s. This view has been supported by an increasing number of studies that analysed the co-evolution within ecological systems. Secondary metabolites produced by plants can facilitate host-symbiont signalling as it has been demonstrated for luteolin, a widespread plant flavone that induces *nod* gene expression in rhizobacteria [35]. Plants have also evolved many ways to cope with their insect herbivores. If plants are attacked by generalist or specialist herbivores they react with direct or indirect defence strategies [45]. Examples for plant secondary metabolites conferring direct defense include nicotine in *Nicotiana attenuata* [46] or pyrrolizidine alkaloids of angiosperms [29]. Volatile organic substances (VOCs) of plants are known to attract natural enemies of herbivores. A well studied example is the indole emission of maize that was shown to be induced by fatty acid derivatives in the regurgitants of beet armyworms [47]. Plants were also shown to react on insect egg deposition with emis-

sion of attractants of egg parasitoids [48]. Secondary metabolites of plants may also act as hormones, as it has been discussed for brassinosteroids of *Arabidopsis* that play an important role in control of cell elongation and de-etiolation [36]. The multiplicity of ecological roles of plant secondary metabolites that were described is impressive and has strengthened the opinion that secondary metabolites are not only fortuitously formed substances, but are rather physiologically important compounds, which originated from adaptation processes driven by natural selection. Nevertheless, there is still one facet that is not clarified satisfyingly: Is it really reasonable to assume that *all* secondary metabolites fulfil a specific function? As we shall see, most of the dissent comes from the issue of natural selection and its role for the evolution of secondary metabolism.

6 Role of selection and adaptation on the evolution of secondary metabolites

A prime example for the dissent about how we can explain the evolution of secondary metabolism is a review article published in 1996 with the title: *Phytochemical diversity: adaptation or random variation?* [49], as many publications discussing the topic gave the impression that a decision had to be made between these two alternatives under the tacit assumption that they were mutually exclusive. In this context, we have to remind ourselves of several principle insights of evolutionary biology. Natural selection is the only explanation for adaptation and natural selection can act on populations only if there is variation among its members and this variation is random with respect to the direction of adaptation. In other words: random (i.e., undirected) variation is the basis of natural selection and natural selection, in turn, is the basis of adaptation. Therefore, it makes no sense to regard variation and adaptation as alternatives.

Applying this to the 'secondary metabolism problem', we can say: the more variation in molecular structures an organism is able to produce, the higher the chance to produce a 'starting point' for subsequent adaptation. The process of adaptation then potentially leads to substances that intervene into very specific physiological responses and bind to receptors with sophisticated complementarity, as it is for example unveiled by DNA-binding antibiotics like the bleomycins [50]. How to explain such a

molecular fitting accuracy? It is very likely that at first there was a more or less fitting ancestor molecule, which came into existence by mere chance. In the course of an ecological interaction this molecule and the respective enzyme/enzyme system was passed through a process of fine-tuning. Without that interaction with the environment the ancestor molecule would have possibly been removed from the metabolism. This concept, however, inevitably implies that a certain fraction of the compounds that we subsume under the category 'secondary metabolism' has no function or practically no function at present. Starting from that objection, Firn and Jones proposed a new hypothesis in the early 1990s [51] and have enhanced it continuously over the last 15 years [2, 28]. This so-called Screening Hypothesis is based on the realisation that, for any molecule, it is a rare property to have a potent biological activity. Thus, organisms have to generate as much chemical variability as possible in order to increase the probability to find a molecule that fits into a certain function. From that point of view, the appearance of secondary metabolism has originated from the interplay of creating a vast majority of chemical structures and screening these compounds in their environment for new useful functions. Yet, many scientists seem to feel uncomfortable with that realisation and there is still an intense debate in the literature concerning the question whether we have to interpret our state of knowledge as a lack of knowledge or simply as the actual existence of functionless compounds. Recent contributions to the debate deal with plant volatiles [52–55], starting with a review article about the emission patterns of volatile isoprenoids and possible explanations for these findings and ending up in a dispute about the correctness of the Screening Hypothesis, which indeed declares many secondary metabolites functionless. The critical objection against compounds without current function is that their existence, at least if their fraction exceeds a marginal level, would be inconsistent with evolutionary theory. The reasoning behind this statement is that the production of secondary metabolites causes fitness costs and therefore, when there is no selective advantage conferred by those metabolites genotypes producing them would be selected against [49]. Since selection operates on the current fitness and not on future potentials there would be no or hardly any substance without a function. Consequently, the vast majority of natural products would have to be assumed to fulfil a physiological function that would bring a selective advantage.

There is no doubt that many more compounds will provide an advantage for the producing organisms than we know to date. Nevertheless the 'fitness costs' argument in this general form is not completely convincing. For instance, we have to take into account the time periods necessary to select against secondary metabolites producing genotypes. Strikingly, the discussion seems to be restricted to qualitative statements, whereas quantitative studies are missing. Thinking only in terms of 'fitness costs' without regarding time and without any estimation of these costs could also be used to justify the opinion that there is no range for genetic invention at all: All gene duplications temporarily lead to increased fitness costs, as it will take a while to develop a new beneficial function, and this would mean that all duplication products will be removed more or less immediately due to the fitness costs.

Although there is evidence for unique evolutionary mechanisms leading to diverse chemical structures we still don't understand why so many different structures are selected and maintained even in closely related strains. We are also far from being able to estimate the percentage of secondary metabolites that fulfil true physiological functions. Laboratory conditions that are commonly used to probe functional relevance of secondary metabolites hardly represent the environment. In order to judge the impact of natural selection on diversification and maintenance of secondary metabolites we therefore need a deeper understanding of the ecological roles of secondary metabolites.

Acknowledgements

Rolf Müller and Elke Dittmann were supported by a grant from the German Research Foundation (DFG) within the frame of the Priority Program 'Evolution of Metabolic Diversity'.

References

1 Kossel A (1891) Über die chemische Zusammensetzung der Zelle (On the chemical composition of the cell). *Arch Physiol* 181

2 Firn RD, Jones CG (2000) The evolution of secondary metabolism – a unifying model. *Mol Microbiol* 37: 989–994

3 Vining LC (1992) Roles of secondary metabolites from microbes. *Ciba Found Symp* 171: 184–194; discussion 195–188
4 Altmann KH (2003) Epothilone B and its analogs – a new family of anticancer agents. *Mini Rev Med Chem* 3: 149–158
5 Staunton J, Weissman KJ (2001) Polyketide biosynthesis: a millennium review. *Nat Prod Rep* 18: 380–416
6 Ohno S (1970) *Evolution by gene duplication* , Springer, New York
7 Jenke-Kodama H, Sandmann A, Müller R, Dittmann E (2005) Evolutionary implications of bacterial polyketide synthases. *Mol Biol Evol* 22: 2027–2039
8 Gross F, Luniak N, Perlova O, Gaitatzis N, Jenke-Kodama H, Gerth K, Gottschalk D, Dittmann E, Müller R (2006) Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads. *Arch Microbiol* 185: 28–38
9 Saxena P, Yadav G, Mohanty D, Gokhale RS (2003) A new family of type III polyketide synthases in *Mycobacterium tuberculosis*. *J Biol Chem* 278: 44780–44790
10 Ueda K, Kim KM, Beppu T, Horinouchi S (1995) Overexpression of a gene cluster encoding a chalcone synthase-like protein confers redbrown pigment production in *Streptomyces griseus*. *J Antibiot (Tokyo)* 48: 638–646
11 Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* 39: 121–152
12 Jenke-Kodama H, Börner T, Dittmann E (2006) Natural biocombinatorics in the polyketide synthase genes of the actinobacterium *Streptomyces avermitilis*. *PLoS Comput Biol* 2: e132
13 Lerat E, Ochman H (2005) Recognizing the pseudogenes in bacterial genomes. *Nucleic Acids Res* 33: 3125–3132
14 Lawrence JG, Hendrickson H (2003) Lateral gene transfer: when will adolescence end? *Mol Microbiol* 50: 739–749
15 Bethal V (2006) Mode of action of microbial bioactive metabolites. *Folia Microbiol (Praha)* 51: 359–369
16 Sugiyama M, Thompson CJ, Kumagai T, Suzuki K, Deblaere R, Villarroel R, Davies J (1994) Characterisation by molecular cloning of two genes from *Streptomyces verticillus* encoding resistance to bleomycin. *Gene* 151: 11–16
17 Matsuo H, Mochizuki H, Davies J, Sugiyama M (1997) Production of bleomycin N-acetyltransferase in *Escherichia coli* and *Streptomyces verticillus*. *FEMS Microbiol Lett* 153: 83–88
18 Tang GL, Cheng YQ, Shen B (2004) Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. *Chem Biol* 11: 33–45
19 Brodhagen M, Paulsen I, Loper JE (2005) Reciprocal regulation of pyoluteorin production with membrane transporter gene expression in *Pseudomonas fluorescens* Pf-5. *Appl Environ Microbiol* 71: 6900–6909
20 Ballesta JP, Cundliffe E (1991) Site-specific methylation of 16S rRNA caused by pct, a pactamycin resistance determinant from the producing organism, *Streptomyces pactum*. *J Bacteriol* 173: 7213–7218
21 Liu M, Douthwaite S (2002) Resistance to the macrolide antibiotic tylosin is conferred by single methylations at 23S rRNA nucleotides G748 and A2058 acting in synergy. *Proc Natl Acad Sci USA* 99: 14658–14663
22 Antonenka U, Nolting C, Heesemann J, Rakin A (2005) Horizontal transfer of Yersinia

high-pathogenicity island by the conjugative RP4 attB target-presenting shuttle plasmid. *Mol Microbiol* 57: 727–734

23 Alarcon-Chaidez FJ, Penaloza-Vazquez A, Ullrich M, Bender CL (1999) Characterization of plasmids encoding the phytotoxin coronatine in *Pseudomonas syringae*. *Plasmid* 42: 210–220

24 Dobrindt U, Hochhut B, Hentschel U, Hacker J (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* 2: 414–424

25 Gonzalez-Lergier J, Broadbelt LJ, Hatzimanikatis V (2005) Theoretical considerations and computational analysis of the complexity in polyketide synthesis pathways. *J Am Chem Soc* 127: 9930–9938

26 Oliynyk M, Brown MJ, Cortes J, Staunton J, Leadlay PF (1996) A hybrid modular polyketide synthase obtained by domain swapping. *Chem Biol* 3: 833–839

27 Stinear TP, Mve-Obiang A, Small PL, Frigui W, Pryor MJ, Brosch R, Jenkin GA, Johnson PD, Davies JK, Lee RE et al (2004) Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proc Natl Acad Sci USA* 101: 1345–1349

28 Firn RD, Jones CG (2003) Natural products – a simple model to explain chemical diversity. *Nat Prod Rep* 20: 382–391

29 Croteau R, Karp F, Wagschal KC, Satterwhite DM, Hyatt DC, Skotland CB (1991) Biochemical characterization of a spearmint mutant that resembles peppermint in monoterpene content. *Plant Physiol* 96: 744–752

30 Misawa N, Satomi Y, Kondo K, Yokoyama A, Kajiwara S, Saito T, Ohtani T, Miki W (1995) Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. *J Bacteriol* 177: 6575–6584

31 Durbin ML, Learn GH Jr, Huttley GA, Clegg MT (1995) Evolution of the chalcone synthase gene family in the genus Ipomoea. *Proc Natl Acad Sci USA* 92: 3338–3342

32 Helariutta Y, Kotilainen M, Elomaa P, Kalkkinen N, Bremer K, Teeri TH, Albert VA (1996) Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: evolution with substrate change and catalytic simplification. *Proc Natl Acad Sci USA* 93: 9033–9038

33 Tropf S, Lanz T, Rensing SA, Schroder J, Schroder G (1994) Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. *J Mol Evol* 38: 610–618

34 Zenk MH (1991) Chasing the enzymes of secondary metabolism: plant cell cultures as a pot of gold. *Phytochemistry* 30: 3861–3863

35 Malik VS (1980) Microbial secondary metabolism. *Trends Biochem Sci* 5: 68–72

36 Baba T, Schneewind O (1998) Instruments of microbial warfare: bacteriocin synthesis, toxicity and immunity. *Trends Microbiol* 6: 66–71

37 Zähner H, Anke T (1983) Evolution of secondary pathways. In: JW Bennet, E Ciegler (eds): *Differentiation and secondary metabolism in fungi*. Marcel Dekker, New York, 153–171

38 Davies J (1990) What are antibiotics? Archaic functions for modern activities. *Mol Microbiol* 4: 1227–1232

39 Berenbaum MR, Robinson GE (2003) Chemical communication in a post-genomic world. *Proc Natl Acad Sci USA* 100 Suppl 2: 14513

40 Hartmann T (2004) Plant-derived secondary metabolites as defensive chemicals in herbivorous insects: a case study in chemical ecology. *Planta* 219: 1–4

41 Davies J (2006) Are antibiotics naturally antibiotics? *J Ind Microbiol Biotechnol* 33: 496–499
42 Goh EB, Yim G, Tsui W, McClure J, Surette MG, Davies J (2002) Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc Natl Acad Sci USA* 99: 17025–17030
43 Meiser P, Bode HB, Müller R (2006) The unique DKxanthene secondary metabolite family from the myxobacterium *Myxococcus xanthus* is required for developmental sporulation. *Proc Natl Acad Sci USA* 103: 19128–19133
44 Kehr JC, Zilliges Y, Springer A, Disney MD, Ratner DD, Bouchier C, Seeberger PH, de Marsac NT, Dittmann E (2006) A mannan binding lectin is involved in cell–cell attachment in a toxic strain of *Microcystis aeruginosa*. *Mol Microbiol* 59: 893–906
45 Kessler A, Baldwin IT (2002) Plant responses to insect herbivory: the emerging molecular analysis. *Annu Rev Plant Biol* 53: 299–328
46 Baldwin IT, Halitschke R, Kessler A, Schittko U (2001) Merging molecular and ecological approaches in plant–insect interactions. *Curr Opin Plant Biol* 4: 351–358
47 Frey M, Stettner C, Pare PW, Schmelz EA, Tumlinson JH, Gierl A (2000) An herbivore elicitor activates the gene for indole emission in maize. *Proc Natl Acad Sci USA* 97: 14801–14806
48 Hilker M, Kobs C, Varama M, Schrank K (2002) Insect egg deposition induces *Pinus sylvestris* to attract egg parasitoids. *J Exp Biol* 205: 455–461
49 Berenbaum MR, Zangerl AR (1996) Phytochemical diversity: adaptation or random variation? In: JT Romeo, JA Saunders, P Barbosa (eds): *Phytochemical diversity and redundancy in ecological interactions*, Volume 30. Plenum Press, New York, 1–24
50 Williams DH, Stone MJ, Hauck PR, Rahman SK (1989) Why are secondary metabolites (natural products) biosynthesized? *J Nat Prod* 52: 1189–1208
51 Jones CG, Firn RD (1991) On the evolution of plant secondary metabolite chemical diversity. *Phil Trans R Soc Lon B* 333: 273–280
52 Owen SM, Penuelas J (2005) Opportunistic emissions of volatile isoprenoids. *Trends Plant Sci* 10: 420–426
53 Pichersky E, Sharkey TD, Gershenzon J (2006) Plant volatiles: a lack of function or a lack of knowledge? *Trends Plant Sci* 11: 421; author reply 422–423
54 Firn RD, Jones CG (2006) Do we need a new hypothesis to explain plant VOC emissions? *Trends Plant Sci* 11: 112–113; author reply 113–114
55 Firn RD, Jones CG (2006) Response to Pichersky et al: Correcting a misconception about the screening hypothesis. *Trends Plant Sci* 11: 422

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Biodiversity, chemical diversity and drug discovery

By Sheo B. Singh[1] and Fernando Pelaez[2]

[1]Merck Research Laboratories,
Rahway, New Jersey 07065, USA
<Sheo_singh@merck.com>

[2]Centro de Investigación Básica,
Merck, Sharp & Dohme de España,
S.A., Josefa Valcárcel 38
28027 Madrid, Spain
<Fernando_pelaez@merck.com>

Abstract

Drugs developed from microbial natural products are in the fundaments of modern pharmaceutical companies. Despite decades of research, all evidences suggest that there must remain many interesting natural molecules with potential therapeutic application yet to be discovered. Any efforts to successfully exploit the chemical diversity of microbial secondary metabolites need to rely heavily on a good understanding of microbial diversity, being the working hypothesis that maximizing biological diversity is the key strategy to maximizing chemical diversity. This chapter presents an overview of diverse topics related with this basic principle, always in relation with the discovery of novel secondary metabolites. The types of microorganisms more frequently used for natural products discovery are briefly reviewed, as well as the differences between terrestrial and marine habitats as sources of bioactive secondary metabolite producers. The concepts about microbial diversity as applied to prokaryotes have evolved in the last years, but recent data suggest the existence of true biogeographic patterns of bacterial diversity, which are also discussed. Special attention is dedicated to the existing strategies to exploit the microbial diversity that is not easy to tackle by conventional approaches. This refers explicitly to the current attempts to isolate and cultivate the previously uncultured bacteria, including the application of high throughput techniques. Likewise, the advances of microbial molecular biology has allowed the development of metagenomic approaches, i.e., the expression of biosynthetic pathways directly obtained from environmental DNA and cloned in a suitable host, as another way of accessing microbial genetic resources. Also, approaches relying on the genomics of metabolite producers are reviewed.

1 Introduction

Biodiversity is a concept usually referring to the number of species (or other taxonomical entities) in a given ecosystem or geographic area. For the purposes of this article, biodiversity will be defined as concerning those biological species that are studied specifically for their value as producers of small molecule organic compounds, so-called secondary metabolites. These compounds are isolated, characterized and studied by chemists and biologists to evaluate their biological, pharmacological and/or chemotaxonomic potential, or simply to explore chemical diversity. Terrestrial plants have been the most and longest studied sources, having afforded a large number of natural products that have been used for treatment of various diseases dating back to historic times. Most of the old remedies were based on plant extracts, and this led to the discovery of very useful drugs, such as aspirin and many others. The discovery of penicillin from *Penicillium* spp. in 1930s introduced microbial sources for the discovery of biologically and therapeutically relevant natural products, followed by a slow

but steady progress on discoveries of natural products from marine organisms. All in all about 194,000 natural compounds have been listed in the first volume of the 2006 *Chapman and Hall Dictionary of Natural Products* from all sources. Year by year comparison of the list indicate an increase of about 5% new natural products per year. Of these approximately 18,000 are derived from marine sources (courtesy of Prof. John W. Blunt, University of Canterbury, New Zealand) and reported annually in reviews [1–4]). There are approximately 32,500 natural products reported from microbial sources (Antibase database), including about 1,000 derived from marine microbes. While plants and marine macroorganisms are great sources of secondary metabolites with unsurpassed chemical diversity, they have lost some popularity in the pharmaceutical industry in recent times, mainly due to real or perceived fear of the difficulty of large scale supply. Most of the discussions in this article will be on microbial sources, which in principle can be grown in the artificial conditions to any desirable scale, thus guaranteeing an unlimited and uninterrupted supply of the quantities of raw material needed for drug development.

2 Microbial sources

One of the basic questions to address in any pharmaceutically relevant drug discovery effort from natural products is on what group of organisms is most productive and should be focused to improve the probability of success of finding new chemical entities as leads to build drug candidates. Filamentous fungi and bacteria from the Actinomycetales have been traditionally the focus of industrial and academic groups for more than the last 60 years. The list of breakthrough medicines derived from secondary metabolites produced by these organisms is long and its impact is widespread across many therapeutic areas, from infectious diseases to oncology, immunosuppression, atherosclerosis, and others (Tab. 1). In addition, the number of metabolites with some type of biological activity reported from actinomycetes or fungi is astounding (around 19,000, according to Bérdy [5]), and account for about 60% of all the bioactive microbial metabolites reported to date [5]. It is not clear whether the higher proportion of biologically active compounds reported from microorganisms is due to the fact that they have been historically the major focus of larger pharmaceuti-

cal companies, and were subjected to systematic bioassay guided isolation compared to the other sources such as plants and marines.

It is not unreasonable therefore to argue that since these microorganisms have been already so extensively studied, the chances of finding truly novel metabolites from either actinomycetes or fungi are too low to be worth the effort. However, the evidence suggests that only a minor fraction of all the species or genetically distinct strains of bacteria and fungi existing in nature have been cultured in the laboratories. About 6,000 species of prokaryotes have been described to date, and although the number of existing bacterial species is difficult to estimate [6, 7], the existence of 10^6 to 10^9 bacterial strains has been estimated [8]. Of course, it is not known how many of those will be actinomycetes. As for fungi, about 70,000 fungal species have been formally described (though not all of them have been cultured), but the number of species existing in the environment is likely to be in the range of 1.5×10^6 [9].

Likewise, it is tempting to speculate that the number of known metabolites from actinomycetes and fungi represent only a minor fraction of the total chemical universe produced by these organisms. In fact, mathematical models suggest that the number of antibiotics still to be discovered from actinomycetes could well be above 10^5, suggesting that the number of compounds characterized to date would represent less than $< 5\%$ of the total [10]. New species and even major taxa of fungi and actinomycetes are being discovered every day, opening windows of opportunities and providing evidence that our knowledge of these microorganisms is far from exhaustive. Moreover, these species are producing new chemistries with brand new structural classes possessing interesting biological activities, suggesting that their potential in natural products research is by no means exhausted [11]. Last but not least, the availability of the first sequences of complete genomes of organisms from these groups has confirmed also their enormous potential. For example, the analysis of the genomic sequences of two *Streptomyces* species, *S. coelicolor* [12] and *S. avermitilis* [13], revealed the presence of more than 20 gene clusters encoding for the synthesis of polyketides (PKs) or non-ribosomal peptides (NRPs), the two most important biosynthetic classes of microbial secondary metabolites. These gene clusters are significantly more than the number of compounds actually isolated from these strains. Similar results have been shown in fungi [14]. This suggests that we are still far from having a complete understanding

Table 1
Examples of microbial natural products relevant for the pharmaceutical industry

Biological activity	Metabolite	Biosynthetic type*	Commercial product**	Producing organism
Antibiotic	Penicillins	NRP	Penicillin G, V, Ampicillin, Methicillin, Amoxicillin, Carbenicillin	*Penicillium* spp. *Aspergillus* spp.
	Cephalosporins	NRP	MEFOXIN (cefoxitin), CECLOR (cefaclor), CLAFORAN (cefotaxime)	*Acremonium* spp. *Emericellopsis* spp. *Amycolatopsis lactamdurans* *Streptomyces clavuligerus*
	Thienamycin	NRP	PRIMAXIN (imipenem)	*Streptomyces cattleya*
	Erythromycin	PK	ZITHROMAX (azithromycin)	*Saccharopolyspora erythraea*
	Vancomycin	Glycopeptide	VANCOCIN	*Streptomyces orientalis*
	Streptomycin	AG		*Streptomyces griseus*
	Gentamycin	AG		*Micromonospora purpurea*
	Tetracyclines	PK	TYGACIL (tygecycline)	*Streptomyces* spp., *Dactosporangium* spp., *Actinomadura brunnea*
	Mupirocin (pseudomonic acid)		BACTROBAN	*Pseudomonas fluorescens*
	Daptomycin	NRP	CUBICIN	*Streptomyces roseosporus*
Antifungal	Amphotericin B	PK	AMBISOME, AMPHOTEC	*Streptomyces nodosus*
	Pneumocandin B_0	NRP + PK	CANCIDAS (caspofungin acetate)	*Glarea lozoyensis*
Anthelminthic	Avermectin	PK	IVOMEC (ivermectin)	*Streptomyces avermitilis*
Immunosuppressant	Cyclosporin A	NRP	SANDIMMUNE	*Tolypocladium* spp., other *Hypocreales*
	Tacrolimus (FK506)	PK	PROGRAF	*Streptomyces tsukubaensis*
	Rapamycin (sirolimus)	PK	RAPAMUNE	*Streptomyces hygroscopicus*
	Mycophenolic acid	PK-ISOP	CELLCEPT (mycophenolate mofetil)	*Penicillium* spp. *Verticicladiella abientina* *Septoria nodorum*

Table 1 (continued)

Biological activity	Metabolite	Biosynthetic type*	Commercial product**	Producing organism
Antitumor	Bleomycin	Glycopeptide		*Streptomyces verticillus*
	Doxorubicin	PK	ADRIAMYCIN, DOXIL	*Streptomyces peucetius*
	Daunorubicin	PK	DAUNOXOME	*Streptomyces peucetius* and other *Streptomyces* spp.
Lipase inhibitor	Lipstatin	PK	XENICAL (orlistat)	*Streptomyces toxytricini*
Cholesterol-lowering	Lovastatin	PK	MEVACOR, ZOCOR (simvastatin)	*Aspergillus terreus, Monascus ruber*, other fungi
	Mevastatin	PK	PRAVACHOL (pravastatin)	*Penicillium* spp.
Anti-migraine	Ergotamine	Tryptophan-ISOP	ERGOSTAT, CAFERGOT	*Claviceps* spp.

(*) NRP = non-ribosomal peptide; PK = polyketide; AG = aminoglycoside; ISOP = isoprenoid
(**) Trade names in capitals. Non-capitalized names or names in parenthesis refer to marketed semisynthetic derivatives from the original natural compound. Only some representatives are indicated.

about how to fully exploit the metabolic potential of these microbes under laboratory conditions.

All these arguments, combined with the new technologies used in modern drug discovery, improved cultivation techniques, and highly refined and sensitive analytical techniques that today allow rapid dereplication of known compounds and the characterization of new molecules present at low concentrations, reinforce the idea that there is significant promise in continuing exploring the potential of both filamentous fungi and actinomycetes as sources of new leads for drug discovery. It goes without saying that the new and sensitive biology must play a key role if natural products discovery is to succeed.

Interestingly, secondary metabolic pathways are not evenly distributed across these two large microbial groups. Thus, about 70% of all the metabolites described from *Actinobacteria* are produced by species of one single

genus, *Streptomyces* [15], the rest being distributed across other families, mainly *Micromonosporaceae* and *Pseudonocardiaceae*. Likewise, there are certain genera within fungi, such as *Penicillium*, *Aspergillus*, *Trichoderma* or *Fusarium*, which are much more productive than any of the remaining fungal genera [16].

Other microbial major taxa well known to produce bioactive secondary metabolites include the cyanobacteria and the myxobacteria [5, 17]. Similar to actinomycetes, cyanobacterial genomes are large in size and dedicate a significant portion of genes to the biosynthesis of secondary metabolites, particularly those belonging to the PKs and NRP classes. For example, the genome of *Nostoc punctiforme* appears to carry up to 15 gene clusters of these two classes [18]. A study using degenerate PCR probes revealed the existence of a significant diversity and novelty in the genes encoding for both PK and NRP synthases across a number of strains of cyanobacteria, suggesting that undifferentiated filamentous and heterocystous strains are the most likely source of novel metabolites [19]. More than 600 bioactive metabolites have been described to date from cyanobacteria [5, 20, 21], and some compounds have progressed to clinical trials in the oncology area, exemplified by several derivatives of the dolastatins, metabolites originally isolated by a mollusk (the sea hare *Dolabella auricularia*) but later shown to be produced by cyanobacteria [18, 20, 22].

Myxobacteria (class Deltaproteobacteria, phylum Proteobacteria) also bear very large genomes (around 10 Mbp). Detailed analysis and mapping of the genes encoding for polyketides and non-ribosomal peptides in species of *Sorangium* have revealed the existence of at least 10 clusters of genes encoding the biosynthesis of secondary metabolites [23]. About 80 basic scaffolds and 450 structural variants have been described from myxobacteria [24]. Although no products from myxobacteria have reached the market so far, several compounds are in clinical development for treatment of various forms of cancer. Several derivatives of epothilone, produced by the myxobacterium *Sorangium cellulosum* [23], are among the most promising.

Despite the promise shown by these two classes of bacteria, a more efficient exploitation of these microbial groups in modern drug discovery efforts needs significant improvement in the methods for isolation and cultivation. Axenic cultures are often difficult to obtain, at least in part because of co-cultivation requirements of these with other bacteria in many cases [18, 25, 26]. This certainly represents an obstacle not only

when one needs to generate the large number of isolates required for industrial purposes, but also for the scale up of the cultures needed to obtain enough material for drug development [23, 26, 27]. Despite these difficulties, new approaches to the industrial exploitation of cyanobacteria have been reported. These include improved cultivation techniques to harvest material and combinatorial biosynthesis to generate analogs with improved properties [18].

Other bacteria within the phyla Proteobacteria (e.g., *Pseudomonas*), the Firmicutes (*Bacillus*) and the *Bacterioidetes* have also shown to produce bioactive secondary metabolites, but the level of productivity does not match with that of actinomycetes or fungi. Approximately 800 compounds have been reported from *Bacillus* species, and a similar number from Pseudomonads [5]. One of the most relevant and medically useful secondary metabolite produced by bacteria out of the Actinobacteria is pseudomonic acid (mupirocin) (Tab. 1). The latter is a marketed antibiotic used in clinic for topical application, and is produced by *Pseudomonas fluorescens* [28]. Secondary metabolites produced by *Bacillus* species include many examples of PKs and NRPs, several of which have been studied in some detail regarding the genetics of their biosynthetic pathways [29]. Several NRPs produced by bacilli, such as polymyxin and bacitracin, have been marketed as antibiotics useful for topical application [30].

3 Microbial diversity

When talking about increasing microbial diversity for a natural products screening, even within well studied groups such as fungi and actinomycetes, it is critical to understand the factors driving microbial diversity in natural environments. For fungi there is ample evidence for a fair degree of ecological specialization. Plant endophytes, enthomopathogenic, nematode-trapping, coprophilous, marine or freshwater fungi, to mention just a few, represent different fungal assemblages, notwithstanding the existence of some limited overlap as well as a number of cosmopolitan species, which can be recovered from almost every substrate. Moreover, within some of those groups, a high degree of affinity to specific types of substrata may be seen. For instance, many endophytic and phytopathogenic fungi are specifically associated to particular plant species or groups of species (e.g., [31,

32]). In addition, like terrestrial plants and animals, fungi also exhibit biogeographic patterns of diversity, so much so that many species are not only ecologically specialized but geographically restricted as well. Thus, many of the fungi found in the tropics are never found in temperate regions and *vice versa* (e.g., [33, 34]). Thus, it is possible to find substantially different fungal species by changing the focus of the isolation efforts to different types of natural substrata, ecological niches and geographic sources.

This situation is however far less clear for prokaryotes. In essence, there remains significant scientific debate on whether bacteria exhibit any biogeographic distribution patterns at all. The statement "everything is everywhere – the environment selects", as originally coined by Beijerink and Baas-Becking [35] in the early years of the 20th Century, has pervaded our understanding and concepts on microbial biogeography for decades. According to this hypothesis, and likewise for fungi, different environments contain distinctive microbial assemblages, but the huge dispersal capabilities of bacteria erases the effects of past evolutionary and ecological events, thus preventing the existence of any biogeographic distribution patterns, which are based on the genetic differentiation of populations geographically distant. Cosmopolitan distribution of microbes is usually justified by the enormous size of microbial populations, which would increase dispersal probabilities and decrease the chance of local extinction. However, the fact is that little work has been done to verify this hypothesis in any systematic way (that is, distinguishing between the roles of historical contingencies from contemporary environmental factors), or to study what environmental factors play greater influence on the structure of microbial communities. Many reports have appeared that show that microbial communities are not spatially random and correlations with environmental or geographic characteristics such as latitude, depth, salinity, vegetation and other factors do exist [35]. However, very limited studies have been undertaken to differentiate the role of environment factors *versus* geographic distances for the variation of microbial species. A recent report [36] suggests that bacterial diversity follows very different biogeographical patterns to those that are known for plants and animals. In this study, neither temperature, latitude and those variables typically predicting animal or plant diversity had much influence on microbial diversity. The diversity and richness of soil bacterial communities was largely independent of geographical distance, differing mainly by ecosystem type. The

only factor that was apparently related to bacterial diversity was the pH of soil, with acidic soils showing lower diversity than soils at neutral or basic pH. However, there are a number of contrasting reports in which the effect of both distance and environmental factors were assessed, and that suggest the existence of true biogeographic patterns for different types of prokaryotes [35]. Likewise, free living microbial eukaryotes in marine ecosystems have been suggested to have a worldwide distribution [37]. Importantly, environmental effects have a predominant influence at smaller spatial scales (below a few kilometers); it is at larger scales (hundreds or thousands of kilometers) where the influence of distance is more likely to be detected. An additional evidence for the existence of microbial biogeography comes from studies showing patterns of taxa-area relationships similar to those seen in plants and animals [35, 38].

In any case, even if it were true that microbial diversity is just driven by ecological forces, a prospecting strategy aiming to increase the diversity of microbes for studies of natural products research should include a good range of geographic areas where multiple environmental conditions may be found, preferably those including uncommon and under-explored habitats. Likewise, any sound prospecting strategy should include a thorough exploitation of even the small areas, since often multiple microhabitats can be found within short distances. Just as an example, as many as 13 different eco-regions are identified within the state of Massachusetts, with several sub-ecosystems within each [39].

Another aspect that is not clear is what role time plays in the changes in microbial biodiversity. Changes in environmental conditions might cause changes in microbial population due not only to invasion or dominance by new species better adapted to the new circumstances, but also due to mutations in the same original species, leading to the production of new metabolites to adapt to the new environment (for instance, to fight new predators). It would be interesting to test if a soil sample collected at the same location after significant time (~25 years), during which significant environmental changes may have occurred, would contain identical microorganisms producing identical compounds, or the microbial biota would have changed to adapt to the new environmental conditions, including the appearance of new genes for biosynthetic pathways producing different compounds. Unfortunately, those kinds of experiments are hard to undertake, for obvious reasons.

4 Role of terrestrial and marine habitats

Historically, soil-derived actinomycetes and fungi dominated early natural products screening efforts. Indeed, most of the microbial-derived drugs were produced by (or based on metabolites produced by) soil microorganisms, with few exceptions (e.g., the antifungal agent caspofungin – Cancidas™, a semisynthetic derivative of pneumocandin B_0, produced by *Glarea lozoyensis*, a fungus isolated from water [40]). However, it is well known that there are many other ecological niches and types of substrates besides soil harboring abundant and diverse microbial communities, including actinomycetes and fungi. Some of those habitats have been explored since the late 1980s or early 1990s, and they have shown to be a productive source of metabolically talented microbes. A good example is that of fungal endophytes, living asymptomatically invading plant tissues, that produce interesting metabolites such as the antifungal agents enfumafungin [41] and moriniafungin [42], the parasiticide nodulisporic acids [34, 43], the insulin mimetic desmethyl-asterriquinone B1 [43, 44] and many others (e.g., [45]). It has been claimed that some fungal endophytes could be able to synthesize the same metabolites that are produced by the host plant (due to horizontal gene transfer), the classical example being that of taxol, originally isolated from *Taxus brevifolia*, but apparently also produced by some endophytes of yew species [45]. Other terrestrial habitats such as leaf litter, herbivore dung, lichens, rock surfaces or even arthropods, have shown to be a valuable source of filamentous fungi, actinomycetes and other bacteria that can be used for natural products research purposes [46–49].

The marine environment has been investigated for several decades as a source of bioactive molecules isolated mainly by invertebrates such as sponges, tunicates, bryozoans and mollusks. Not surprisingly, due to the US National Cancer Institute funding for marine natural products initiative, the most relevant examples are a number of antitumor compounds that have reached clinical trials, such as bryostatin, ecteinascidin 743, kahalalide F, aplidine, discodermolide and several others [22]. It has been hypothesized that at least some of the compounds isolated from marine invertebrates would actually be produced by bacterial symbionts. This hypothesis was initially based on the structural and biosynthetic similarities of metabolites isolated from marine macroorganisms that were

of unequivocal bacterial origin, but recent works have shown evidences that this could be indeed the case, since parts of the putative biosynthesis genes for some compounds initially isolated from invertebrates have been cloned from metagenomic libraries obtained from the associated microbial flora [50, 51].

In any case, the marine environment has also attracted a lot of attention as the source of new microbial taxa with potential interest in natural products research. The existence of fungi exquisitely adapted to the marine environment has been known for more than a century, although most of the cataloged species were described after 1950. Obligate marine fungi comprise a list of a few hundred species that are adapted to growth and dispersal in an aquatic environment [52], but in addition, marine substrates such as driftwood, seaweeds, or mangrove parts allow the isolation of many fungi belonging to terrestrial taxa. It is unclear whether these are present in marine habitats just as viable propagules or actually they are physiologically active (i.e., facultative marine fungi). In any case, almost 300 biologically active molecules have been reported from marine fungi or, more generally speaking, from fungi isolated from marine substrates, mostly during the last two decades [53].

The study of marine actinomycetes has quite a different history. Although strains of actinomycetes had been isolated from marine sediments at least during the 1970s, actinomycetes were at that time considered to be rare in the marine environment, and the predominant thought was that they were metabolically inactive in the sediments, present just as dormant spores of terrestrial origin contributed by rivers [54, 55]. It was not until the late 1990s and early years of the new millennium that the interest on marine actinomycetes, from the ecological point of view as well as from the perspective of natural products research, was revived and experienced a considerable boost. Until that point, only one marine actinomycete species (*Rhodococcus marinonascens*) had been formally described [56]. In recent years though, several new genera and species of marine Actinobacteria have been described which are widespread in marine sediments from different geographic origins (e.g., [55, 57–60]). Some of these actinomycetes can be isolated from other marine substrates as well, such as sponges and tunicates [60, 61]. Even more interestingly, strains from these new taxa produce several novel metabolites with antibiotic and antitumor properties, such as the salinosporamides [62], sporolides [63], the terpe-

noid chloro-dihydroquinones [64], the marinomycins [65] and others [58], as well as known metabolites such as rifamycins [66].

In summary, it is clear that expanding the diversity of microbial species, particularly actinomycetes and fungi, is feasible and can be achieved by exploring less explored ecological niches and habitats such as marine environments [17].

5 Microbial diversity and dereplication of strains

Maximizing diversity within the microbial groups, even if it does not ensure success, is one of the cornerstones of any rational strategy designed to find new natural products with biological activity, but it remains an imperfect science and is a critical challenge. It is relatively easy to isolate hundreds or even thousands of strains of fungi, actinomycetes and other bacteria in a relatively short time by using indirect methods ([16, 67]). Moreover, semi-automated platforms for microbial isolation have been described [67]. However, deciding which strains are worth being tested in a screening process is a much more daunting task. Many of the isolates will be clones, and therefore with identical metabolic potential, and it is necessary to minimize the inclusion of redundant isolates in the screening process. A number of strategies have been described to facilitate the 'dereplication' of genetically identical strains, using tools ranging from morphology to chemotaxonomy markers, including molecular approaches [16]. A critical point though is that even a minimum level of taxonomic analysis requires considerable resources. It is necessary to carefully evaluate at what point the dereplication efforts should be focused, considering other alternatives such as the identification of metabolites in the fermentation broths by analytical methods. However, while the latter process has the obvious advantage of being a dereplication based on actual chemical compounds, it does necessitate a fermentation process, and as a result it needs considerable efforts compared to taxonomically oriented approaches, which are applied before the fermentation step. To make things even worse, some secondary metabolites are produced by not only different species of the same genera but also different genera and even relatively distant taxa [68, 69]. This makes biodiversity an imperfect surrogate of chemical diversity, which is actually what one needs to maximize. However, the principle that

maximizing diversity increases the chances of success to find structural novelty remains one of the few valid axioms in industrial natural products screening programs.

6 Exploiting the hidden microbial diversity

It is now commonly accepted that the number of bacterial species living in natural substrates such as soil is much larger than what can be isolated by conventional culturing methods [6, 7]. The estimates range from about 10^4 species per gram of soil, mainly based on DNA reassociation experiments [7, 70, 71] to 10^6 species per gram of clean non-polluted soils [72]. This broad range of estimates is at least in part related to the inherent difficulties associated with making accurate predictions using mathematical models based on limited data [73]. In addition, the ambiguity of the species concept as applied to microbes adds one more level of complexity [35, 70, 71]. In general, the level of taxonomic resolution in studies on microbial diversity is usually much coarser for macroorganisms, so much so that a single Latin binomial may actually refer to a complex of 'cryptic' species, a phenomenon well documented for microbial eukaryotes but likely to apply to other microorganisms as well [74]. As for what natural product research groups may be concerned, there is enough evidence that the number of genetically distinct microbial strains (i.e., different potential producers of novel metabolites) is huge and we have only scraped the surface of this chemical universe after decades of work.

However, if it is indeed true that the number of species to be grown in pure culture largely exceeds those that have been cultured to date, it is not yet clear how many – if any – 'unculturable' bacteria will bear the necessary genes to produce the new compounds that we are seeking. As discussed above, it is well known that the distribution of secondary metabolic pathways seems not to be present across all prokaryotic phyla, and is apparently concentrated in five groups (Actinobacteria, Firmicutes, Bacterioidetes, Cyanobacteria and Proteobacteria) across the more than 50 phyla so far described [6], about half of which have been cultivated in laboratories. Interestingly, those five phyla contain more than 90% of all cultivated and published species [75]. This leaves the question wide open as to how many bacterial species across those that have shown to be difficult to cultivate

under standard conditions will be really productive and remains to be seen how useful these species would be for drug discovery purposes. Actually, the fact that about half the recognized phyla within the prokaryotes do not have any cultured representative [15] may be interpreted in two opposite directions. Those phyla could be either 'talented' or 'non-talented', i.e., capable or incapable of producing secondary metabolites, respectively. In the absence of more substantial genomic information or direct evidence of metabolites produced in culture, it is not easy to predict whether the vast number of species that remain to be cultivated will indeed produce secondary metabolites of interest. However, it is not unreasonable to keep expectations high in the developing strategies aiming to dig deeper in this black box of 'unculturable' bacteria.

In essence, two types of strategies have been developed in the last decade that focus on addressing the exploitation of these genetic resources. The first strategy is directed towards improving the success rate in the cultivation of prokaryotes formerly thought to be reluctant to growth in laboratory conditions. The second strategy refers to the exploitation of those hidden biosynthetic pathways in the absence of any cultivation process of the organisms harboring those genes, by using molecular genetic techniques on DNA directly extracted from the environmental sample; these have often been referred to as 'metagenomic' approaches.

6.1 Improvement of cultivation of intact organisms

The huge difference between the number of microbial species predicted in a given environmental sample and the actual number of strains that can be isolated and cultured from the same sample using conventional microbiological techniques may be due to several factors, from the existence of truly non-cultivable cells to the damage by oxidative stress of otherwise viable cells during the manipulation of the sample, the inhibition of growth by high concentrations of substrate or the absence of signaling molecules produced by neighboring cells from other species [39]. Not all these factors are equally prone to manipulation or correction under laboratory conditions, but at least two of them have been successfully used to increase the number of cultivated species from a sample, namely the addition of cell-signaling agents and the use of oligotrophic isolation media.

The use of oligotrophic isolation media as a way to improve the ability to culture bacteria has been widely reported [75]. For instance, Janssen et al. [76] used diluted nutrient broth as isolation medium on soil samples, combined with extended incubation times (>10 weeks), to recover bacterial species that had not been previously cultured, some of them represented groups of organisms that had few cultivated representatives, such as the *Acidobacteria* or the *Verrucomicrobia*. As for the effect of signaling molecules, the addition of cyclic AMP and several acyl homoserine lactones known to work as signaling molecules in prokaryotes has been reported to significantly increase the percentage of planktonic bacteria that can be cultivated [75, 77].

Other related strategies rely on the use of natural substrata containing living cells as the source of nutrients and potential signaling agents, using membrane systems that allow the diffusion of small molecules but prevent the contamination by undesired bacteria. This strategy was originally described for the isolation of marine bacteria using diffusion chambers incubated in an aquarium simulating the natural environmental conditions [78]. Using this approach a large number of diverse colonies could be recovered, although new phylotypes were not obtained. Some of these bacteria were unable to grow when placed on artificial media, but they did when put in the presence of other microorganisms. Likewise, Ferrari et al. [79] reported the use of polycarbonate membranes as growth support and non-sterile soil slurry as a culture medium, which was expected to provide not only the limited nutrients but also the signaling molecules required for the growth of oligophilic microcolony-forming bacteria. Using this system, the authors reported the recovery of colonies from bacterial species not previously cultivated, including members of the candidate division TM7, which lacked any cultivated representative to date [75, 79].

6.2 Environmental metagenome (eDNA) approach

An alternative way to access the unculturable microorganisms is represented by various approaches that utilize extraction of total DNA from environmental samples, which is digested and ligated into a vector and then introduced into a surrogate or heterologous host. These strategies are based on the knowledge that genes involved in the biosynthesis of

secondary metabolites are usually arrayed in clusters, physically contiguous in the microbial chromosome. Thus, large fragments of DNA can be cloned that allow access to metagenomic DNA or environmental DNA (eDNA) that in theory is capable of expressing partial or complete natural biosynthetic pathways [80–83] in heterologous hosts that do not produce secondary metabolites. *Escherichia coli* was used as a heterologous host in a first approach whereas *Streptomyces lividans* was used in the second approach. A series of novel N-hydroxy amides called terragines were isolated from soil eDNA expressed into a *Streptomyces lividans* host [84], and a series of long-chain N-acyl amino acids [85] and its synthases [86], violacein [87], long-chain acyl phenols [88], long-chain fatty acid enol esters [89], palmitoylputrescine [90], N-acyl arginine and tryptophan derivatives [91], isocyanide [92, 93] and antibiotic tubromycins [94] were isolated using *E. coli* as a host. This approach has the potential to represent the true Holy Grail of natural products discovery, but has yet to yield compounds of any significance and it has been rather disappointing so far. At present, this approach is mainly in the realms of academic laboratories and industry is waiting for initial successes before it can jump in. Actually, some of the companies that pioneered the initial work in this field have either totally abandoned this strategy or re-focused their efforts in a different direction.

A related transgenic approach has been reported for the cultivation of slow-growing fungi, in which cosmid-size genomic DNA isolated from individual fungal colonies was cloned and introduced into *Aspergillus nidulans* host. The resultant transgenic strains were fermented and analyzed for secondary metabolites, leading to identification of two compounds that were not produced by the host control [95]. Many fungi are not amenable to be grown in sufficient scale in artificial culture conditions, but it is feasible to get enough biomass for DNA extraction from very small colonies cultured in the laboratory (as in the case described above), or directly from the field, e.g., as fruiting bodies (e.g., mycorrhizogenous basidiomycetes) or lichens (in the case of lichenized ascomycetes). Thus, this approach presents in principle an attractive method to exploit the metabolic potential of fungi refractory to current cultivation techniques.

In any case, a lot more work needs to be done to validate these approaches and to re-engineer them into authentic industrial-scale platforms if they are to become real novel natural product discovery engines.

7 High throughput isolation approaches

Recent progress has been made in combining the new cultivation techniques described above with systems that allow the high throughput isolation that is desirable to sustain natural products screening efforts. Most of these approaches are based on the concept of extinction culturing, by which the cell populations in the environmental samples are diluted down to the point of containing only one cell type per vessel (plate, tube, well in a microtiter plate, etc.). This idea was originally applied to the recovery and cultivation of marine bacteria by Button et al. [96]. The authors of this study concluded that the number of viable cells in marine samples was much higher than originally expected (approximately 60%), though most marine bacteria reached stationary phase before attaining any visible turbidity, therefore, the low estimates of viability observed by the traditional techniques. However, this concept was not really amenable to high throughput until the use of microtiter plates, automated cell array and imaging systems replaced the use of conventional Petri dishes and visual inspection.

A first high throughput cultivation method using low nutrient media was described by Connon and Giovanonni [97], which allowed the isolation and cultivation of many new marine bacteria (orders of magnitude higher than that obtained by using conventional agar plates), including members of the clade SAR11, which are the dominant bacterial taxa in seawater [98]. Full automated inoculation of microtiter plates for the rapid generation of extinction cultures can be achieved by using automatic dispensing devices. A good example is the MicroDrop® AutoDrop microdispenser system, which is able to deliver very small droplets (as small as 170 pico liter), working similar to an ink jet printer. This system is not only very fast but also avoids having to perform an initial dilution of the samples [99]. This technique has been extensively used to characterize bacterioplankton communities, allowing the cultivation of previously uncultured bacteria [99, 100].

Another related technique that uses gel microdroplets to encapsulate single cells was developed by Diversa Corporation. This encapsulation system is combined with parallel cultivation under low nutrient flux conditions, after which flow cytometry is used to separate those microdroplets containing microcolonies [101]. In addition to high throughput,

this system allows growing the encapsulated microcolonies in a diversity of conditions (including those mimicking the natural conditions of the environmental samples), thus maximizing the chances of success. Apparently, after sorting the microcolonies, most of the cultures (67%) grew into microtiter plates containing organic-rich media to cell densities higher than 0.1 (measured by optical density at 600 nm).

Despite current knowledge that a lot more organisms exist than have been cultured and with high dilution and high throughput techniques can be cultured, the question remains whether these organisms possess biosynthetic genes capable of producing interesting secondary metabolites. Furthermore, even if the organisms were able to produce interesting compounds; will they have the robust growth that is typically required to support the production of detectable metabolites in industrial scale screening laboratories? It is not clear how microcolonies of a small number of cells can be made compatible with the needs of natural products research, where significant biomass is required for production of secondary metabolites in enough mass not only to be detected by the primary and the secondary screens but also to provide sufficient amount for isolation and structure elucidation, leaving aside the needs of additional material for drug development. Although it has been claimed that these new cultivation methods will allow the use of these new prokaryotes in drug discovery programs [39, 75], the true potential of these new approaches, expressed as new biologically active molecules produced by previously unculturable bacteria, remains to be seen. In any case, it is clear that these alternatives present a significant opportunity and need more exploration.

8 Genomics-guided natural products discovery

As mentioned above, microorganisms may contain significantly more biosynthetic genes than compounds actually isolated after their growth in laboratory media [102]. In some cases this is due to (a) bioassay-guided isolation leading to detection only of compounds with a given type of bioactivity in particular, (b) not all biosynthetic pathways are activated in the growth conditions used or (c) totally dormant genes. The knowledge of gene clusters is relatively limited, and as a result the nature of dormant genes could not be known. This could be circumvented by the genome

scanning technology described by the group of Ecopia BioSciences Inc. [102, 103], by which natural product gene clusters are scanned in a high throughput fashion without sequencing entire genomes. This is then coupled with a growing database consisting of genes from a variety of microorganisms encoding diverse known natural products with maximal chemical diversity. This approach appears to be quite successful in the identification of a large number of gene clusters predicting a large number of chemically diverse novel structures. Mining of this gene cluster database with specific bioinformatics tools allows prediction of novel structures which could be potentially produced after culturing the organisms in variety of growth conditions, particularly using high throughput cultivation methods. Structure prediction provides the ability to analyze the fermentation broths by analytical methods for specific compounds.

In a recent report [102], this technology was used to study 60 microbial strains that were known to produce 65 natural products. In addition to 65 gene clusters that encode the same number of natural products, they found >600 gene clusters encoding for the biosynthesis of potentially new natural products. This is exemplified by the discovery of ECO-02301, a novel antifungal agent from *Streptomyces aizunensis* NRRL B-11277, a well characterized organism which is a known producer of bicyclomycin (Fig. 1). The genome scanning identified 11 natural product gene clusters in addition to the gene cluster responsible for the synthesis of bicyclomycin. One of these gene clusters encoded for the synthesis of ECO-02301, and this was predicted by the bioinformatics tool with sufficient accuracy to allow its detection by HPLC in fermentation extracts, and peak guided isolation [104]. As has been illustrated earlier, this method further confirms that even known microorganisms have potential to produce many more compounds (may be more than 10-fold) than have been isolated so far. This method has potential to allow prediction of chemical diversity and allows the scanning for structural derivatives for potential SAR studies. It does not however allow prediction of biological activity. This approach, if really high throughput, and provided that it become cost effective, can be a very effective tool for dereplication of newly isolated microorganism from natural habitats.

It is clear from the discussion above that microorganisms may be capable of producing more compounds than have been isolated so far. Although it is not clear whether some biosynthetic pathways may be silent and never

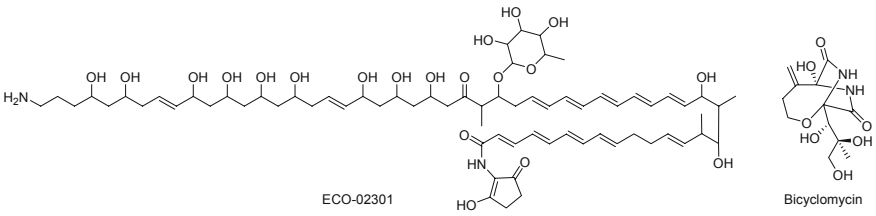


Figure 1
Chemical structures of ECO-02301 and bicyclomycin

expressed, it is likely that the reason why many of those metabolites have not been isolated is due to the limited number and diversity of growth conditions that can be applied to each organism, in the absence of high throughput culturing techniques. However, if that were the case, these technical challenges are being addressed now. Of particular note is the recent introduction by Duetz and collaborators [105, 106] of a new method for cultivation of microorganisms in 96-well plates. This method provides for the first time real opportunities to perform effective fermentations of microbial strains in multiple conditions, minimizing both scale and cost. This system is compatible with standard automation equipment, allowing the parallel handling of the strains and the optimization of culture conditions with manageable resources. It is well known that growth conditions play a critical role in the activation of biosynthetic genes and thus in the production of a variety of secondary metabolites [16]. This high throughput cultivation approach provides the possibility to maximize the exploitation of the metabolic potential of microorganisms.

9 Chemical diversity

It is well recognized that natural products provide unparalleled chemical diversity and has been demonstrated by computational methods that they are more akin to drugs than random chemicals [107, 108]. Natural products are represented by molecules spanning from the simplest of the structures such as aspirin to the highly complex such as palytoxin (Fig. 2) [109]. The former is a highly valuable clinical anti-inflammatory drug, whereas the

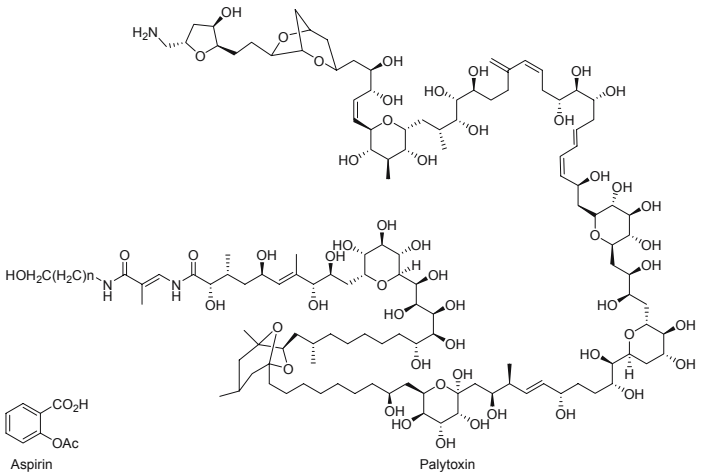


Figure 2
Chemical structures of aspirin and palytoxin

latter is one of the most toxic compounds known to mankind. Structural diversity represented by natural product drugs or potential drug candidates is shown in Table 1 and Figures 3–9. The natural product structures are produced by specific biosynthetic pathways inherent to producing organisms. The complexity of the structure is clearly defined by the complexity of genetics of the organisms. What controls biosynthetic machineries in each organism is unclear. The simplicity and complexity of structures represent advantages and challenges when natural products are identified as chemical leads against a drug target or as drugs. Simpler and smaller structures with differentiating functional groups are more amenable for selective chemical modifications (e.g., from lovastatin to simvastatin and other statins, Fig. 7) than larger complex multifunctional structures (e.g., FK506, cyclosporine (Fig. 8), and caspofungin (Fig. 3)). However, many times those complex structures provide exquisite potencies and selectivity against their biological targets because of multicenter ligand–protein interactions. In several cases, resources and intellect has been dedicated by synthetic chemists to reduce the complexity of the structure and not only retain the activity but occasionally produce simpler yet more potent compounds. This is best exemplified by the preparation of simplified totally synthetic analogs (e.g., analog 2) of bryostatin 1 (Fig. 9) [110, 111].

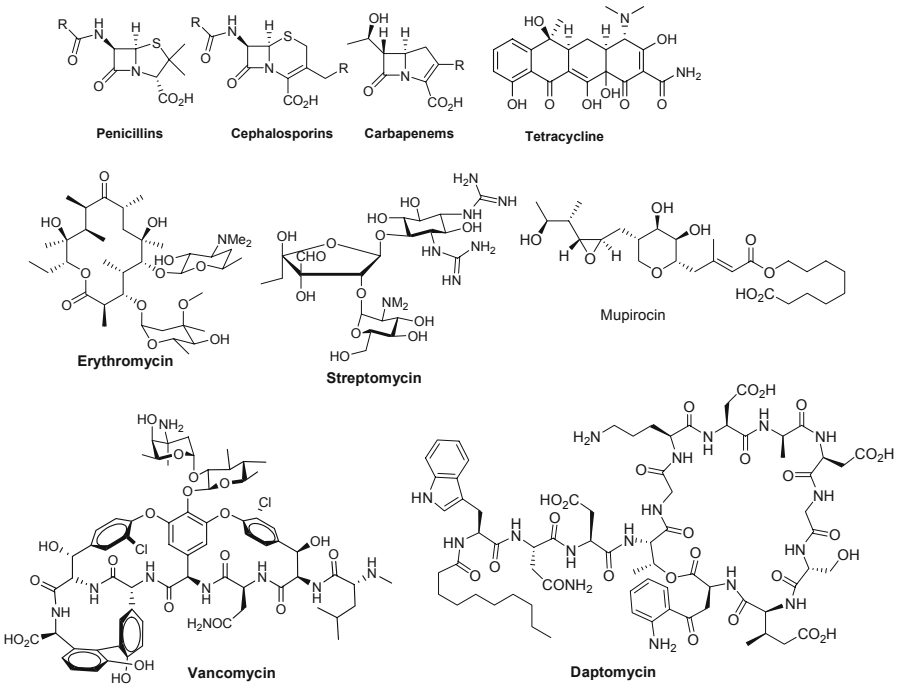


Figure 3
Chemical structures of key antibiotics

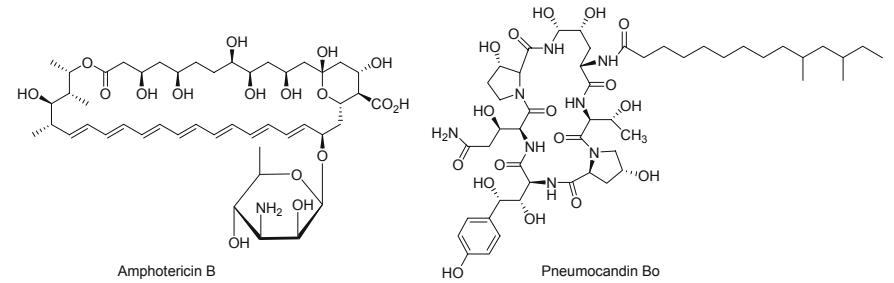


Figure 4
Chemical structures of key antifungal agents

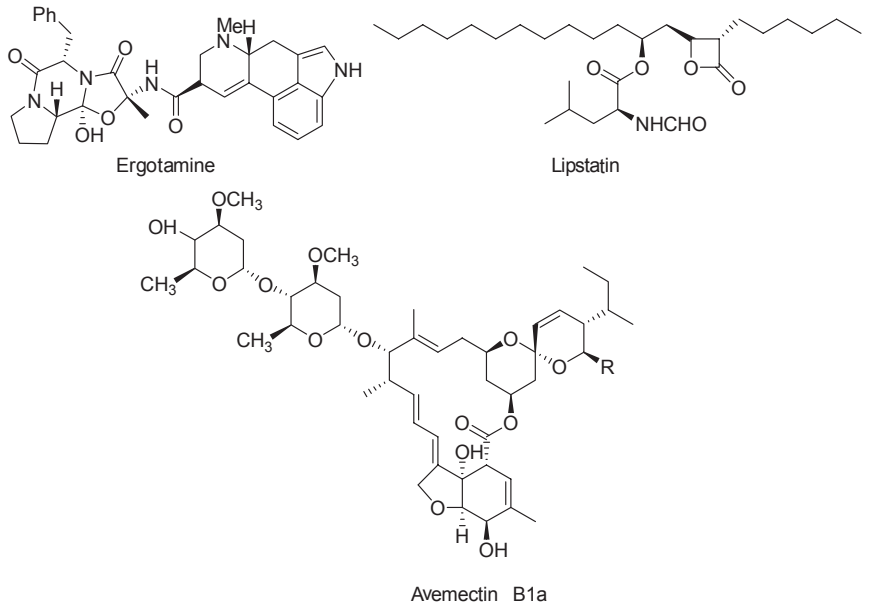


Figure 5
Chemical structures of anthelmintic (avermectin), lipase inhibitor (lipstatin), and anti-migraine (ergotamine) agents

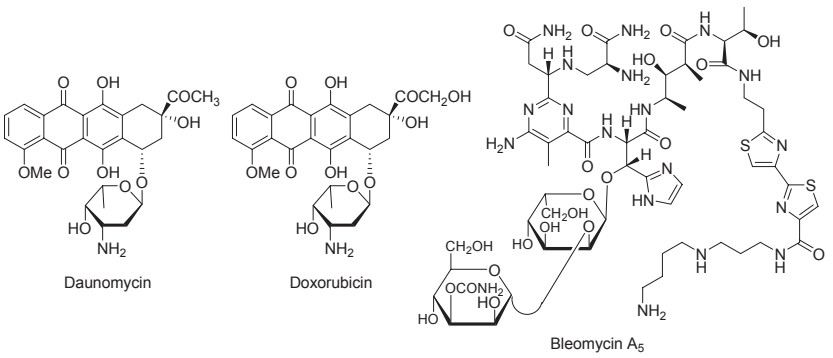


Figure 6
Chemical structures of antitumor agents

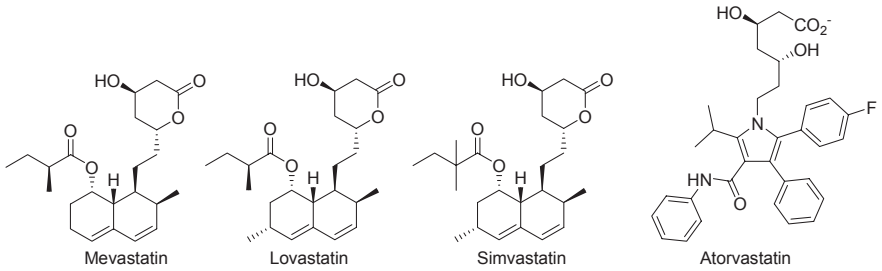


Figure 7
Chemical structures of natural (mevastatin and lovastatin) and semisynthetic (simvastatin) and synthetic (atorvastatin) cholesterol-lowering agents with common natural pharmacophore

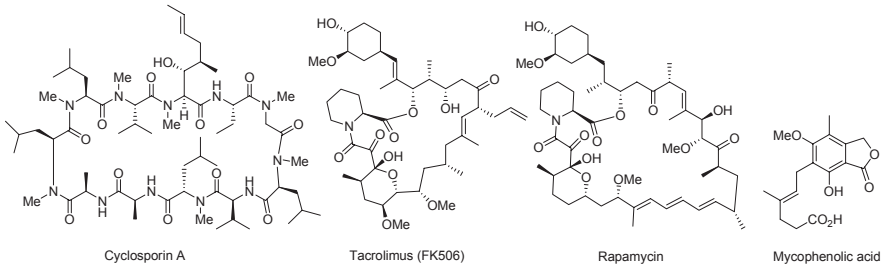


Figure 8
Chemical structures of key immunosuppressants

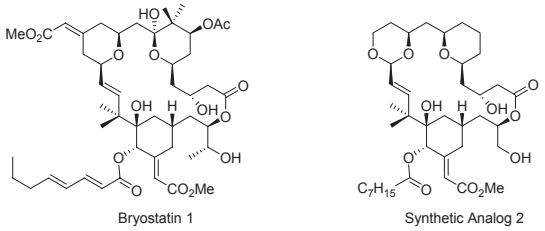


Figure 9
Chemical structures of bryostatin 1 and analog 2 [110, 111]

It is noteworthy that some classes of biosynthetic genes are cosmopolitan, being present in most living organisms such as plants, marine invertebrates, fungi and prokaryotes. These widespread biosynthetic pathways would include polyketide synthases, non-ribosomal peptide synthases, fatty acid synthases and isoprenoid synthases. Interestingly, although they use the same building blocks (aminoacids, acetyl-CoA, etc.) they produce distinctly different compounds, often due to selective decoration by enzymes at specific steps. However, significant differences exist in the basic biosynthetic machineries leading to biosynthesis of drastically different chemical structures as well. For example, *Streptomyces* are less likely to produce isoprenoids than eukaryotic organisms. It is this difference in the biosynthetic machinery that intrigues the natural product community which may lead to the biosynthesis of novel natural products not yet discovered. The question does arise though whether completely new biosynthetic machineries exist that are capable of using different yet undiscovered or known pools of small building blocks in arrays different than known machineries, and thus producing yet unimagined/undiscovered molecules. The question is where those organisms are? How can we find them? And how can we make them express those potential biosynthetic pathways?

10 Drug discovery

The question of why secondary metabolites are produced by living organisms continues to evade us and is vigorously debated. It has been proposed that some of the higher eukaryotes such as plants and animals produce secondary metabolites as a way to defend themselves from predators; microorganisms would produce them to defend from other organisms that surround them or use them as signaling molecules with other organisms [112]. However, it is hard to unequivocally prove either of these hypotheses and only circumstantial evidence has been proposed. Until we understand the reason why these secondary metabolites are produced and how they are used by the producing organisms, the debate as to what organisms should be screened for what drug targets is nothing more than a random approach and mainly based on previous successes. Nevertheless, natural products have taught us numerous lessons, as nicely summarized

by Clardy and Walsh [113]. It is unreasonable to expect that all natural products will interact with specific and required biological targets, and actually only a small percentage do. Of those which do interact, only even smaller numbers will have drugable properties and can be converted into drugs. Therefore, to identify a natural product-biological target pair, one needs to have specific biological assays that can selectively detect these compounds from the mixture of many others. Therefore the success of natural products drug discovery depends on the availability of highly robust, sensitive and high throughput biological screening assays. Tight coupling of robust assays with broad biodiversity and efficient natural products chemistry is the key to success of natural products programs [114], as exemplified by the discovery of platensimycin by Wang et al. [11]. Lack of efficiency in any of the pieces can significantly derail natural products programs. Recently, significant improvements have been seen in all aspects of the technologies and disciplines associated with natural products (microbiology, natural products chemistry and assay technologies). Those groups who can combine them effectively at one place will have the potential to make a profound impact on the discovery of natural product drugs and leads. Many practitioners of today's drug discovery and development rely heavily on the Lipinski's rule of five [115], which is violated by many highly successful natural products drugs that have been critical to saving millions of human lives. If one has to rely on such a rule then natural products and the drug discovery inevitably will benefit from a similar rule tailor made for natural products.

References

1 Blunt JW, Copp BR, Munro MH, Northcote PT, Prinsep MR (2003) Marine natural products. *Nat Prod Rep* 20: 1–48
2 Blunt JW, Copp BR, Munro MH, Northcote PT, Prinsep MR (2004) Marine natural products. *Nat Prod Rep* 21: 1–49
3 Blunt JW, Copp BR, Munro MH, Northcote PT, Prinsep MR (2005) Marine natural products. *Nat Prod Rep* 22: 15–61
4 Blunt JW, Copp BR, Munro MH, Northcote PT, Prinsep MR (2006) Marine natural products. *Nat Prod Rep* 23: 26–78
5 Berdy J (2005) Bioactive microbial metabolites. *J Antibiot* (Tokyo) 58: 1–26
6 Ward BB (2002) How many species of prokaryotes are there? *Proc Natl Acad Sci USA* 99: 10234–10236

7 Curtis TP, Sloan WT, Scannell JW (2002) Estimating prokaryotic diversity and its limits, *Proc Natl Acad Sci USA* 99: 10494–10499

8 Pedros-Alio C (2006) Marine microbial diversity: can it be determined? *Trends Microbiol* 14: 257–263

9 Hawksworth DL (1991) The fungal dimension of biodiversity, magnitude, significance and conservation. *Mycological Res* 95: 641–655

10 Watve MG, Tickoo R, Jog MM, Bhole BD (2001) How many antibiotics are produced by the genus *Streptomyces*? *Arch Microbiol* 176: 386–390

11 Wang J, Soisson SM, Young K, Shoop W, Kodali S, Galgoci A, Painter R, Parthasarathy G, Tang YS, Cummings R et al (2006) Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* 441: 358–361

12 Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D et al (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417: 141–147

13 Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, Sakaki Y, Hattori M, Omura S (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* 21: 526–531

14 Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H et al (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 434: 980–986

15 Zengler K, Paradkar A, Keller M (2005) New methods to access microbial diversity for small molecule discovery. In: L Zhang, AL Demain (eds): *Natural products drug discovery and therapeutic medicine*. Humana Press, Totowa, New Jersey, 275–293

16 Pelaez F, Genilloud O (2003) Discovering new drugs from microbial natural products. In: JL Barredo (ed): *Microorganisms for health care, foods and enzyme production*. Research Signpost, Trivendrum, 1–22

17 Harvey A (2000) Strategies for discovering drugs from previously unexplored natural products. *Drug Discov Today* 5: 294–300

18 Sielaff H, Christiansen G, Schwecke T (2006) Natural products from cyanobacteria: Exploiting a new source for drug discovery. *IDrugs* 9: 119–127

19 Ehrenreich IM, Waterbury JB, Webb EA (2005) Distribution and diversity of natural product genes in marine and freshwater cyanobacterial cultures and genomes. *Appl Environ Microbiol* 71: 7401–7413

20 Burja AM, Banaigs B, About-Mansour E, Burgess JG, Wright P (2001) Marine cyanobacteria – a prolific source of natural products. *Tetrahedron* 57: 9347–9377

21 Harada K (2004) Production of secondary metabolites by freshwater cyanobacteria. *Chem Pharm Bull (Tokyo)* 52: 889–899

22 Simmons TL, Andrianasolo E, McPhail K, Flatt P, Gerwick WH (2005) Marine natural products as anticancer drugs. *Mol Cancer Ther* 4: 333–342

23 Gerth K, Pradella S, Perlova O, Beyer S, Muller R (2003) Myxobacteria: proficient producers of novel natural products with various biological activities – past and future biotechnological aspects with the focus on the genus *Sorangium*. *J Biotechnol* 106: 233–253

24 Reichenbach H (2001) Myxobacteria, producers of novel bioactive substances. *J Ind Microbiol Biotechnol* 27: 149–156

25 Reichenbach H (1999) The ecology of the myxobacteria. *Environ Microbiol* 1: 15–21

26 Gaspari F, Paitan Y, Mainini M, Losi D, Ron EZ, Marinelli F (2005) Myxobacteria isolated in Israel as potential source of new anti-infectives. *J Appl Microbiol* 98: 429–439

27 Burja A, Wright P (2003) Look first to nature before invention industrial applications for marine cyanobacteria. *SIM News* 53: 4–9
28 Pelaez F (2006) The historical delivery of antibiotics from microbial natural products – can history repeat? *Biochem Pharmacol* 71: 981–990
29 Chen XH, Vater J, Piel J, Franke P, Scholz R, Schneider K, Koumoutsi A, Hitzeroth G, Grammel N, Strittmatter AW et al (2006) Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB 42. *J Bacteriol* 188: 4024–4036
30 Walsh C (2003) *Antibiotics actions, origins, resistance*. ASP Press, Washington, DC
31 Shivas RG, Hyde KD (1997) Biodiversity of plant pathogenic fungi in the tropics. In: KD Hyde (ed): *Biodiversity of tropical microfungi*. Hong Kong University Press, Hong Kong, 47–56
32 Collado J, Platas G, Pelaez F (2001) Identification of an endophytic *Nodulisporium* sp from Quercus ilex in central Spain as the anamorph of *Biscogniauxia mediterranea* by rDNA sequencing analysis and effect of different ecological factors on distribution of the fungus. *Mycologia* 93: 875–886
33 Rodrigues KF, Petrini O (1997) Biodiversity of endophytic fungi in tropical regions. In: KD Hyde (eds): *Biodiversity of tropical microfungi*. Hong Kong University Press, Hong Kong, 57–69
34 Polishook J, Pelaez F, Platas G, Ondeyka JG, Dombrowski AW, Teran AM (2001) Biogeography and relatedness of *Nodulisporium* strains producing nodulisporic acid. Mycologia 93: 1125–1137
35 Martiny JB, Bohannan BJ, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR et al (2006) Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* 4: 102–112
36 Fierer N, Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* 103: 626–631
37 Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061–1063
38 Bell T, Ager D, Song JI, Newman JA, Thompson IP, Lilley AK, van der Gast CJ (2005) Larger islands house more bacterial taxa. *Science* 308: 1884
39 Zhang L (2005) Integrated approaches for discovering novel drugs from microbial natural products. In: L Zhang, AL Demain (eds): *Natural products drug discovery and therapeutic medicine*. Humana Press, Totowa, New Jersey, 33–55
40 Bills GF, Platas G, Pelaez F, Masurekar P (1999) Reclassification of a pneumocandin-producing anamorph, *Glea lozoyensis* gen et sp nov, previously identified as *Zalerion arboricola*. *Mycol Res* 103: 179–192
41 Pelaez F, Cabello A, Platas G, Diez MT, Gonzalez del Val A, Basilio A, Martan I, Vicente F, Bills GE, Giacobbe RA et al (2000) The discovery of enfumafungin, a novel antifungal compound produced by an endophytic *Hormonema* species biological activity and taxonomy of the producing organisms. *Syst Appl Microbiol* 23: 333–343
42 Basilio A, Justice M, Harris G, Bills G, Collado J, de la Cruz M, Diez MT, Hernandez P, Liberator P, Nielsen Kahn J et al (2006) The discovery of moriniafungin, a novel sordarin derivative produced by *Morinia pestalozzioides*. *Bioorg Med Chem* 14: 560–566
43 Bills GF, Dombrowski AW, Pelaez F, Polishook J, An Z (2002) Recent and future discoveries of pharmacologically active metabolites from tropical fungi. In: R Walting, JC Fran-

kland, AM Ainsworth, S Isaac, CH Robinson (eds): *Tropical mycology*. CABI Publishing, CAB International, Wallingford, UK, 165–194

44 Salituro GM, Pelaez F, Zhang BB (2001) Discovery of a small molecule insulin receptor activator. *Recent Prog Horm Res* 56: 107–126

45 Strobel G, Daisy B, Castillo U (2005) *Novel natural products from rainforest endophytes*. Humana Press, Totowa, New Jersey

46 Ruibal C, Platas G, Bills GF (2005) Isolation and characterization of melanized fungi from limestone formations in Mallorca. *Mycological Prog* 4: 23–38

47 Salazar O, Valverde A, Genilloud O (2006) Real-time PCR for the detection and quantification of geodermatophilaceae from stone samples and identification of new members of the genus *Blastococcus*. *Appl Environ Microbiol* 72: 346–352

48 Gonzalez I, Ayuso-Sacido A, Anderson A, Genilloud O (2005) Actinomycetes isolated from lichens: evaluation of their diversity and detection of biosynthetic gene sequences. *FEMS Microbiol Ecol* 54: 401–415

49 Gebhardt K, Schimana J, Muller J, Fiedler HP, Kallenborn HG, Holzenkampfer M, Krastel P, Zeeck A, Vater J, Holtzel A et al (2002) Screening for biologically active metabolites with endosymbiotic bacilli isolated from arthropods. *FEMS Microbiol Lett* 217: 199–205

50 Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S (2004) Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc Natl Acad Sci USA* 101: 16222–16227

51 Schirmer A, Gadkari R, Reeves CD, Ibrahim F, DeLong EF, Hutchinson CR (2005) Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge *Discodermia dissoluta*. *Appl Environ Microbiol* 71: 4840–4849

52 Kohlmeyer K, Kohlmeyer E (1979) *Marine mycology.* Academic Press, London

53 Bhaduri P, Mohammad BT, Wright PC (2006) The current status of natural products from marine fungi and their potential as anti-infective agents. *J Ind Microbiol Biotechnol* 33: 325–337

54 Fenical W (1993) Chemical studies of marine bacteria: developing a new resource. *Chem Rev* 93: 1673–1683

55 Mincer TJ, Jensen PR, Kauffman CA, Fenical W (2002) Widespread and persistent populations of a major new marine actinomycete taxon in ocean sediments. *Appl Environ Microbiol* 68: 5005–5011

56 Helmke E, Weyland H (1984) *Rhodococcus marinonascens* sp nov, an actinomycete from the sea. *Int J Syst Bacteriol* 34: 127–134

57 Magarvey NA, Keller JM, Bernan V, Dworkin M, Sherman DH (2004) Isolation and characterization of novel marine-derived actinomycete taxa rich in bioactive metabolites. *Appl Environ Microbiol* 70: 7520–7529

58 Jensen PR, Fenical W (2005) New natural-product diversity from marine actinomycetes. In: L Zhang, AL Demain (eds): *Natural products drug discovery and therapeutic medicine*. Humana Press, Totowa, New Jersey, 315–328

59 Jensen PR, Mincer TJ, Williams PG, Fenical W (2005) Marine actinomycete diversity and natural product discovery. *Antonie Van Leeuwenhoek* 87: 43–48

60 Jensen PR, Gontang E, Mafnas C, Mincer TJ, Fenical W (2005) Culturable marine actinomycete diversity from tropical Pacific Ocean sediments. *Environ Microbiol* 7: 1039–1048

61 Kim TK, Garson MJ, Fuerst JA (2005) Marine actinomycetes related to the 'Salinospora'

group from the Great Barrier Reef sponge *Pseudoceratina clavata*. *Environ Microbiol* 7: 509–518

62 Williams PG, Buchanan GO, Feling RH, Kauffman CA, Jensen PR, Fenical W (2005) New cytotoxic salinosporamides from the marine Actinomycete *Salinispora tropica*. *J Org Chem* 70: 6196–6203

63 Buchanan GO, Williams PG, Feling RH, Kauffman CA, Jensen PR, Fenical W (2005) Sporolides A and B: structurally unprecedented halogenated macrolides from the marine actinomycete *Salinispora tropica*. *Org Lett* 7: 2731–2734

64 Soria-Mercado IE, Prieto-Davo A, Jensen PR, Fenical W (2005) Antibiotic terpenoid chloro-dihydroquinones from a new marine actinomycete. *J Nat Prod* 68: 904–910

65 Kwon HC, Kauffman CA, Jensen PR, Fenical W (2006) Marinomycins A-D, antitumor-antibiotics of a new structure class from a marine actinomycete of the recently discovered genus 'marinispora'. *J Am Chem Soc* 128: 1622–1632

66 Kim TK, Hewavitharana AK, Shaw PN, Fuerst JA (2006) Discovery of a new source of rifamycin antibiotics in marine sponge actinobacteria by phylogenetic prediction. *Appl Environ Microbiol* 72: 2118–2125

67 Okuda T, Ando K, Bills G (2004) Fungal germplasm for drug discovery and industrial applications In: Z An (ed): *Handbook of industrial mycology*. Marcel Dekker Inc, New York 123–166

68 Bills GF, Pelaez F, Polishook J, Diez MT, Harris GH, Clapp WH, Dufresne C, Byrne KM, Nallin-Omstead M, Jenkins RG et al (1994) Distribution of zaragozic acids (squalestatins) among filamentous ascomycetes. *Mycol Res* 98: 733–739

69 Vilella D, Sanchez M, Platas G, Salazar O, Genilloud O, Royo I, Cascales C, Martin I, Diez T, Silverman KC et al (2000) Inhibitors of farnesylation of Ras from a microbial natural products screening program. *J Ind Microbiol Biotechnol* 25: 315–327

70 Torsvik V, Ovreas L, Thingstad TF (2002) Prokaryotic diversity – magnitude, dynamics, and controlling factors. *Science* 296: 1064–1066

71 Torsvik V, Ovreas L (2002) Microbial diversity and function in soil: from genes to ecosystems. *Curr Opin Microbiol* 5: 240–245

72 Gans J, Wolinsky M, Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* 309: 1387–1390

73 Hughes JB, Hellmann JJ, Ricketts TH, Bohannan BJ (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl Environ Microbiol* 67: 4399–4406

74 de Vargas C, Norris R, Zaninetti L, Gibb SW, Pawlowski J (1999) Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. *Proc Natl Acad Sci USA* 96: 2864–2868

75 Keller M, Zengler K (2004) Tapping into microbial diversity. *Nat Rev Microbiol* 2: 141–150

76 Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M (2002) Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl Environ Microbiol* 68: 2391–2396

77 Bruns A, Cypionka H, Overmann J (2002) Cyclic AMP and acyl homoserine lactones increase the cultivation efficiency of heterotrophic bacteria from the central Baltic Sea. *Appl Environ Microbiol* 68: 3978–3987

78 Kaeberlein T, Lewis K, Epstein SS (2002) Isolating 'uncultivable' microorganisms in pure culture in a simulated natural environment. *Science* 296: 1127–1129

79 Ferrari BC, Binnerup SJ, Gillings M (2005) Microcolony cultivation on a soil substrate membrane system selects for previously uncultured soil bacteria. *Appl Environ Microbiol* 71: 8714–8720

80 Yap WH, Li X, Soong TW, Davies JE (1996) Genetic diversity of soil microorganisms assessed by analysis of hsp70 (dnaK) sequences. *J Ind Microbiol* 17: 179–184

81 Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5: R245–249

82 Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C et al (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* 66: 2541–2547

83 Handelsman J (2005) How to find new antibiotics. *The Scientist* October 10: 20–21

84 Wang GY, Graziani E, Waters B, Pan W, Li X, McDermott J, Meurer G, Saxena G, Andersen RJ, Davies J (2000) Novel natural products from soil DNA libraries in a streptomycete host. *Org Lett* 2: 2401–2404

85 Brady SF, Clardy J (2000) Long-chain N-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA. *J Am Chem Soc* 122: 12903–12904

86 Brady SF, Chao CJ, Clardy J (2004) Long-chain N-acyltyrosine synthases from environmental DNA. *Appl Environ Microbiol* 70: 6865–6870

87 Brady SF, Chao CJ, Handelsman J, Clardy J (2001) Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA. *Org Lett* 3: 1981–1984

88 Brady SF, Chao CJ, Clardy J (2002) New natural product families from an environmental DNA (eDNA) gene cluster. *J Am Chem Soc* 124: 9968–9969

89 Brady SF, Clardy J (2003) Synthesis of long-chain fatty acid enol esters isolated from an environmental DNA clone. *Org Lett* 5: 121–124

90 Brady SF, Clardy J (2004) Palmitoylputrescine, an antibiotic isolated from the heterologous expression of DNA extracted from bromeliad tank water. *J Nat Prod* 67: 1283–1286

91 Brady SF, Clardy J (2005) N-acyl derivatives of arginine and tryptophan isolated from environmental DNA expressed in *Escherichia coli*. *Org Lett* 7: 3613–3616

92 Brady SF, Clardy J (2005) Cloning and heterologous expression of isocyanide biosynthetic genes from environmental DNA. *Angew Chem Int Ed Engl* 44: 7063–7065

93 Brady SF, Clardy J (2005) Systematic investigation of the *Escherichia coli* metabolome for the biosynthetic origin of an isocyanide carbon atom. *Angew Chem Int Ed Engl* 44: 7045–7048

94 Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon MR, Clardy J, Goodman RM, Handelsman J (2002) Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. *Appl Environ Microbiol* 68: 4301–4306

95 An Z, Harris GH, Zink D, Giacobbe RA, Lu P, Sangari R, Svetnik V, Gunter B, Liaw A, Masurekar P et al (2005) Expression of cosmid-size DNA of slow-growing fungi in *Aspergillus nidulans* for secondary metabolite screening. In: Z An (ed): *Handbook of industrial mycology*. Marcel Dekker, New York, 167–186

96 Button DK, Schut F, Quang P, Martin R, Robertson BR (1993) Viability and isolation of marine bacteria by dilution culture: theory, procedures, and initial results. *Appl Environ Microbiol* 59: 881–891

97 Connon SA, Giovannoni SJ (2002) High-throughput methods for culturing microorgan-

isms in very-low-nutrient media yield diverse new marine isolates. *Appl Environ Microbiol* 68: 3878–3885
98 Rappe MS, Connon SA, Vergin KL, Giovannoni SJ (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 418: 630–633
99 Bruns A, Hoffelner H, Overmann J (2003) A novel approach for the high throughput cultivation assays and the isolation of planktonic bacteria. *FEMS Microbiol Ecol* 45: 161–171
100 Gich F, Schubert K, Bruns A, Hoffelner H, Overmann J (2005) Specific detection, isolation, and characterization of selected, previously uncultured members of the freshwater bacterioplankton community. *Appl Environ Microbiol* 71: 5908–5919
101 Zengler K, Toledo G, Rappe M, Elkins J, Mathur EJ, Short JM, Keller M (2002) Cultivating the uncultured. *Proc Natl Acad Sci USA* 99: 15681–15686
102 Farnet CM, Zazopoulos E (2005) Improving drug discovery from microorganisms. In: L Zhang L, AL Demain (eds): *Natural products drug discovery and therapeutic medicine*. Humana Press, Totowa, New Jersey, 95–106
103 Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, Ahlert J, Thorson JS, Shen B, Farnet CM (2003) A genomics-guided approach for discovering and expressing cryptic metabolic pathways. *Nat Biotechnol* 21: 187–190
104 McAlpine JB, Bachmann BO, Piraee M, Tremblay S, Alarco AM, Zazopoulos E, Farnet CM (2005) Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example. *J Nat Prod* 68: 493–496
105 Minas W, Bailey JE, Duetz W (2000) Streptomycetes in micro-cultures: growth, production of secondary metabolites, and storage and retrieval in the 96-well format. *Antonie Van Leeuwenhoek* 78: 297–305
106 Duetz WA, Ruedi L, Hermann R, O'Connor K, Buchs J, Witholt B (2000) Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. *Appl Environ Microbiol* 66: 2641–2646
107 Ortholand JY, Ganesan A (2004) Natural products and combinatorial chemistry: back to the future. *Curr Opin Chem Biol* 8: 271–280
108 Feher M, Schmidt JM (2003) Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci* 43: 218–227
109 Tosteson MT, Scriven DR, Bharadwaj AK, Kishi Y, Tosteson DC (1995) Interaction of palytoxin with red cells: structure–function studies. *Toxicon* 33: 799–807
110 Baryza JL, Brenner SE, Craske ML, Meyer T, Wender PA (2004) Simplified analogs of bryostatin with anticancer activity display greater potency for translocation of PKCdelta-GFP. *Chem Biol* 11: 1261–1267
111 Wender PA, Hinkle KW, Koehler MF, Lippa B (1999) The rational design of potential chemotherapeutic agents: synthesis of bryostatin analogues. *Med Res Rev* 19: 388–407
112 Davies J (2006) Are antibiotics naturally antibiotics? *J Ind Microbiol Biotechnol* 33: 496–499
113 Clardy J, Walsh C (2004) Lessons from natural molecules. *Nature* 432: 829–837
114 Singh SB, Barrett JF (2006) Empirical antibacterial drug discovery – foundation in natural products. *Biochem Pharmacol* 71: 1006–1015
115 Lipinski CA (2003) Chris Lipinski discusses life and chemistry after the Rule of Five. *Drug Discov Today* 8: 12–16

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

High impact technologies for natural products screening

By Frank E. Koehn

Natural Products Discovery Research
– Chemical and Screening Sciences,
Wyeth Research,
Pearl River, NY 10965, USA
<koehnf@wyeth.com>

Abstract

Natural products have historically been a rich source of lead molecules in drug discovery. However, natural products have been de-emphasized as high throughput screening resources in the recent past, in part because of difficulties in obtaining high quality natural products screening libraries, or in applying modern screening assays to these libraries. In addition, natural products programs based on screening of extract libraries, bioassay-guided isolation, structure elucidation and subsequent production scale-up are challenged to meet the rapid cycle times that are characteristic of the modern HTS approach. Fortunately, new technologies in mass spectrometry, NMR and other spectroscopic techniques can greatly facilitate the first components of the process – namely the efficient creation of high-quality natural products libraries, bimolecular target or cell-based screening, and early hit characterization.

The success of any high throughput screening campaign is dependent on the quality of the chemical library. The construction and maintenance of a high quality natural products library, whether based on microbial, plant, marine or other sources is a costly endeavor. The library itself may be composed of samples that are themselves mixtures – such as crude extracts, semi-pure mixtures or single purified natural products. Each of these library designs carries with it distinctive advantages and disadvantages. Crude extract libraries have lower resource requirements for sample preparation, but high requirements for identification of the bioactive constituents. Pre-fractionated libraries can be an effective strategy to alleviate interferences encountered with crude libraries, and may shorten the time needed to identify the active principle. Purified natural product libraries require substantial resources for preparation, but offer the advantage that the hit detection process is reduced to that of synthetic single component libraries. Whether the natural products library consists of crude or partially fractionated mixtures, the library contents should be profiled to identify the known components present – a process known as dereplication. The use of mass spectrometry and HPLC-mass spectrometry together with spectral databases is a powerful tool in the chemometric profiling of bio-sources for natural product production. High throughput, high sensitivity flow NMR is an emerging tool in this area as well. Whether by cell based or biomolecular target based assays, screening of natural product extract libraries continues to furnish novel lead molecules for further drug development, despite challenges in the analysis and prioritization of natural products hits.

Spectroscopic techniques are now being used to directly screen natural product and synthetic libraries. Mass spectrometry in the form of methods such as ESI-ICRFTMS, and FACS-MS as well as NMR methods such as SAR by NMR and STD-NMR have been utilized to effectively screen molecular libraries. Overall, emerging advances in mass spectrometry, NMR and other technologies are making it possible to overcome the challenges encountered in screening natural products libraries in today's drug discovery environment. As we apply these technologies and develop them even further, we can look forward to increased impact of natural products in the HTS based drug discovery.

1 Introduction

Successful drug discovery programs can be based on a number of strategies. In screening approaches, limited information is available concerning

the structural characteristics of potential lead molecules. Thus, collections of compounds (libraries) are assayed against the biomolecular target in order to identify compounds (hits) that can then be chemically elaborated into so-called 'lead' molecules. Optimization of these drug leads in terms of potency and physical properties subsequently furnishes compounds suitable for development into clinical candidates. By contrast, so-called 'rational' approaches are based on more detailed information about the target, and this key information then is used to develop chemical strategies to obtain chemical leads. This information may take the form of a small molecule which is already known to bind the target, or conversely, an x-ray structure or homology model of the target into which molecular candidates may be designed or computationally 'docked'. Natural products can furnish leads using either of these approaches. Natural products often possess exquisite potency, and frequently have mechanistic specificity and 'biological friendliness' sufficient to bypass the hit-to-lead phase and or even lead-optimization phase entirely. Natural products are also a rich source of drug leads *via* screening approaches, but have been overlooked in the recent past, partly because of difficulties in creating high quality natural products screening libraries or in applying modern screening assays to these libraries.

Drug discovery in the pharmaceutical industry is often based on a process-driven strategy that incorporates target-directed, high throughput screening (HTS) of chemical libraries followed by rapid hit-to-lead and lead optimization phases [1]. This process is illustrated in Figure 1. In the traditional natural products approach, following detection of a library hit, the active natural product is extracted from the source, concentrated, fractionated and purified, yielding one or more pure biologically active compounds. In order to keep pace with synthetic library screening hits these activities must be accomplished within the time frame for screening and hit identification of synthetic compounds (Fig. 1). Thus, traditional natural products programs based on screening of extract libraries, bioassay-guided isolation, structure elucidation and subsequent production scale-up are greatly challenged to meet the rapid cycle times that are a characteristic of the modern HTS approach [2]. The classical natural products process must often overcome a number of hurdles. The first is to reliably detect robust hits and identify known or otherwise uninteresting compounds (dereplication), in order to avoid replication of effort on compounds that

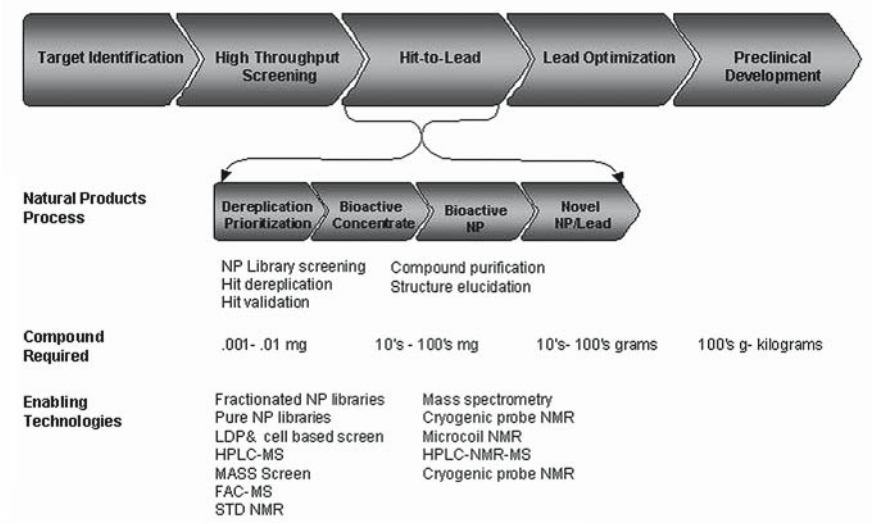


Figure 1.
High throughput screening-based small molecule drug discovery. Upper bar shows general phases of the process while lower bar shows unique natural products screening processes that should be accomplished within the hit-to-lead phase. Amounts of pure compound required for each phase are shown. Enabling technologies for natural products screening are listed at the point where they can maximally impact
Abbreviations: FAC-MS, frontal affinity chromatography mass spectrometry; HPLC-MS, liquid chromatography-mass spectrometry; LDP HTS, lifetime discriminated polarization high throughput screening; MASS, multi-target affinity specificity screening; NP fractionated libraries, natural product fractionated libraries; STD NMR, saturation transfer difference NMR; FT-ICR-MS, Fourier transform ion cyclotron resonance mass spectrometry; HPLC-NMR-MS, liquid chromatography-nuclear magnetic resonance-mass spectrometry.

are not of interest. The second major hurdle in the process, *de novo* structure determination of compounds that are new molecular entities (NME), is an area that has been revolutionized by many advances in spectroscopic techniques, particularly high resolution NMR technologies. While many approaches have been developed to lower the third hurdle, that being the isolation and purification of the active principles (often a minor component) from a complex matrix, it remains as a major bottleneck that continues to impact natural product drug discovery [3]. In the HTS approach, it is essential that new technologies be leveraged in order to facilitate the rapid navigation of natural product hits through the detection, validation,

isolation, hit-to-lead and lead optimization phases. This article will examine the issues surrounding HTS-based natural products drug discovery, and focus specifically on technologies that can be utilized to facilitate the first components of the process – namely the efficient creation of natural products libraries, their screening, and early hit characterization. Of particular note are new technologies in mass spectrometry, NMR and other spectroscopic techniques that have greatly facilitated HTS-based natural product drug discovery programs.

2 Construction and characterization of natural products libraries

2.1 Screening of natural products libraries – general aspects

Historically the pharmaceutical industry has enjoyed considerable success from lead generation programs based on screening crude extract libraries, for a sizable proportion of current therapeutic agents are derived from these efforts [4]. Before the advent of biotechnology, the difficulty of obtaining purified protein targets directly from tissues necessitated that screening be performed using cellular *in vitro* or whole animal systems. The expense and low throughput of these assays made it necessary to screen crude natural products extracts in order to achieve suitable chemical diversity exposure for the assay. When applied to crude or partially purified natural products libraries in the 1970s and 1980s, these screens had good ability to detect active components, in particular with respect to the microbial targets that were the focus of drug discovery efforts. Recent advances in cloning and protein expression have facilitated the production of purified protein targets in sufficient quantities for large-scale cell-free assays to directly detect catalytic inhibition or target binding [5]. These developments in conjunction with sophisticated laboratory automation and miniaturization have greatly expanded the number of targets amenable to HTS and have at the same time expanded screening capacity and throughput. They have also introduced more demanding requirements on screening library integrity and compound physical properties. The end result is the requirement to screen ever-larger libraries against an ever-greater number of targets, in

a limited amount of time, with assays less friendly to natural products extracts.

As an example we consider the case of protein kinases – enzymes which catalyze the phosphorylation by ATP of protein tyrosine or serine/threonine residues, and which play a central role in the cellular signaling pathways controlling activation, growth and proliferation of cells. Numerous HTS technologies have been developed to detect specific kinase inhibitors using expressed and purified kinases. Initially assays were based on measuring the transfer of radioactive phosphate (^{32}P or ^{33}P) from ATP to a protein or synthetic peptide substrate, a robust but low throughput technique. Newer HTS kinase assay platforms are based on fluorescence intensity, time resolved fluorescence or fluorescence polarization [6]. However, early complications in using assays of this type for screening natural products libraries came from two sources. First, in order to detect competitive substrate inhibitors, it is best to screen at a compound concentration close to the K_m of the enzyme for ATP, which is typically micromolar or greater. This poses a challenge with natural products libraries where the relative concentrations of individual components in a sample are often not known precisely, and may vary by three orders of magnitude or more. Consequently, at any given assay dose, the levels of trace components may not be high enough for detection while the highly abundant inactive components may exhibit inhibition *via* non-specific binding, perturbation of the assay pH or other physical properties. Alternatively, the amount of active compounds present may be far in excess, leading to dose independent non-specific inhibition. This problem can be overcome with new approaches such as quantitative high throughput screening (QHTS). QHTS is a titration based approach which can be effective in screening libraries which contain compounds that vary greatly in potency or abundance, and at the same time can furnish concentration-response curves in the primary screen. The approach was demonstrated on a 60,000-member library screened against pyruvate kinase. QHTS relies on assay robustness, automation, and miniaturization in the form of a 1,536 well format [7].

In assays based on readout of fluorescence, difficulties are encountered when natural product samples contain compounds that either emit or absorb radiation at excitation or emission wavelengths of the fluorophore (typically flourescein), or when insoluble components cause light scattering [8]. This liability is shared with synthetic screening libraries as well,

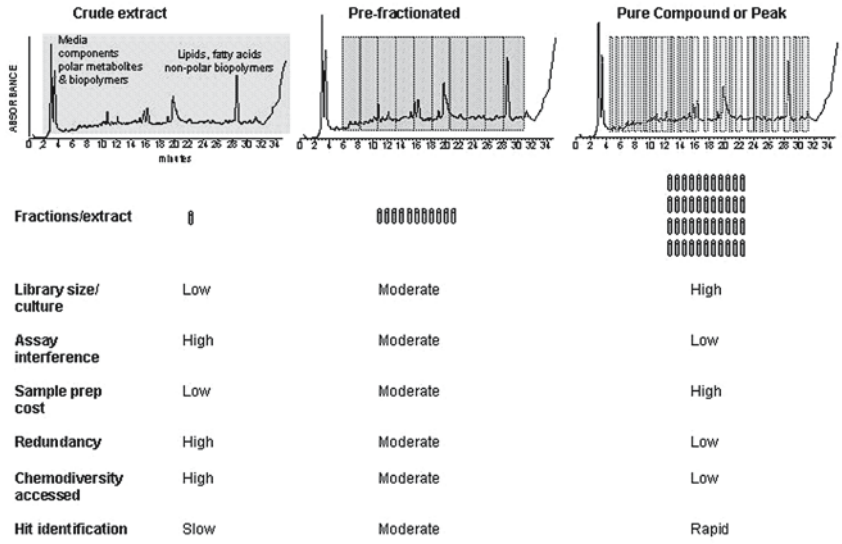


Figure 2.
Comparison of general schemes for constructing natural products screening libraries.
Representative HPLC chromatograms of natural products extracts are shown. Shaded areas depict those components of the extract to be included in the library. Vertical lines show fractionation pattern for pre-fractionated and pure compound or peak-based library.

but the issue is exacerbated with natural products since the presence of these interfering compounds may not be fully characterized. It has now been shown that in kinase, protease and phosphatase assays, fluorescent compound and light scattering interferences can be overcome by increasing the fluorophore concentration in the assay, by using red-shifted wavelength dyes [9], or by the technique of lifetime discriminated polarization [8]. Further, an effective strategy that helps to alleviate these kinds of interferences as well as shorten the time needed to isolate the active principle is the use of purified or pre-fractionated screening samples from the original crude extract (Fig. 2).

When developing an overall screening strategy for natural products it is essential to consider the assays to be employed, the strengths and weaknesses of those assays as they relate to natural product library design, and the technologies and resources available to address the anticipated drawbacks. It cannot be over-emphasized that the results of any high

throughput screening exercise are ultimately determined by the quality of the library itself. High throughput screening of poorly designed or constructed natural products libraries yield few viable hits. Secondly, the construction and maintenance of a high quality natural products library, whether it be based on microbial, plant, marine or other sources remains a costly endeavor. Some basic approaches to natural products screening library design are shown in Figure 2. The library itself may be composed of samples that are themselves mixtures – such as crude extracts (10's to 100's of components per sample), semi-pure mixtures (roughly 3–10 compounds per sample) or single purified natural products. For crude libraries, samples can be prepared with a minimum of processing, often by merely extracting the fermentation broth, plant or animal material with solvent followed by evaporation and plating. Crude extract libraries have lower resource requirements for sample preparation, but require the most resources for identification and dereplication of the bioactive constituents. The heterogeneity of the library samples adds two additional levels of complexity once the hits are detected. The first is that the complexity of natural products crude libraries, and the chemical nature of many of the components found therein, often challenges the robustness of HTS technology. The highly polar or highly lipophilic components of the crude extract can interfere with the functioning of the assay, causing false positives or false negatives (Fig. 2). The second is that one or more rounds of chemical purification and biological assay may be necessary to identify and isolate the active component(s) from the extract. This requires the continued availability of assay resources to support the isolation and purification along with additional time to resolve the hit and furnish pure compound for further biological evaluation. In fast-paced hit-to-lead programs, this constraint is a liability.

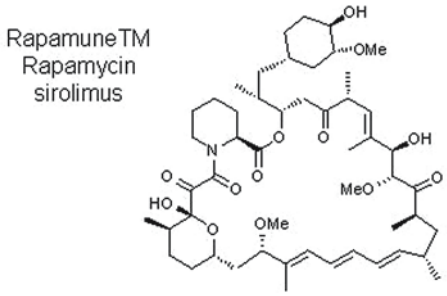
Pre-fractionated libraries can be an effective strategy to alleviate interferences encountered with crude libraries, and may also shorten the time needed to identify the active principle (Fig. 2). There are many variations of this approach and each offers advantages of expediency or purity gained at the cost of up-front partial purification [10–12]. Samples produced by the pre-fractionation approach are simpler mixtures and the final resolution of active components requires fewer purification steps. Interferences are reduced due to the fact that extremely polar and extremely non-polar components are separated from the bulk of the library samples. Moreover,

the relative concentration of minor components is increased over that in the crude, thereby enhancing the opportunity to uncover novel biologically active metabolites. The advantages of the pre-fractionation approach need to be balanced against the resource investment necessary to select, prepare, characterize and maintain such a partially purified natural product library. Since it creates several samples from a single extract, pre-fractionation increases the size (and cost) of the library for a given number of extracts. Given the often-substantial costs associated with HTS assay reagents, especially against high value targets, it is essential that redundancy in screening libraries be minimized. This entails analytical characterization of the natural product library contents in the form of HPLC-MS or other techniques, in order to assure a minimum of redundancy and a maximum of chemical diversity.

Purified natural product libraries offer the advantage that the hit detection process is similar to that for synthetic single component libraries, and the robustness of the hit identification process depends primarily on the purity and chemical integrity of the library itself. Purified libraries offer considerable advantage in the detection of quality hits and in moving forward immediately, since the bioactive principle requires no isolation from a mixture. The downside to this approach is that even with modern methods of separation and automation, substantial resources are required to prepare pure natural product libraries and trace components will not be fully captured in pure form [12]. However, the increasing migration of the industry toward precision ultra high throughput screening coupled with the overall accelerated pace of drug discovery have prompted a move toward highly processed or even pure natural product libraries [13]. Advances in automated separations and sample processing have made purified natural products libraries much more achievable [11]. Besides familiar separation methods such as reversed phase HPLC, newer applications of countercurrent and super critical fluid extraction and chromatography are finding use in natural products library construction [14, 15]. In any library strategy it is essential to accurately characterize the content of the library in order to minimize duplication and maximize chemical diversity and integrity [16]. The purity and integrity maintenance of a sizable pure compound library is an often-underestimated technical challenge [17]. This is particularly true for pure natural products libraries that contain complex unidentified molecules.

Pure natural products work well for screening if care is taken to avoid unstable or so-called problematic compounds. They have high chemical diversity and biochemical specificity, which differentiate them from synthetic and combinatorial compounds [18, 19] and which make them favorable lead structures for drug discovery [20]. Natural products are marked by higher numbers of chiral centers and greater stearic complexity than either synthetic drugs or combinatorial libraries, and while synthetic molecules contain significantly higher numbers of nitrogen, sulfur and halogen containing groups, natural products bear higher numbers of oxygen atoms [19, 21]. Natural products can be differentiated from trade drugs or other synthetic molecular libraries on the basis of scaffold architecture and pharmacophoric properties [22]. They differ significantly from synthetic drugs and combinatorial libraries in the ratio of aromatic ring atoms to total heavy atoms (lower), number of solvated hydrogen bond donors and acceptors (higher), by greater molecular rigidity [23], and they have a broader distribution of molecular properties such as molecular weight, logP and diversity of ring systems. Indeed, less than one fifth of the ring systems found in natural products are represented in current trade drugs.

It is often cited that natural products lack suitable drug-like properties or are structurally too complex for efficient post-screen hit-to-lead development. Detailed analysis however shows that high-quality natural product libraries compare quite favorably in terms of drug-like properties as well. Feher and Schmid examined representative combinatorial, synthetic and natural product compound libraries on the basis of molecular diversity and 'drug-likedness' properties such as molecular weight, number of chiral centers, molecular flexibility as measured by number of rotatable bonds and ring topology, distribution of heavy atoms, and Linpiski-type descriptors [21, 24]. In the overall picture for drug-like properties, Schneider and Lee determined that the fraction of natural product structures with 2 or more 'rule-of-5' violations is equal to that of trade drugs, approximately 10% [22]. These differentiating characteristics of natural products *versus* synthetic drugs are illustrated in Figure 3, which presents a comparison of drug-like properties for the natural product rapamycin and the synthetic drug Pantoprazole.



Molecular weight	914	383
H-bond Donors	3	1
H-Bond Acceptors	13	8
Polar Surface Area	195	102
Rotatable bonds	6	7
Heavy atoms	65	26
N+O	14	7
Lipinski violations	3	0

Figure 3.
Comparison of drug-like properties of natural product and synthetic drug molecules

2.2 Analysis of natural products HTS hits

The large size of libraries now being subjected to ultra high throughput screening has introduced new challenges in analysis and prioritization of natural products hits. It is important to consider the nature of HTS data itself in order to appreciate this issue. High throughput screening in large pharmaceutical companies is frequently performed in 'campaigns' during which an entire library is rapidly assayed in one or more phases. Such a screening campaign often consists of a so-called primary phase, followed by confirmation assay of the primary 'hits' and perhaps one or more additional assays to identify 'false positives'. A 'typical' HTS data set profile, whether it is an enzyme-based or cell-based assay, is shown in Figure 4. The assay responses are distributed around a central value in a manner that can be statistically characterized [25]. Positive 'hits' are contained in an extended upper value 'foot' area. If the hit rate for the primary assay is 1%, (a typical value) and the natural products library consists of 50,000–

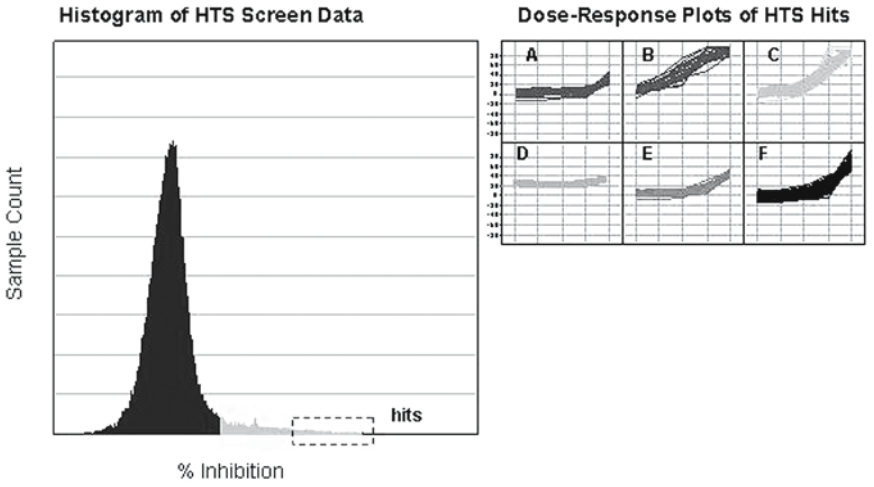


Figure 4.
Profiles of high throughput screen data
Left panel is plot of sample number *versus* % inhibition. Boxed area denotes samples designated as 'active'. Right panel depicts dose response curves of active samples. Those samples in bins E and F show curves dose-response suggesting specific active components. Curves in other bins suggest non-specific activity.

100,000 samples, then the primary screen yields approximately 500–1,000 hits. These primary hits are often re-assayed to confirm the initial assay results. Depending on the statistical method of designating primary hits and the noise characteristics of the primary assay, 30–75% of the primary hits will be confirmed in the confirmation assay, giving anywhere between 150–750 confirmed hits to be investigated. If the natural products library is composed of pure compounds, then the confirmed hits can immediately be identified and prioritized according to hit-to-lead potential and other important criteria. Many of these criteria are the same as those for synthetic hits, and would normally include purity, integrity, molecular complexity, potency, intellectual property constraints, and other drug-like or lead-like properties such as molecular weight, solubility, membrane permeability and stability [26].

When the natural products library consists of crude or partially fractionated mixtures, then the samples should be profiled, often by LCMS, to identify the known components present – a process known as derepli-

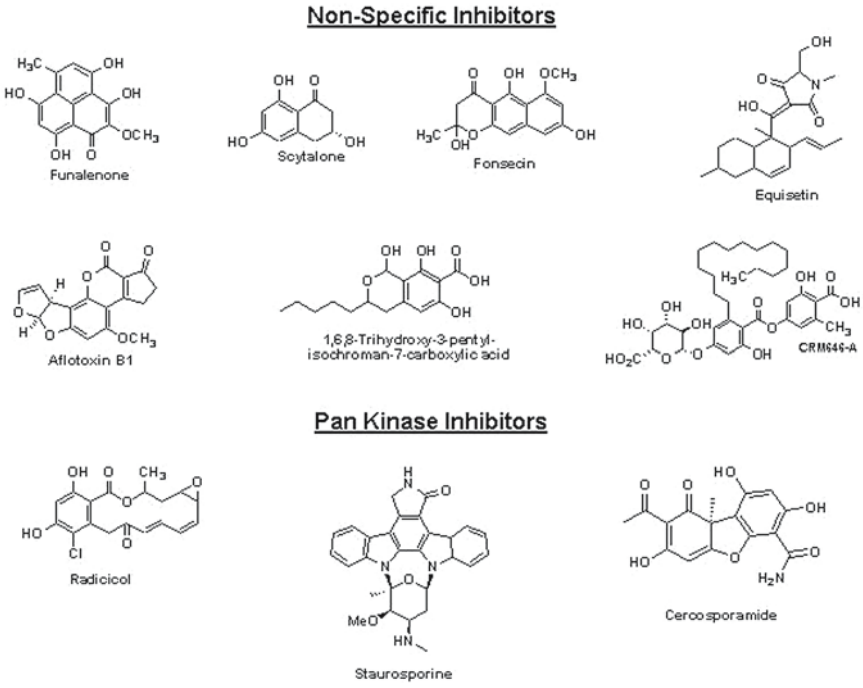


Figure 5.
Examples of common non-specific and pan inhibitors frequently identified in high throughput screening of microbial natural products

cation. Recent advances in HPLC-MS have greatly facilitated this process. Additional criteria are employed to prioritize the hits if the library is of significant size, because the number of initial active samples is likely too large for each hit to undergo bio-assay guided fractionation. Criteria must be employed to ascertain the likelihood of the observed activity being due to the presence of a natural product of interest. One of these criteria is the presence of so-called non-specific compounds or pan-inhibitors such as those shown in Figure 5. For example, the potent pan-kinase inhibitor staurosporine is produced in many actinomycete fermentations and is frequently identified as active in kinase assays [27]. Indeed, staurosporine itself has served as a lead for the development of several clinical candidates [28], however it might not be of interest if novel structural types were

desired. Staurosporine itself might also mask the presence of additional inhibitors if the extract is not well fractionated.

It is important to select a platform for high throughput screening which has a high degree of compatibility with the natural products library to be screened. A wide variety of screening platforms have been developed for the major categories of drug targets – enzymes [29], G-protein coupled receptors (GPCRs) [30] and ion channels [31]. A full review of these specific HTS technologies is beyond the scope of this chapter, but a mention is made here of the increasing usage in high throughput screening of cell based reporter assays engineered for specific targets [32]. These assays appear to be more amenable to natural product screening than many enzyme assays. For example, a cell-based approach led to the isolation of icariin (Fig. 6), a strong agonist of the neuromedin-U2 receptor [33]. To search for inhibitors of tumor cell multi-drug transporters, a fluorescent cell based assay was developed to screen a natural products library of 3,523 compounds for modulators of the multi-drug transporter protein component ABCG-2 in an engineered cell line. A total of 11 compounds of moderate potency, including eupatin (Fig. 6) were identified [34]. While there is as yet insufficient data to support conclusions, it may be that cell-based systems will yield improved effectiveness with natural product libraries, given the biological compatibility of natural products themselves.

Whether by screening of large extract libraries, or smaller more defined natural products compound libraries, the HTS approach continues to yield highly active natural products with novel structures, biological activity and drug potential. These are shown in Figure 6. To discover platensimycin investigators at Merck screened a library of 250,000 microbial extracts using a combination of target based, whole-cell and biochemical assays. These efforts led to the isolation of a new structural class of potent and selective inhibitors of β-ketoacyl-(acyl-carrier-protein synthase I/II) (FabF/B), a key enzyme in the biosynthetic pathway of fatty acids in bacteria [35]. Additional screening using the same anti-sense based whole-cell approach led to the isolation of phomalenic acids A-C, which do not posses suitable drug-like properties, but which further validate the approach [36]. Similarly, the novel antibacterial suaveolindole was discovered using natural products library screening and high-throughput natural products microchemistry [37]. An elegant example has been the

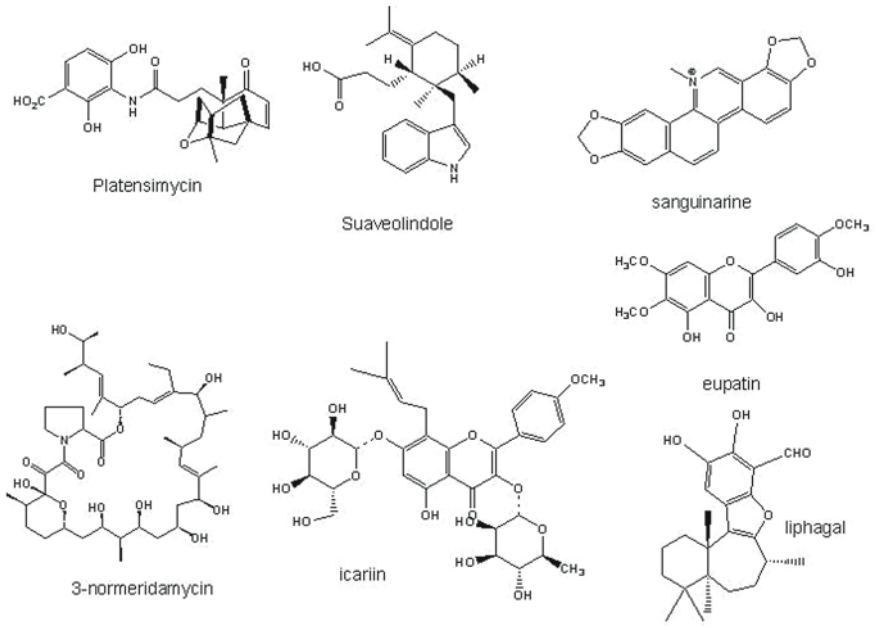


Figure 6.
Recently isolated bioactive natural products from high throughput screening

discovery by screening that sanguinarine is a selective and cellular active inhibitor of mitogen activated protein kinase phophatase-1 of MKP-1 [38]. This work employed a high content chemical complementation assay to screen a diverse but compact natural products library for cellular MKP-1 activity. The significance lies in the fact that there currently is insufficient structural information on MKP-1 to enable the rational design of inhibitors. New signal transduction inhibitors recently discovered include normeridamycin [39]. In this case, screening of a microbial extract library for compounds which bound FKBP12 identified normeridamycin as a trace component in an actinomycete culture. The compound was found later to have potent neuroprotective activity. Marion and co-workers employed a whole cell fluorescence polarization-based assay using human PI3 kinase α expressed in SF9 insect cells to screen a marine extract library and isolated liphagal, a selective inhibitor of PI3 kinase α [40].

2.3 Mass spectrometry and natural products libraries

Mass spectrometry has long been a primary tool in the structure elucidation of natural products and in post-screening stages of drug discovery such as hit validation, dereplication and characterization. The high sensitivity and unique molecular information provided by mass spectrometry have also made it an indispensable tool in the characterization of natural products libraries as well. This is due mainly to the development of reliable directly coupled high performance liquid chromatography-mass spectrometer (HPLC-MS) systems, and the general availability of natural product databases [41].

To avoid redundancy, reduce time and cost, and perhaps most importantly, maximize the 'hit quality' of a natural products library it is essential to select extracts for the library which contain the desired types of secondary metabolites. Whether or not a microbial culture produces secondary metabolites is primarily determined by the genetic capacity of the organism coupled with response of the microbe to the fermentation conditions. Given the fact that most microbial cultures do not produce secondary metabolites, it is important to identify collections of organisms and culture conditions that result in secondary metabolite production, as well as extraction and processing methods which retain secondary metabolites while maintaining compatibility with HTS platforms. In terms of analytical methodology, HPLC and HPLC-MS have become the tools of choice to address these needs. The reason is that a single run LCMS data set of a natural product mixture contains an unsurpassed wealth of chemometric data useful in characterizing the chemical content of an extract [42]. This information includes chromatographic retention time, UV-VIS absorbance spectrum, mass abundance, molecular mass, and molecular formula if sufficient mass accuracy is obtained. All of these data can be incorporated into a database and statistically leveraged to gain an accurate profile of the library sample. The choice of which of these parameters to use and the type of screening algorithm employed should be determined for the most part by the design of the natural products library. As described above, crude or semi-crude natural products libraries offer lower cost of construction and processing, but the evaluation of cultures for unique metabolite production presents a challenge because of the complexity of the composition. Candidate extracts for library inclusion should show the presence of

unique secondary metabolites that enhance the existing diversity of the library. The relative abundance of these molecules should be such that they would give a detectable assay response at micromolar activity. The extract should have a minimum of the many so-called 'promiscuous' metabolites, compounds known to possess nonspecific activity in biological assays, or so-called 'Pan-inhibitors'. Common examples of these are shown in Figure 5. The presence of these compounds, particularly in large amounts, can mask the activity of other specific, perhaps novel actives, and so they are problematic. Media components in the extracts should be identified as well. Thus the chemical fingerprint by HPLC-MS must be processed and ranked according to the number of compounds present, their uniqueness, relative abundance, and the presence of interferences. Crude, pre-fractionated and purified compound libraries might differ in the relative weight of these parameters in the culture selection process.

A number of approaches have been developed to compare and rank chromatographic profiles. Initial approaches were based on transforming the HPLC chromatogram of an extract into a retention time-peak area matrix, emphasizing the most abundant peaks. This approach can be difficult to apply to complex chromatograms with widely different amounts of metabolites. In an early chemometric approach, Higgs and co-workers found that the mean natural log of the area under the entire HPLC-ELSD chromatogram could be used to determine the presence of secondary metabolites. This same group then developed a direct infusion electrospray mass spectrometry technique that gave comparable utility for assessing microbial production. They utilized this approach to select fermentation conditions for Actinomycete secondary metabolism [42]. Recently, approaches have been developed which apply pattern recognition routines such as factor analysis, principal components analysis and cluster analysis to chromatograms and mass spectra. These methods can be automated and utilize a much greater proportion of the chromatogram, perhaps even the entire trace. Other methods have been developed to compensate for the misalignment of peaks due to drifts in compound retention in natural products extracts between runs. These operate either through alignment algorithms [43, 44] or inclusion of internal standards for calibration. Another approach, applied effectively to the selection of extracts from organisms fermented under multiple media conditions, makes use of the principle of a virtual chromatogram generated by co-addition of mul-

tiple individual chromatograms. The various methods described above, coupled with rapid advances in computer and instrument capabilities have made the automated analysis and comparison of chromatographic extract profiles a valuable tool in increasing the chemodiversity of collections, media development and even investigation of taxonomic relationships between organisms [45].

Mass spectral databases are now a mainstay of the dereplication process, where molecular ion and fragment product masses contained in HPLC-MS spectral libraries facilitate rapid assessment of HTS hits and library components [46]. The pivotal development responsible for the success of the LC-MS has been the introduction of efficient and general methods for producing ions from the effluent of HPLC separations. The most general of these methods known as electrospray ionization (ESI) and atmospheric pressure ionization (API), can generate the ions essential for mass spectrometric analysis for greater than 90% of analytes, ranging from amino acids to proteins and nucleic acids [47]. Correlation of both molecular weight and UV absorption data with known compounds by database searching ordinarily is sufficient to classify sets of compounds [41]. There are many such mass spectral database libraries available, and while there are limitations to these libraries, they have become powerful tools for natural products profiling and dereplication [48]. The advent of multichannel HPLC-MS technology has made it possible to assay the chemical integrity and content of sizable natural product libraries, such as a 36,000 sample fractionated plant library [49]. Often it is the in-house developed databases that are most effective, since they are tailored to meet the particular needs of the specific laboratory. For example, Fredenhagen and co-workers used an ESI-HPLC-MS system interfaced with an electrospray (ESI) source and ion-trap mass spectrometer to develop an extensive MS/MS dereplication spectral library [50]. Natural products libraries can also be effectively profiled and the source organisms can be classified using a metabolomics based analysis of HPLC-MS data. [51].

2.4 NMR and natural products libraries

The rich chemical information, versatility, resistance to interference and high dynamic range of NMR spectroscopy make it an unsurpassed tool in

the investigation of natural product mixtures [52]. NMR spectra can be qualitatively or quantitatively analyzed with one of many pattern recognition techniques in order to characterize natural products libraries or evaluate whether an extract should be included in a library. The most popular of these methods is principal components analysis (PCA). These multivariate methods arose out of the field of metabolomics of biofluids [53], and they are gaining acceptance in the library analysis of natural products, particularly plant libraries [54, 55]. Pierens and co-workers have extended this approach by developing clustering methods for NMR spectra of marine natural product mixtures. The aim of the group was to use flow-injection microcoil NMR methods to characterize and dereplicate extracts in microtitre plates [56]. High throughput capillary scale microcoil NMR can effectively be used to assess the chemical novelty and provide prompt structure information on plated screening library components without the necessity of re-isolation in larger quantities. Using microcoil probes at 600 MHz, one and two-dimensional homonuclear spectra can be readily obtained on a few micrograms of sample. Two-dimensional heteronuclear measurements can be performed on 200 micrograms or less. These compound amounts are readily obtained in a typical library primary fractionation [57]. The NMR samples can be loaded into the probe for measurement by automated direct sampling of the 96-multiwell plate [58]. Highly complex mixtures in limited quantity can be extensively characterized by NMR, even without separation – for example crude spider venoms [59]. At the time of this writing, there are no literature reports of microcoil applications to microbial fermentation libraries. Other high-sensitivity NMR technologies such as cryoprobe technology and 1 mm capillary technologies have been applied in similar fashion to synthetic libraries, and reports on applications to natural products are now beginning to appear in the literature for these NMR technologies [60, 61].

Natural products extracts are a logical choice for characterization by hyphenated HPLC-NMR-MS. In HPLC-NMR effluent from a chromatographic column is directly analyzed in the NMR spectrometer without the need for peak isolation and sample preparation. Although LC-NMR was introduced over 20 years ago [62], its application to direct analysis of natural products did not become feasible until several technical hurdles were overcome. Some of these, such as chromatographic peak diffusion, are derived from the process of hyphenation. Others are inherent in the NMR

technique itself, such as low sensitivity and detection of weak compound signals in the presence of background solvent signals. These were a stumbling block to early LC-NMR applications [63]. Recent advances in NMR, mass spectrometry and HPLC and SFC technology have made hyphenated LC-NMR, and LC-NMR-MS practical options for screening [64] and analysis of complex natural products mixtures, particularly those from plants [65, 66]. Early applications of LCNMR used either 1) on-flow detection, where the spectrum is measured as the column effluent passes through the NMR flow-probe, or 2) stopped-flow detection, where the chromatographic flow is switched off in a timed fashion to 'park' the peak of interest in the flow probe for measurement. The drawback to these early methods is that in on-flow measurements, the limited measuring time and sample flow reduce the quality of the spectrum. In stopped-flow applications, peak diffusion out of the NMR flow cell and on the column broadens the sample peak, reducing sensitivity and resolution. Advances in the HPLC-NMR technique were offered by the development of peak capture and peak management methods that effectively eliminated the need to perform the NMR measurement during the chromatographic run. In each of these methods, chromatographic peaks are trapped in a sample handling device during the chromatographic run and later selectively delivered to the NMR probe for measurement. The first is the use of loop storage units that reduced peak the diffusion problems encountered with earlier stopped-flow or on-flow methods [67]. Recent developments include on-line automated solid phase extraction and peak trapping which improve sensitivity by narrowing of the peaks while at the same time reduce the need for solvent suppression. [68, 69].

HPLC-NMR requirements for solvent suppression have benefited from methods which were initially developed for protein NMR [70]. Finally, the inherent low sensitivity of HPLC-NMR (approximately 1,000-fold less sensitive than mass spectrometry) has been improved to a great degree by the development of cryogenic flow probes [71], microcoil (nanoliter volume) NMR probes, and sample management methods described above [72]. Each of these sensitivity enhancement technologies has inherent advantages and disadvantages, depending on the type of chromatography used and whether or not the sample of interest is truly mass limiting [73]. In any event, it is now feasible to fully characterize complex natural products mixtures using one or more LC-NMR methodologies, even to the extent of

minor components [74]. Direct coupling of electrospray mass spectrometry for LC-NMR-MS has proven an effective combination for characterization of compounds that might be difficult to study by LC-NMR alone. Since the NMR measurement is the most costly in terms of measuring time, the mass spectral information is very useful in determining which of the many metabolites encountered in the natural products mixture is of interest for measurement. [75]. Using LC-UV-SPE-NMR-MS it is now possible to perform automated analysis of natural products extracts where the individual components are present in 10–50 microgram amounts [69, 76].

3 Screening of natural products libraries

3.1 Screening of natural products libraries by mass spectrometry

Mass spectrometry has become a viable assay method for high throughput screening of natural products. This is due to the ability to detect and analyze with high sensitivity and mass accuracy ligand-target complexes *via* electrospray, APCI and other 'soft' ionization techniques. Significant successes have been achieved in the use of electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR) to screen natural product microbial libraries against RNA targets. In a method termed Multitarget Affinity/Specificity Screening (MASS), Hofstadler has shown the ability to detect specific binding of the aminoglycoside antibiotic paromomycin in chromatographic fractions derived from cultures of *Streptomyces rimosus* ssp. *paromomycinus* [77]. Non-covalent adducts of paromomycins with a synthetic RNA oligomer containing the *E. coli* A-site (site of action of aminoglycoside antibiotics) were observed in the mass spectrometer. The molecular weights of the aminoglycosides were determined by the differences in mass between adducts and the free RNA. Mass spectrometry can also be used to screen for covalent adducts of protein targets. Liu and co-workers developed a method based on matrix-assisted laser desorption time-of-flight mass spectrometry as an initial high-throughput screening step for natural products which covalently modify cysteine groups in Keap1. This approach is based on the hypothesis that ubiquitination and proteosome-mediated degradation of Nrf2 in the cytoplasm decreases upon the covalent modification of cysteine sulfhydryl groups on Keap1 (a

protein that sequesters Nrf2 in the cytoplasm) and results in higher Nrf2 levels both in the cytoplasm and in the nucleus. Two active constituents were identified to be xanthohumol and xanthohumol D. In a subsequent cell-based assay, xanthohumol and xanthohumol D were confirmed to be potent inducers of quinone reductase, and reaction with Keap1 was also confirmed [58].

Fourier transform ion cyclotron resonance mass spectrometry has also been effective in screening a natural product-based combinatorial library for binding to bovine carbonic anhydrase II (βCAII) [78]. The fungal product 3-chloro-4-hydroxyphenylacetamide was the library template, with 11 secondary amide analogs of this template constituting the combinatorial library. Tight binding inhibitors of βCAII were identified by detection of a non-covalent complex corresponding to [βCAII + ligand] in the mass spectrum. The equilibrium dissociation constant (K_i) for the complex was measured to be 77.4 nM. A competitive βCAII enzyme binding assay validated the mass spectrometry screening result. Mass spectrometry has recently been used to identify inhibitors of protein aggregation [79], and also protein–protein interactions as well, when used in conjunction with NMR based screening methods. Examples here include screens for inhibitors of collagenase and RGS4 protein–protein interactions [80].

A key development in the validation of natural product hits and subsequent bio-assay guided fractionation has been the use of hyphenated mass spectral methods [46]. One wishes to identify each of the 'active' binding components as well as assign rank order to multiple binders according to binding affinity. An effective method to accomplish this comes by the coupling of mass spectrometry with frontal affinity chromatography (FAC-MS), first developed in 1998 [81, 82]. FAC-MS works *via* the continuous infusion of the ligand mixture over an immobilized target (protein). A continuous binding equilibrium is reached between the bound and unbound populations of ligand, the result being that individual ligands that bind the target are retained and the elution volume required for the 'front' of ligand to exit the column is increased over that for non-binding molecules. The degree to which the breakthrough volume is increased over the void is a direct measure of the binding affinity. Those ligands with the greatest affinity for the target will have the longest 'breakthrough' times. Using mass spectrometry as a detector in FAC is invaluable since it allows for selected or universal monitoring of the breakthrough times of each mixture

component by molecular mass. In natural products samples, the masses of the binding principles are often not known. The utility of FAC-MS can be extended by the inclusion of an indicator molecule in the assay, in this case a molecule which binds the target and who's known 'breakthrough' time is perturbed (typically reduced) by the binding of other ligands. By the use of weak ligands as indicators, the effective range of detection for FAC-MS extends from picomolar to 0.5 millimolar. The K_d values for specific ligands can be accurately measured by plotting the breakthrough volume as a function of ligand concentration [83]. Furthermore, the screening capacity of FAC-MS can be as high as 10,000 compounds/day, depending on the number of compounds per run and conditions [83]. In the case of natural products, non-ideal characteristics of the extracts would limit the screening capacity for initial hit identification. However, the potential for FAC-MS as an analytical biosensor to characterize natural products screening hits, once they are detected, is intriguing. To date, the method has been most effective in the deconvolution and post-screen identification of active components in crude natural product extracts [84].

3.2 Screening of natural product libraries by NMR

NMR methods for detecting ligand–protein binding to targets have existed since the early 1980s, although the effective use of NMR to screen chemical libraries for drug discovery began a decade later [85]. The delayed implementation was due to the need for improvements in automation, sensitivity and software, design and synthesis of suitable chemical libraries, and availability of protein targets suitable for NMR binding experiments. Although NMR requires the use of expensive instrumentation, and throughput is lower than screening methods based on photometric or chemical detection, it is becoming clear that NMR methods can fill a unique niche in the screening of chemical libraries for low affinity ligands [86]. An NMR screen detects ligand–protein binding *via* binding-induced modulation of the either 1) the target (protein) spectrum (*target-detected*) or 2) the ligand spectrum (*ligand-detected*). In the first case, the altered protein spectrum gives information on the protein sites affected by the interaction. In the second case (ligand-detected methods), binding information about the ligand molecule is provided. Numerous experimental methods

and NMR pulse sequences have been developed for both approaches, each with its own distinct advantages and disadvantages [87]. Which approach is most effective depends on the nature of the information desired, the availability and size of the target protein, and the binding constant (K_D) of the protein–ligand complex.

In general, target-detected methods rely on the use of isotopically labeled protein, either ^{15}N, ^{13}C, or both. Upon exposure in solution to binding ligands (either single or mixture) the native protein spectrum is perturbed, i.e., chemical shifts of interacting resonances are altered. Various types of high throughput two-dimensional spectra can be recorded for the native and bound protein. Depending on the type of experiment measured and the extent of spectral assignment possible for native protein and complex, the pattern of spectral perturbation yields information as to the specific protein residues involved in the binding interaction. Ligand mixtures are often used in the binding experiment to improve throughput. Additional NMR experiments are performed to deconvolute the identity of the specific binding components once a 'hit' is detected. The pioneering work in this area was reported by the Abbot group, using a fragment-based library approach combined with ^{1}H-^{15}N HSQC two-dimensional spectra, a method they described as SAR by NMR [88]. Under the SAR by NMR protocol, a synthetic chemical library composed of small molecular weight fragments is combined with the target protein solution. Based on perturbations of the chemical shifts of the ^{1}H-^{15}N HSQC signals in the protein (which are previously assigned to their sequence position), the protein binding regions for each ligand can be estimated. Binding interactions are noted for those fragments that bind to different, but proximal (within a few Angstroms) sites on the protein. The molecular fragments, which typically bind very weakly due to their sub-optimal size, are then joined *via* a chemical linker to furnish a starting drug lead. The rationale is that binding synergy will occur upon linkage of the fragments. In the first report of the method, the Abbot group used SAR by NMR to find high affinity ligands for FK506 binding protein [88]. The Abbot group later extended the method to use protein labeled with ^{13}C in the methyl groups of valine, leucine and isoleucine in order to improve sensitivity [89]. By using cryogenic probe technology and compound libraries of 100 fragments per sample, they showed it is possible to screen libraries of 200,000 compounds in a few weeks with a single 500 MHz NMR spectrometer, using less than ten

nanomoles of protein [90]. This can still be a sizable protein requirement, depending on the difficulty of protein expression and purification. There are also limitations as to the size of proteins amenable to the SAR by NMR method. When the target becomes large, i.e., MW > 30 kDa then signal overlap in the HSQC spectrum and relaxation effects can diminish the quality of the results. The protein size limit can be extended by he use of ^{13}C and deuterium labeling [89] or TROSY methods [91].

The use of fragment libraries and compound mixtures in NMR screening is a popular strategy owing to the specific advantages associated with fragment-based libraries [92]. Arguably, molecular fragment based approaches access greater chemical diversity per unit effort when compared to other molecular library methods. Lack of chemical diversity however, is not a shortcoming of natural product libraries – rather it is their accurate characterization. In any event, the use of protein detected NMR methods such as SAR by NMR has made an impact in the screening of synthetic libraries [90], but its use to directly screen natural products libraries has been limited. Natural product libraries are most often themselves mixtures of unknown composition, so it is likely that the difficulty of natural product hit deconvolution has been an obstacle to the use of SAR by NMR for natural products screening [93].

Ligand-detected NMR screening methods are effective for detecting relatively weak binding complexes with K_D's in the high nanomolar to low millimolar range. As a result, they have received greater application in natural products screening. Ligand-detected methods are based on a number of NMR observables, including binding induced line broadening, magnetization transfer of saturation, binding-induced diffusion effects, and binding competition with known ligands [86]. The most popular of these in terms of natural products screening has been saturation transfer difference (STD) spectroscopy [94]. In the STD experiment, the target protein is mixed with a large ligand excess (ca. 100:1 ligand:protein) and the sample is irradiated at the edges of the spectrum to excite the protein resonances. When ligand molecules bind to the protein, rapid spin diffusion results in the signals of the ligand being saturated along with the protein. Molecules that do not bind the protein are unaffected. A difference spectrum is measured between the spectrum resulting from the saturation transfer and an off-resonance control, resulting in a spectrum bearing signals only from molecules which bind to the protein. Signals from the non-binding ligands are filtered out.

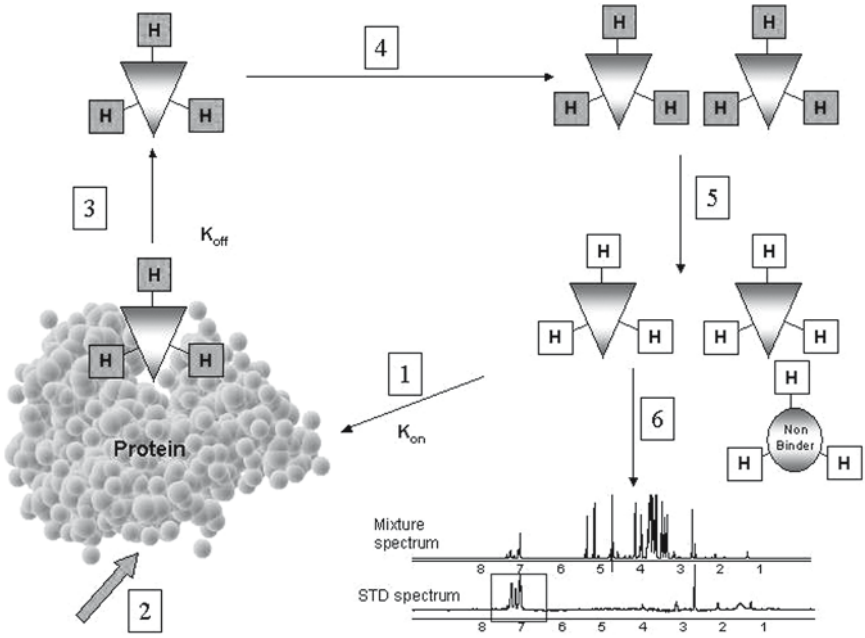


Figure 7
Principles of saturation transfer difference (STD) NMR in the STD experiment, a large excess of screening small molecule mixture composed of *binders* (triangle) and *non-binders* (ellipse) is exposed to target protein. The process is composed of the following steps: 1) Free ligand molecule (triangle) binds protein at rate constant K_{on}, 2) Irradiation saturation is applied to protein resonance but away from the ligand resonances. Bound ligand molecule (triangle) is saturated via spin diffusion, 3) Saturated ligand dissociates at rate constant K_{off} and free ligand population (binders only) becomes saturated, 4) Saturation of free ligand is built up in bulk population, 5) Saturation of free ligand is lost by relaxation. Re-association to protein occurs and saturation transfer cycle is repeated. 6) Difference spectrum is measured between on and off protein resonance, resulting in signals (in box) only from those protons on binding ligands.

This is depicted in Figure 7. STD NMR spectroscopy utilizes significantly less protein than target-detected methods, on the order of 0.1 nanomole per sample. An additional advantage is that for larger molecular weight ligands such as natural products, ligand protons in direct contact to the protein show stronger STD enhancements those that are not in direct contact. This can be especially advantageous for complex natural products, for the STD enhancements yield important information about the structural elements responsible for binding. In addition, the chemical exchange rates between

bound and unbound ligands which make a system favorable for STD studies, namely binding in the micromolar range, also make the ligand–protein binding favorable for transferred NOE studies. A recent example of this application is the epitope mapping of he antimitotic tubulin binding agent HTI-286, a close analog of the marine peptide hemiasterlin [95]. Through adjustment of the STD experimental parameters to account for the effects of differential relaxation, Milton and co-workers were able to characterize the essential structural elements in HTI-286 that were involved in tubulin binding. They also determined the bound conformation of the molecule using STD and transferred NOE methods. Perhaps the best example to date on the use of transferred NOE NMR methods to characterize the tubulin binding properties of natural products is that of epothilone A and its binding conformation [96]. The STD technique offers the advantages of relative robustness with respect to ligand type and lack of dependence on protein size. In fact, larger target proteins have been shown to give superior results due to increased rates of spin diffusion. STD NMR has been applied to the direct screening of crude aqueous natural products extracts from the mushroom *Pleurotus ostreatus* against a target lectin protein [97].

Several other ligand-detected NMR screening methods based on some type of saturation or NOE transfer have been developed. Some of these either have been applied to natural products screening directly, or hold out the potential for successful use. The NOE pumping technique is in essence a reversal of the STD method, where the loss of a ligand signal intensity caused by NOE-mediated magnetization transfer from the ligand to the target is detected in a one-dimensional experiment [98]. Since this method relies on specific excitation of the ligand, its utility for screening mixtures is somewhat limited. However, the method is effective for determining binding characteristics of ligand molecules once they have been identified [98]. The WaterLOGSY method is based on magnetization relay from the bulk solvent H_2O to ligand protons in the complex *via* the water molecules trapped by the protein–ligand interaction. WaterLOGSY offers the advantage of increased sensitivity over other methods such as STD [99]. When used in conjunction with a SHAPES-based fragment library to find ligands for an RNA target, it was found to be the most sensitive of the techniques evaluated [100].

Diffusion based methods rely on the fact that the diffusion rate of a free ligand is diminished upon binding to a macromolecule target [101].

Surprisingly, diffusion based methods have received less attention for natural product screening than one might expect, given the initial successes reported in use of the techniques on natural products type samples. One such early example for drug discovery is the use of Affinity-NMR to unravel the specific binding components in a mixture of tetrapeptides that bind vancomycin [102]. As in most applications, the primary limiting factor in the application of NMR screening to natural products has been the limited availability of well-characterized natural products libraries. As NMR instrumentation becomes increasingly more powerful, we can anticipate that NMR screening of natural products libraries will see increased application in the years ahead.

4 Conclusion

Natural products research is rich in its contribution to drug discovery. It can be argued that no other source of chemical diversity can compare with the depth and breadth of therapeutic agents that have their origins in the secondary metabolites produced by living organisms. These contributions have come *via* numerous strategies, some direct and focused, some by serendipitous routes, and many by screening. The rational approach to lead discovery, using natural products as biologically validated starting points, continues to be an effective means of developing drug candidates. Today's drug discovery environment however, with its increasing reliance on automated high throughput screening of chemical libraries and rapid hit-to-lead development, does not favor traditional natural products screen-based approaches. This is because the enhanced chemical diversity offered by natural products libraries is offset by the often-lengthy (by contemporary standards) periods needed for screening, hit validation and dereplication. If screening of natural products is to continue as an important engine for lead generation, it will be necessary to reduce the time and resources needed to assay natural products libraries and to identify and isolate the important bio-active components. Many of the necessary economies can be found in the construction of improved natural products libraries themselves, and in the means by which they are screened. Current and emerging advances in mass spectrometry, NMR and other technologies are making it possible to overcome these challenges. As we apply

these technologies and develop them even further, we can look forward to continued and even increased impact of natural products in HTS based drug discovery.

Acknowledgements

The author gratefully acknowledges Guy Carter, John Morin and Jeff Paslay for helpful discussions. The author is also indebted to colleagues at Wyeth Natural Products Discovery Research, whose daily research contributions provide the basis for views cited in this chapter.

References

1 Steinmeyer A (2006) The hit-to-lead process at Schering AG: strategic aspects. *ChemMedChem* 1: 31–36

2 Baker D, Mocek U, Garr C (2000) Natural products vs. combinatorials: a case study. *Special Publication – Royal Soc Chem* 257: 66–72

3 Potterat O, Hamburger M (2006) Natural products in drug discovery – concepts and approaches for tracking bioactivity. *Curr Org Chem* 10: 899–920

4 Newman DJ, Cragg GM, Snader KM (2003) Natural products as sources of new drugs over the period 1981–2002. *J Nat Prod* 66: 1022–1037

5 Silverman L, Campbell R, Broach JR (1998) New assay technologies for high-throughput screening. *Curr Opin Chem Biol* 2: 397–403

6 Zaman GJR, Garritsen A, de Boer T, van Boeckel CAA (2003) Fluorescence assays for high-throughput screening of protein kinases. *Combinatorial Chem High Throughput Screening* 6: 313–320

7 Inglese J, Auld DS, Jadhav A, Johnson RL, Simeonov A, Yasgar A, Zheng W, Austin CP (2006) Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc Natl Acad Sci USA* 103: 11473–11478

8 Fowler A, Swift D, Longman E, Acornley A, Hemsley P, Murray D, Unitt J, Dale I, Sullivan E, Coldwell M (2002) An evaluation of fluorescence polarization and lifetime discriminated polarization for high throughput screening of serine/threonine kinases. *Anal Biochem* 308: 223–231

9 Turek-Etienne TC, Lei M, Terracciano JS, Langsdorf EF, Bryant RW, Hart RF, Horan AC (2004) Use of red-shifted dyes in a fluorescence polarization AKT kinase assay for detection of biological activity in natural product extracts. *J Biomol Screening* 9: 52–61

10 Eldridge GR, Vervoort HC, Lee CM, Cremin PA, Williams CT, Hart SM, Goering MG, O'Neil-Johnson M, Zeng L (2002) High-throughput method for the production and analysis of large natural product libraries for drug discovery. *Anal Chem* 74: 3963–3971

11 Abel U, Koch C, Speitling M, Hansske FG (2002) Modern methods to produce natural-product libraries. *Curr Opin Chem Biol* 6: 453–458
12 Bindseil KU, Jakupovic J, Wolf D, Lavayre J, Leboul J, van der Pyl D (2001) Pure compound libraries; a new perspective for natural product based drug discovery. *Drug Dis Today* 6: 840–847
13 Koehn FE, Carter GT (2005) The evolving role of natural products in drug discovery. *Nature Rev Drug Dis* 4: 206–220
14 Barrow RA (2005) Isolation of microbial natural products. *Meth Biotech* 20: 391–414
15 Henry MC, Yonker CR (2006) Supercritical fluid chromatography, pressurized liquid extraction, and supercritical fluid extraction. *Anal Chem* 78: 3909–3915
16 Harrigan GG, Goetz GH (2005) Chemical and biological integrity in natural products screening. *Combinatorial Chem High Throughput Screening* 8: 529–534
17 Yan B, Fang L, Irving M, Zhang S, Boldi AM, Woolard F, Johnson CR, Kshirsagar T, Figliozzi GM, Krueger CA et al (2003) Quality control in combinatorial chemistry: determination of the quantity, purity, and quantitative purity of compounds in combinatorial libraries. *J Combinatorial Chem* 5: 547–559
18 Sadowski J, Kubinyi H (1998) A scoring scheme for discriminating between drugs and nondrugs. *J Med Chem* 41: 3325–3329
19 Henkel T, Brunne RM, Muller H, Reichel F (1999) Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angewandte Chemie, Intl Ed* 38: 643–647
20 Newman DJ, Cragg GM, Kingston DGI (2003) Natural products as pharmaceuticals and sources for lead structures. *Practice Med Chem (2nd Ed)*: 91–109
21 Feher M, Schmidt JM (2003) Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inform Comp Sci* 43: 218–227
22 Lee ML, Schneider G (2001) Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries. *J Combinatorial Chem* 3: 284–289
23 Stahura FL, Godden JW, Xue L, Bajorath J (2000) Distinguishing between natural products and synthetic molecules by descriptor Shannon entropy analysis and binary QSAR calculations. *J Chem Info Comp Sci* 40: 1245–1252
24 Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Del Rev* 46: 3–26
25 Malo N, Hanley JA, Cerquozzi S, Pelletier J, Nadon R (2006) Statistical practice in high-throughput screening data analysis. *Nature Biotech* 24: 167–175
26 Lipinski CA (2001) Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods* 44: 235–249
27 Osada H, Magae J, Watanabe C, Isono K (1988) Rapid screening method for inhibitors of protein kinase C. *J Antibiot*ics 41: 925–931
28 Prudhomme M (2005) Staurosporines and structurally related indolocarbazoles as antitumor agents. *Anticancer Agents Nat Prod* 499–517
29 Kumar RA, Clark DS (2006) High-throughput screening of biocatalytic activity: applications in drug discovery. *Curr Opin Chem Biol* 10: 162–168
30 Heilker R, Zemanova L, Valler MJ, Nienhaus GU (2005) Confocal fluorescence micros-

copy for high-throughput screening of G-protein coupled receptors. *Curr Med Chem* 12: 2551–2559

31 Terstappen GC (2005) Ion channel screening technologies today. *Drug Dis Today: Technologies* 2: 133–140

32 Croston GE (2002) Functional cell-based uHTS in chemical genomic drug discovery. *Trends Biotech* 20: 110–115

33 Zheng X, Hu Y, Liu J, Ouyang K (2005) Screening of active compounds as neuromedin U2 receptor agonist from natural products. *Bioorg Med Chem* Lett 15: 4531–4535

34 Henrich CJ, Bokesch HR, Dean M, Bates SE, Robey RW, Goncharova EI, Wilson JA, McMahon JB (2006) A high-throughput cell-based assay for inhibitors of ABCG2 activity. *J Biomolecular Screening* 11: 176–183

35 Wang J, Soisson SM, Young K, Shoop W, Kodali S, Galgoci A, Painter R, Parthasarathy G, Tang YS, Cummings R et al (2006) Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* 441: 358–361

36 Ondeyka JG, Zink DL, Young K, Painter R, Kodali S, Galgoci A, Collado J, Tormo JR, Basilio A, Vicente F et al (2006) Discovery of bacterial fatty acid synthase inhibitors from a Phoma species as antimicrobial agents using a new antisense-based strategy. *J Nat Prod* 69: 377–380

37 Yoo H-D, Cremin PA, Zeng L, Garo E, Williams CT, Lee CM, Goering MG, O'Neil-Johnson M, Eldridge GR, Hu JF (2005) Suaveolindole, a new mass-limited antibacterial indolosesquiterpene from *Greenwayodendron suaveolens* obtained *via* high-throughput natural products chemistry methods. *J Nat Prod* 68: 122–124

38 Vogt A, Tamewitz A, Skoko J, Sikorski RP, Giuliano KA, Lazo JS (2005) The benzo[c] phenanthridine alkaloid, sanguinarine, is a selective, cell-active inhibitor of mitogen-activated protein kinase phosphatase-1. *J Biol Chem* 280: 19078–19086

39 Summers MY, Leighton M, Liu D, Pong K, Graziani EI (2006) 3-normeridamycin: a potent non-immunosuppressive immunophilin ligand is neuroprotective in dopaminergic neurons. *J Antibiotics* 59: 184–189

40 Marion F, Williams DE, Patrick BO, Hollander I, Mallon R, Kim SC, Roll DM, Feldberg L, Van Soest R, Andersen RJ (2006) Liphagal, a selective inhibitor of pi3 kinase a isolated from the sponge aka coralliphaga: structure elucidation and biomimetic synthesis. *Org Lett* 8: 321–324

41 Strege MA (1999) High-performance liquid chromatographic-electrospray ionization mass spectrometric analyses for the integration of natural products with modern high-throughput screening. *J Chromatography B: Biomedical Sci Applications* 725: 67–78

42 Higgs RE, Zahn JA, Gygi JD, Hilton MD (2001) Rapid method to estimate the presence of secondary metabolites in microbial extracts. *Applied Env Microbiol* 67: 371–376

43 Nielsen N-PV, Carstensen JM, Smedsgaard J (1998) Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimized warping. *J Chromatography, A* 805: 17–35

44 van Nederkassel AM, Daszykowski M, Eilers PHC, Vander Heyden Y (2006) A comparison of three algorithms for chromatograms alignment. *J Chromatography, A* 1118: 199–210

45 Tormo JR, Garcia JB (2005) Automated analyses of HPLC profiles of microbial extracts: a new tool for drug discovery screening. *Nat Prod* 57–75

46 Peake DA, Duckworth DC, Perun TJ, Scott WL, Kulanthaivel P, Strege MA (2005) Analytical and biological evaluation of high throughput screen actives using evaporative

light scattering, chemiluminescent nitrogen detection, and accurate mass LC-MS-MS. *Combinatorial Chem High Throughput Screening* 8: 477–487

47 Ganguly AK, Pramanik BN, Chen G, Shipkova PA (2002) Characterization of pharmaceuticals and natural products by electrospray ionization mass spectrometry. *Practl Spectroscopy* 32: 149–185

48 Fuellbeck M, Michalsky E, Dunkel M, Preissner R (2006) Natural products: sources and databases. *Nat Prod Rep* 23: 347–356

49 Cremin PA, Zeng L (2002) High-throughput analysis of natural product compound libraries by parallel LC-MS evaporative light scattering detection. *Anal Chem* 74: 5492–5500

50 Fredenhagen A, Derrien C, Gassmann E (2005) An MS/MS library on an ion-trap instrument for efficient dereplication of natural products. different fragmentation patterns for [M + H]+ and [M + Na]+ ions. *J Nat Prod* 68: 385–391

51 Larsen TO, Smedsgaard J, Nielsen KF, Hansen ME, Frisvad JC (2005) Phenotypic taxonomy and metabolite profiling in microbial drug discovery. *Nat Prod Rep* 22: 672–695

52 Pauli GF, Jaki BU, Lankin DC (2005) Quantitative 1H NMR: development and potential of a method for natural products analysis. *J Nat Prod* 68: 133–149

53 Bollard ME, Stanley EG, Lindon JC, Nicholson JK, Holmes E (2005) NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. *NMR in Biomedicine* 18: 143–162

54 Sehgal R, Nandave M, Ojha SK (2006) Natural products: leads for drug discovery and development. *Pharma Rev* 4: 40–44

55 Holmes E, Tang H, Wang Y, Seger C (2006) The assessment of plant metabolite profiles by NMR-based methodologies. *Planta Medica* 72: 771–785

56 Pierens GK, Palframan ME, Tranter CJ, Carroll AR, Quinn RJ (2005) A robust clustering approach for NMR spectra of natural product extracts. *Magnetic Resonance Chem* 43: 359–365

57 Jansma A, Chuan T, Albrecht RW, Olson DL, Peck TL, Geierstanger BH (2005) Automated microflow NMR: routine analysis of five-microliter samples. *Anal Chem* 77: 6509–6515

58 Hu JF, Garo E, Yoo HD, Cremin PA, Zeng L, Goering MG, O'Neil-Johnson M, Eldridge GR (2005) Application of capillary-scale NMR for the structure determination of phytochemicals. *Phytochem Anal* 16: 127–133

59 Taggi AE, Meinwald J, Schroeder FC (2004) A new approach to natural products discovery exemplified by the identification of sulfated nucleosides in spider venom. *J Am Chem Soc* 126: 10364–10369

60 Zeng L, Cremin P, Lee C, O'Neil-Johnson M, Caporale L (2001) High throughput natural product chemistry for drug discovery. *Abstracts of Papers, 221st ACS National Meeting, San Diego, CA, United States, April 1–5, 2001*: MEDI–193

61 Schlotterbeck G, Ross A, Hochstrasser R, Senn H, Kuehn T, Marek D, Schett O (2002) High-resolution capillary tube NMR. A miniaturized 5-mL high-sensitivity TXI probe for mass-limited samples, off-line LC NMR, and HT NMR. *Anal Chem* 74: 4464–4471

62 Laude DA Jr, Wilkins CL (1984) Direct-linked analytical scale high-performance liquid chromatography/nuclear magnetic resonance spectrometry. *Anal Chem* 56: 2471–2475

63 Albert K (1999) Liquid chromatography-nuclear magnetic resonance spectroscopy. *J Chromatography, A* 856: 199–211

64 Bringmann G, Wohlfarth M, Rischer H, Schlauer J, Brun R (2002) Extract screening by

HPLC coupled to MS-MS, NMR, and CD: a dimeric and three monomeric naphthylisoquinoline alkaloids from *Ancistrocladus griffithii*. *Phytochem* 61: 195–204

65 Jaroszewski JW (2005) Hyphenated NMR methods in natural products research, part 2: HPLC-SPE-NMR and other new trends in NMR hyphenation. *Planta Medica* 71: 795–802

66 Jaroszewski JW (2005) Hyphenated NMR methods in natural products research, part 1: Direct hyphenation. *Planta Medica* 71: 691–700

67 Exarchou V, Krucker M, van Beek TA, Vervoort J, Gerothanassis IP, Albert K (2005) LC-NMR coupling technology: Recent advancements and applications in natural products analysis. *Magnetic Resonance Chem* 43: 681–687

68 Bieri S, Varesio E, Veuthey J-L, Munoz O, Tseng L-H, Braumann U, Spraul M, Christen P (2006) Identification of isomeric tropane alkaloids from *Schizanthus grahamii* by HPLC-NMR with loop storage and HPLC-UV-MS/SPE-NMR using a cryogenic flow probe. *Phytochem Anal* 17: 78–86

69 Lambert M, Staerk D, Hansen SH, Jaroszewski JW (2005) HPLC-SPE-NMR hyphenation in natural products research: optimization of analysis of Croton membranaceus extract. *Magnetic Resonance Chem* 43: 771–775

70 Smallcombe SH, Patt SL, Keifer PA (1995) WET solvent suppression and its applications to LC NMR and high-resolution NMR spectroscopy. *J Magnetic Resonance, Series A* 117: 295–303

71 Spraul M, Freund AS, Nast RE, Withers RS, Maas WE, Corcoran O (2003) Advancing NMR sensitivity for LC-NMR-MS using a cryoflow probe: application to the analysis of acetaminophen metabolites in urine. *Anal Chem* 75: 1536–1541

72 Putzbach K, Krucker M, Grynbaum MD, Hentschel P, Webb AG, Albert K (2005) Hyphenation of capillary high-performance liquid chromatography to microcoil magnetic resonance spectroscopy-determination of various carotenoids in a small-sized spinach sample. *J Pharma Biomed Anal* 38: 910–917

73 Lewis RJ, Bernstein MA, Duncan SJ, Sleigh CJ (2005) A comparison of capillary-scale LC-NMR with alternative techniques: Spectroscopic and practical considerations. *Magnetic Resonance Chem* 43: 783–789

74 Seger C, Godejohann M, Tseng LH, Spraul M, Girtler A, Sturm S, Stuppner H (2005) LC-DAD-MS/SPE-NMR hyphenation. a tool for the analysis of pharmaceutically used plant extracts: identification of isobaric iridoid glycoside regioisomers from *Harpagophytum procumbens*. *Anal Chem* 77: 878–885

75 Lommen A, Godejohann M, Venema DP, Hollman PCH, Spraul M (2000) Application of directly coupled HPLC-NMR-MS to the identification and confirmation of Quercetin glycosides and Phloretin glycosides in apple peel. *Anal Chem* 72: 1793–1797

76 Exarchou V, Godejohann M, Van Beek TA, Gerothanassis IP, Vervoort J (2003) LC-UV-solid-phase extraction-NMR-MS combined with a cryogenic flow probe and its application to the identification of compounds present in greek oregano. *Anal Chem* 75: 6288–6294

77 Cummins LL, Chen S, Blyn LB, Sannes-Lowery KA, Drader JJ, Griffey RH, Hofstadler SA (2003) Multitarget affinity/specificity screening of natural products: finding and characterizing high-affinity ligands from complex mixtures by using high-performance mass spectrometry. *J Nat Prod* 66: 1186–1190

78 Poulsen SA, Davis RA, Keys TG (2006) Screening a natural product-based combinatorial library using FTICR mass spectrometry. *Bioorg Med Chem* 14: 510–515

79 Cheng X, Van Breemen RB (2005) Mass spectrometry-based screening for inhibitors of amyloid protein aggregation. *Anal Chem* 77: 7012–7015

80 Moy FJ, Haraki K, Mobilio D, Walker G, Powers R, Tabei K, Tong H, Siegel MM (2001) MS/NMR: A structure-based approach for discovering protein ligands and for drug design by coupling size exclusion chromatography, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Anal Chem* 73: 571–581

81 Schriemer DC, Bundle DR, Li L, Hindsgaul O (1999) Micro-scale frontal affinity chromatography with mass spectrometric detection: a new method for the screening of compound libraries. *Angewandte Chemie, Intl Ed* 37: 3383–3387

82 Chan NWC, Lewis DF, Rosner PJ, Kelly MA, Schriemer DC (2003) Frontal affinity chromatography-mass spectrometry assay technology for multiple stages of drug discovery: applications of a chromatographic biosensor. *Anal Biochem* 319: 1–12

83 Slon-Usakiewicz JJ, Ng W, Dai JR, Pasternak A, Redden PR (2005) Frontal affinity chromatography with MS detection (FAC-MS) in drug discovery. *Drug Dis Today* 10: 409–416

84 Zhu L, Chen L, Luo H, Xu X (2003) Frontal affinity chromatography combined on-line with mass spectrometry: A tool for the binding study of different epidermal growth factor receptor inhibitors. *Anal Chem* 75: 6388–6393

85 Fejzo J, Lepre C, Xie X (2003) Application of NMR screening in drug discovery. *Curr Topics Med Chem* 3: 81–97

86 Carlomagno T (2005) Ligand-target interactions: what can we learn from NMR? *Ann Rev Biophys Biomol Structure* 34: 245–266, 242 plates

87 Marchioro C, Davalli S, Provera S, Heller M, Ross A, Senn H (2003) Experiments in NMR-based screening. *Methods Principles Med Chem* 16: 321–340

88 Shuker SB, Hajduk PJ, Meadows RP, Fesik SW (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274: 1531–1534

89 Hajduk PJ, Augeri DJ, Mack J, Mendoza R, Yang J, Betz SF, Fesik SW (2000) NMR-based screening of proteins containing 13C-labeled methyl groups. *J Am Chem Soc* 122: 7898–7904

90 Hajduk PJ, Gerfin T, Boehlen J-M, Haeberli M, Marek D, Fesik SW (1999) High-throughput nuclear magnetic resonance-based screening. *J Med Chem* 42: 2315–2317

91 Shapiro MJ, Wareing JR (1999) High resolution NMR for screening ligand/protein binding. *Curr Opin Drug Dis Dev* 2: 396–400

92 Rees DC, Congreve M, Murray CW, Carr R (2004) Fragment-based lead discovery. *Nature Rev Drug Dis 3*: 660–672

93 Mercier KA, Powers R (2005) Determining the optimal size of small molecule mixtures for high throughput NMR screening. *J Biomolecular NMR* 31: 243–258

94 Mayer M, Meyer B (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie, Intl Ed* 38: 1784–1788

95 Milton MJ, Williamson RT, Koehn FE (2006) Mapping the bound conformation and protein interactions of microtubule destabilizing peptides by STD-NMR spectroscopy. *Bioorg Med Chem Lett* 16: 4279–4282

96 Carlomagno T, Blommers Marcel JJ, Meiler J, Jahnke W, Schupp T, Petersen F, Schinzer D, Altmann KH, Griesinger C (2003) The high-resolution solution structure of epothilone A bound to tubulin: an understanding of the structure-activity relationships for a powerful class of antitumor agents. *Angewandte Chemie, Intl Ed* 42: 2511–2515

97 Politi M, Chavez MI, Canada FJ, Jimenez-Barbero J (2005) Screening by NMR: A new

approach for the study of bioactive natural products? The example of *Pleurotus ostreatus* hot water extract. *Eur J Org Chem* 7: 1392–1396

98 Chen A, Shapiro MJ (2000) NOE pumping as high-throughput method to determine compounds with binding affinity to macromolecules by NMR. *J Am Chem Soc* 122: 414–415

99 Dalvit C, Pevarello P, Tato M, Veronesi M, Vulpetti A, Sundstrom M (2000) Identification of compounds with binding affinity to proteins *via* magnetization transfer from bulk water. *J Biomolecular NMR* 18: 65–68

100 Johnson EC, Feher VA, Peng JW, Moore JM, Williamson JR (2003) Application of NMR SHAPES screening to an RNA target. *J Am Chem Soc* 125: 15724–15725

101 Cohen Y, Avram L, Frish L (2005) Diffusion NMR spectroscopy in supramolecular and combinatorial chemistry: an old parameter – new insights. *Angewandte Chemie, Intl Ed* 44: 520–554

102 Bleicher K, Lin M, Shapiro MJ, Wareing JR (1998) Diffusion edited NMR: screening compound mixtures by affinity NMR to detect binding ligands to vancomycin. *J Org Chem* 63: 8486–8490

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Virtual screening for the discovery of bioactive natural products

By Judith M. Rollinger[1], Hermann Stuppner[1,2] and Thierry Langer[2,3]

[1]Institute of Pharmacy/Pharmacognosy, Leopold-Franzens University of Innsbruck, Innrain 52c, 6020 Innsbruck, Austria

[2]Inte:Ligand GmbH, Software Engineering and Consulting, Clemens Maria Hofbauergasse 6, 2344 Maria Enzersdorf, Austria

[3]Institute of Pharmacy/Pharmaceutical Chemistry/Computer Aided Molecular Design Group, Leopold-Franzens University of Innsbruck, Innrain 52c, 6020 Innsbruck, Austria
<Thierry.Langer@uibk.ac.at>

Abstract

In this survey the impact of the virtual screening concept is discussed in the field of drug discovery from nature. Confronted by a steadily increasing number of secondary metabolites and a growing number of molecular targets relevant in the therapy of human disorders, the huge amount of information needs to be handled. Virtual screening filtering experiments already showed great promise for dealing with large libraries of potential bioactive molecules. It can be utilized for browsing databases for molecules fitting either an established pharmacophore model or a three dimensional (3D) structure of a macromolecular target. However, for the discovery of natural lead candidates the application of this *in silico* tool has so far almost been neglected. There are several reasons for that. One concerns the scarce availability of natural product (NP) 3D databases in contrast to synthetic libraries; another reason is the problematic compatibility of NPs with modern robotized high throughput screening (HTS) technologies. Further arguments deal with the incalculable availability of pure natural compounds and their often too complex chemistry. Thus research in this field is time-consuming, highly complex, expensive and ineffective. Nevertheless, naturally derived compounds are among the most favorable source of drug candidates. A more rational and economic search for new lead structures from nature must therefore be a priority in order to overcome these problems.

Here we demonstrate some basic principles, requirements and limitations of virtual screening strategies and support their applicability in NP research with already performed studies. A sensible exploitation of the molecular diversity of secondary metabolites however asks for virtual screening concepts that are interfaced with well-established strategies from classical pharmacognosy that are used in an effort to maximize their efficacy in drug discovery. Such integrated virtual screening workflows are outlined here and shall help to motivate NP researchers to dare a step towards this powerful *in silico* tool.

1 Introduction

In the field of drug discovery we are confronted by a paradox situation: highly efficient tools and advanced technological and molecular know-how, e.g., in the area of genomics, combinatorial chemistry, high throughput screening (HTS), robotized and miniaturized process cycles, could find entrance in big pharmaceutical industries. These costly procedures were expected to raise the number of launched drug substances; however the results were disappointing [1, 2]. In 2002, Adam Smith, the chief-editor of *Nature* presented the sobering data of research and development expenses of the 20 leading pharma companies *versus* new drugs on the market. They have steadily fallen in recent years despite the increasing financial efforts [3].

On the other side we are faced by a high traditional impact of naturally derived medicines and incredible success stories of natural products (NPs)

as potent remedies from the beginnings of human therapeutic activity to modern research and drug development. Nevertheless, most large pharmaceutical companies scaled down or terminated their work in NPs operations. The reasons behind this are that the drug discovery process starting from natural sources is hardly compatible with the today's highly automated drug discovery technologies. Thus, the pre-eminence of combinatorial chemistry as the preferred method for generating new drug leads has led to the comparative neglect of this valuable resource. William Strohl from Merck Research Laboratories summarized the difficulties of NP programs *versus* synthetic chemicals in his editorial remarks in *Drug Discovery Today* [4]. These include (i) the existence of already found potent antimicrobic and antitumor NPs and the lack of sufficient dereplication programs which prevent their repeated discovery; (ii) the fact that – in contrast to the highly sophisticated molecular targets – NP extracts are generally regarded as too 'dirty', too difficult to assay and too time-consuming; (iii) obtaining an assay hit resulting from a bio-guided fractionation, the NPs' structure still has to be elucidated compared with synthetic chemicals; (iv) NPs are often deemed as too structurally complex, possessing multiple hydroxyl moieties, ketones and chiral centers. Strohl nevertheless concluded by listing a number of advantages applying an active NP program, which he finally described as an '*expensive endeavor*' which, however, is '*well worth the cost*'.

The use of NPs has been the single-most successful strategy for the discovery of new drug leads, which is clearly shown by different statistics [5, 6]. With increased calls in recent years for further research on NPs [7, 8] there are again signs that they may play a more active role in the future drug discovery process, since their reintroducing may help to *re-discover the sweet spot in drug discovery* [1].

2 Status of NPs

To date some 200,000 natural compounds [9–11] have been published. The terrestrial flora has been intensively investigated over the last decades; the potential in finding new NPs slumbering in untapped biota is however nearly inconceivable. It is estimated that only 5–15% of the approximately 250,000 described high plant species have ever been in the focus of phytochemical and pharmacological investigations [12]. More sobering is the

percentage in the field of bacterial (less than 1%) and fungal species (less than 5%) [13].

The main part of known NPs belongs to secondary metabolites. These compounds provide living systems with their characteristic features mandatory for surviving. They contain an inherently large-scale of structural diversity. About 40% of the chemical scaffolds of published NPs are unique and have not been made by any chemist [14].

In the past 100 years researchers have discovered many potential therapeutic targets. Since the completion of the human genome, 30,000 to 40,000 genes and at least the same number of proteins are assumed [15]. Thus, we are up against an increasing number of macromolecular targets, like proteins, receptors, enzymes, and ion channels – that might be of pathological concern for humankind. Among them, proteins continue to attract significant attention from pharmaceutical technology as a valuable source of drugable targets [16]. Proteins provide the critical link between genes and disease, and thus are the key to understanding the basic biological processes. Up to now drug discovery has been performed against only approximately 500 targets [17], though the number of potential targets are estimated to be in the range of 2,000 to 5,000 [2, 15].

Taken together, it can be assumed that a large number of drug leads and hits are conserved in the inexhaustible pool of NPs pre-screened by evolution. But how to dig out and to recognize the respective drug leads is a challenging task for both industry and academia, for medicinal chemists, pharmacognosists and pharmacologists. NP research is affected with a wealth of time-consuming and cost intensive investigations. Collection of the natural material, phytochemical analysis, isolation and identification of the constituents is just the basic procedure. A biological screening of extracts or even the arbitrary testing of isolated metabolites is feasible and often performed, though is not at all a focused procedure, thus unpractical and too expensive. The NPs' diversity has to be accessed in a more rational way.

3 Holistic *versus* molecular approaches in drug discovery from nature

During the last century and even today the discovery of bioactive NPs and their development into potential drug candidates are mainly covered by

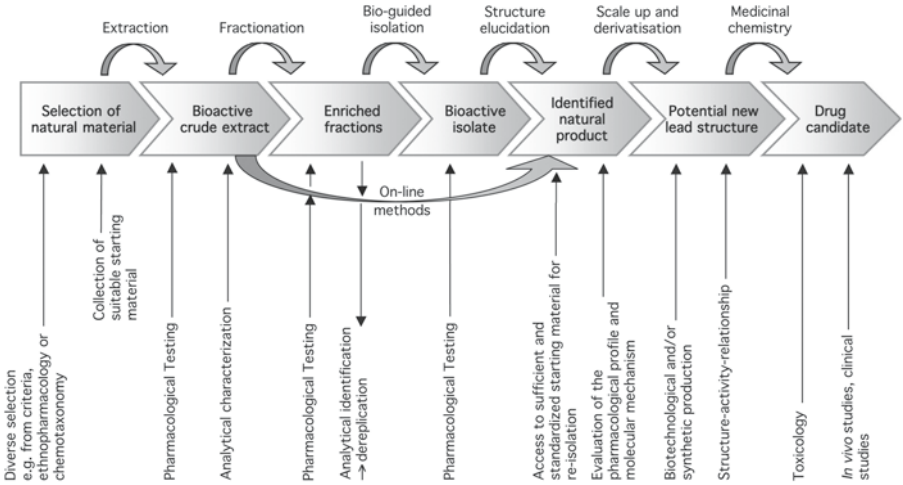


Figure 1.
Traditional early drug development of a nature based drug candidate

a holistic approach. A characteristic workflow of this procedure is given in Figure 1. Starting from the knowledge or assumption about a biological effect the natural material is selected and adequately extracted. If a positive effect in the obtained multi-component extract is recorded, it is attempted to trace it back to the active principle/s by intense phytochemical and analytical investigations (Fig. 1). This can for instance be achieved by a bioactivity-guided fractionation. A more targeted approach focuses on innovative technological tools combining analytical and biological information. An overview of recent developments in this area and successful examples thereof are presented by Potterat and Hamburger [18, 19].

As soon as the constituent regarded to be responsible for the overall effect is isolated, further research focuses on a molecular level including structure elucidation and pharmacological profiling. Synthesis and testing of series of derivatives enable an insight into a structure-activity-relationship and pharmacokinetic aspects. Finally, potential drug leads become drug candidates after some intense toxicological studies and after the verified effectiveness *in vivo* (Fig. 1).

Recent advances in lead identification from nature work on a molecular base more than on a holistic one. A first prerequisite for that is on bioin-

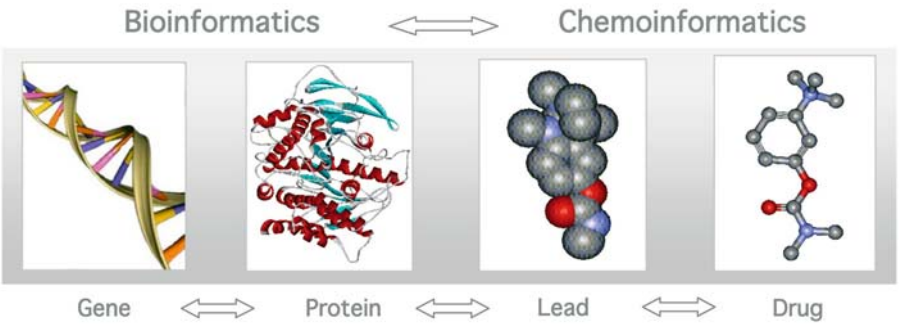


Figure 2.
Schematic correlation between bioinformatics and chemoinformatics

formatics comprising 3D structures from genes and proteins (bioinformatics), substantial knowledge on molecular target functions with accurate structural information and protein–ligand interactions. Secondly, it is essential to refer to unambiguously characterized structures of secondary metabolites preferably with some information to their biological effect. Based on available structural as well as biological knowledge from both sides, information can be deduced from chemoinformatics to bridge the gap between known ligands and the discovery of new lead structures (Fig. 2).

4 Computational approaches for the discovery of lead structures from nature

The increasing understanding of fundamental principles of protein–ligand interactions and the steadily growing number of 3D-structures of potential and experimentally proved ligands provide undreamed of possibilities towards more rationalized concepts in drug discovery. However, too much is expected of the human brain to profit from the already published information. Thus, efficient and effective approaches benefit from today's knowledge about NPs. In the area of medicinal chemistry, computational methods, like virtual screening experiments, have already proved to satisfy these requirements. They are needed to exploit the available structural

information, to understand specific molecular recognition events, and to clarify the function of the target macromolecule. Though rationalized procedures in the search for bioactive natural products are in great demand to find the 'needles in a haystack', computational assistance could hardly break into natural product research.

The common idea of all computational approaches within the early drug discovery process is to mine more or less large compound databases *in silico* and to select a limited number of candidates proposed to have the desired biological activity. For this process the term 'data mining' was coined in 1996 [20], which was concisely defined by Gasteiger and co-authors: '*to extract knowledge from a large set of data in order to make predictions of new events*' [21].

Within the lead discovery process, virtual screening technologies have largely enhanced the impact of computational chemistry and nowadays chemoinformatics plays a predominant role in early phase drug research [22, 23]. The key goal of the use of such methods is to reduce the overall cost associated to the discovery and development of a new drug, by identifying the most promising candidates to focus the experimental efforts on. Recently published books and reviews on the impact of computational chemistry for lead structure determination highlight these efforts [24–27].

If the 3D structure of the biological target is known, high throughput docking turned out to be a valuable structure-based virtual screening method to be used [28–31]. Within this context, the scoring of hits retrieved still remains a question that is often discussed. In fact, currently the major weakness of docking programs lies not in the docking algorithms themselves but still in the inaccuracy of the functions that are used to estimate the affinity between ligand and target, the so-called scoring functions. Previously, Stahl and Rarey analyzed scoring functions for virtual screening [32], giving valuable insight into strengths and weaknesses of currently used models for affinity estimation. The combination of several different scoring functions termed as consensus scoring turns out to be one of the possible answers to the question raised previously. In fact, several authors recently described their efforts in this area; an example is given in reference [33]. In a theoretical study, other authors demonstrate that consensus scoring outperforms any single scoring for simple statistical reasons and that a moderate number of scoring functions (i.e., three or

four) are sufficient for the purpose of consensus scoring [34]. However, it has been shown that consensus scoring alone is not suitable for all cases of docking, and, as highlighted in a recent review by Krovat and co-authors, considerable efforts are still devoted to the optimization of scoring functions [28].

Because of the restricted free access to NP 3D libraries (see below), the number of virtual screening studies published for the rational access to bioactive NPs is limited. Some examples using high throughput docking as a structure-based virtual screening tool will be given here: Liu and Zhou applied a theoretical approach to find natural ligands as potential inhibitors of the SARS-CoV protease, a virus target of the severe acute respiratory syndrome [35]. They used a docking-based virtual screening cycle and applied drug-like filters to finally propose 18 drug candidates out of two 3D databases comprising metabolites from marine organism and compounds from traditional Chinese Medicine. The same virus organism was the main interest in the study performed by Toney et al., who focused on its main proteinase, 3CLpro. The crystal structure of this attractive target was used as the starting point for the virtual docking screening of the NCI database. Searching for non-peptidyl inhibitors, the authors identified the naturally occurring terpenoid alkaloid sabadinine (i.e., cevine; **1**) as potential anti-SARS agent [36].

The author group around Stefano Moro could identify ellagic acid (**2**) as inhibitor of the protein kinase CK2 screening an in-house generated database with almost 2,000 structures of natural compounds [37]. A combination of four docking protocols and five scoring functions has been utilized to dock and rank the molecules in the database. The consensus docking suggested ellagic acid to be one of the most promising candidates. This assumption could be verified by experimental studies revealing this NP as highly potent CK2 inhibitor (K_i = 20 nM).

Estrogen receptor-β plays a key role in regulating brain development and estrogen-induced promotion of neurogenesis and memory. Using the 3D coordinates of the co-crystal structure of human estrogen receptor-β bound with genistein as starting point, Zhao and Brinton pursued a receptor-based molecular docking approach [38]. They focused on the search for natural estrogen receptor-β-selective ligands. Twelve candidate molecules, which had been suggested by the database screening, were selected. The authors determined their binding affinity and selectivity; three of the com-

pounds belonging to the flavanoid family (**3–5**) displayed over 100-fold binding selectivity to the estrogen receptor-β over α. A similar approach was employed by Liu and co-authors. Applying a docking virtual screening filtering experiment, the authors discovered potent inhibitors of the potassium ion channel from a Chinese NP database [39].

5 Pharmacophore concept in NP research

The pharmacophore concept has proven to be extremely successful, not only in rationalizing structure-activity relationships, but also by its large impact in developing the appropriate 3D-tools for efficient virtual screening [40]. Profiling of combinatorial libraries and compound classification are other often-used applications of this concept. Although well established in combinatorial chemistry, it has to be pointed out that the tools described in this section have likewise a considerable impact on the rational finding of new potential lead compounds originating from the immense source of secondary metabolites. The prior use of pharmacophore models in biological screening of NPs is an efficient procedure since it quickly eliminates molecules that do not possess the required features thus leading to a dramatic increase of enrichment, when compared to a purely random screening experiment. In a previous study conducted by Doman and co-authors [41], only 85 molecules or 0.021% revealed as protein tyrosine phosphatase-1B inhibitors ($IC_{50} < 100$ μM) by a HTS of approximately 400,000 compounds. On the other hand, of 365 molecules suggested by molecular docking, 127 or 34.8% were found to be active. Thus, docking-based virtual screening enriched the hit rate by almost 1,700-fold over random screening.

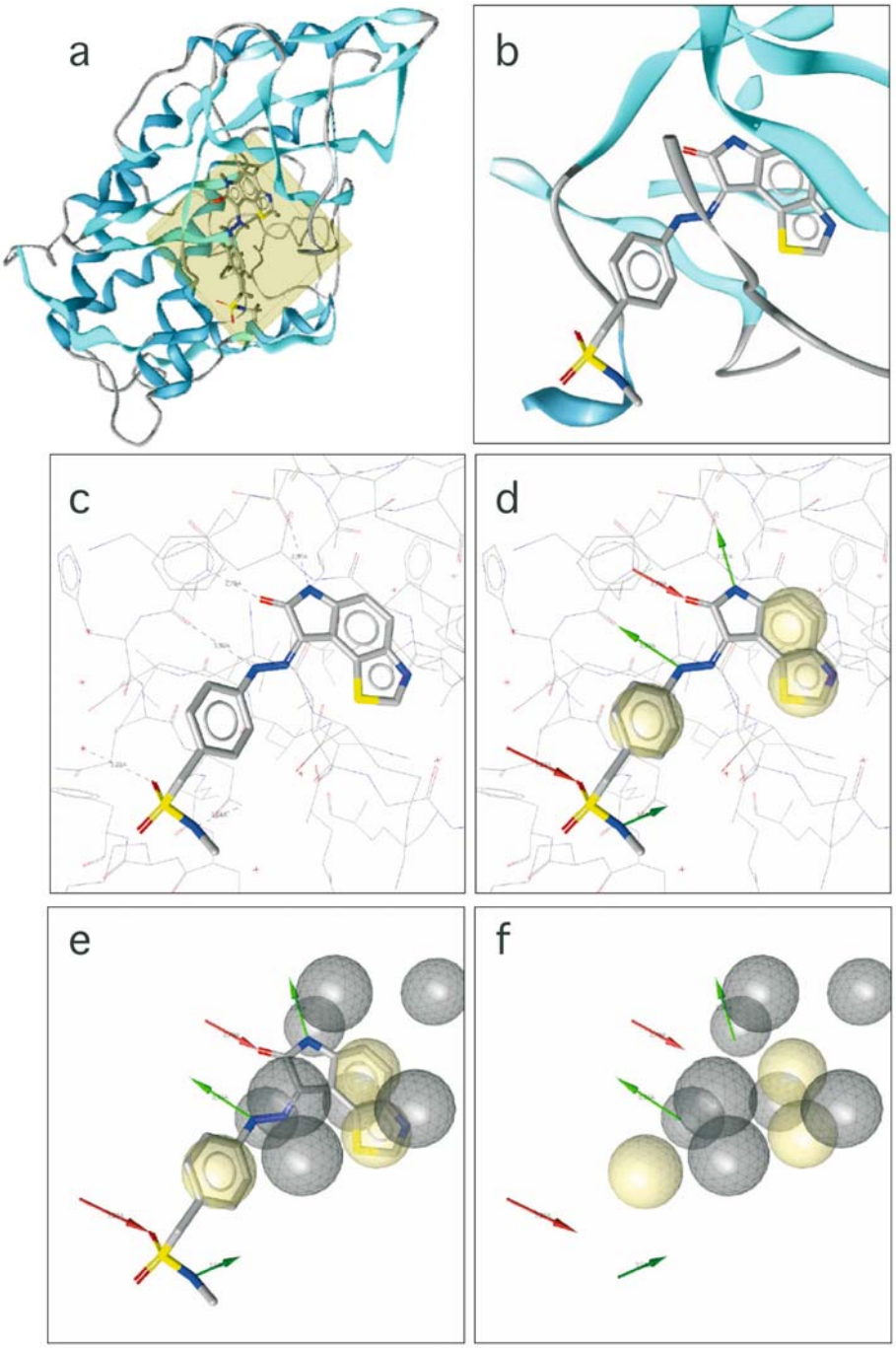
One should not forget, however, that additional molecular characteristics not reflected by pharmacophore models (physicochemical properties relevant for ADME and toxicological properties) must be taken into account when deciding upon which compounds should be further developed [42]. A rapid identification and elimination of compounds with unsuitable physicochemical and pharmacokinetic properties is a pivotal step in the early drug discovery process [43, 44]. They can be evaluated traditionally or by high throughput screening, which are discussed in detail by Avdeef and Testa [45]. This must be considered for synthetics

as well as NPs, though studies revealed secondary metabolites not only high scaffold diversity; biosynthesized molecules also show structural and spatial characteristics that are closer to drug leads than those of synthetic molecules [46, 47]. Typically, NPs include more chiral centers and their stereochemical architecture is much more complex than that of synthetic molecules. Furthermore, they usually contain more carbons, hydrogen and oxygen, however, less nitrogen and other atoms compared to synthetics. Surprisingly, NPs often show a molecular weight higher than 500 Da combined with a high polarity [7], which is in clear contrast to Lipinski's rule of five [48]. Nevertheless only about 10% of NPs contain two or more violations of Lipinski's rules [47]. In summary, natural chemistry can be seen as highly diverse scaffolds endowed with potential drugable pharmacophores.

6 Structure-based pharmacophore model

An inevitable prerequisite for generating a structure-based model is the knowledge about the ligand-target interaction [49] including the availability of the 3D structure of the target either by X-ray crystallography or NMR or constructed on the basis of the structure of homologous proteins. A unique platform containing 3D coordinates of experimentally solved protein structures is the Brookhaven Protein Data Bank (PDB [50]). A crystalline complex with a ligand bound to a protein's active site is the best requirement to start the construction of a structure-based 3D model. In this case, one may profit from the exact information of the ligand's bioactive conformation which is preserved in the binding site of the crystalline complex. The building of a structure-based pharmacophore is depicted in a step by step way in Figure 3.

A new software tool has recently been described for the successful generation of such chemical features-based models: The software LIGANDSCOUT [51] is a program for ligand interpretation and data mining in the PDB. The performance of this program allows the detection of relevant interaction points between ligand and protein. The binding mode of the ligand in the active site of a protein can be visualized in a sophisticated way. LIGANDSCOUT's algorithms perform a stepwise interpretation of the ligand molecules: Planar ring detection, assignment of functional


a
b
c
d
e
f

group patterns, determination of the hybridization state and finally the assignment of Kekulé pattern. The interpretation of the ligand molecules is the basis for the next step, an automated generation of pharmacophore models, derived from the data provided by a crystalline complex of the PDB. An automatic detection and classification of protein–ligand interactions into hydrogen bonds, charge transfer, and lipophilic regions leads to a collection of chemical features in a pharmacophore model. The graphical user-interface can provide an integrated view of protein, ligand, pharmacophore model, and interaction lines. In a previously published study, LIGANDSCOUT was used for the detection and interpretation of crucial interaction patterns between ligands and the factor Xa protein structure [52]. In a second step, the program CATALYST, a state of the art virtual screening platform, was used for rapid virtual screening of multi-conformational 3D structure databases. The information for the pharmacophore pattern (i.e., 3D coordinates of interaction points) was obtained by the interpretation of LIGANDSCOUT pharmacophore definitions and resulted in specific interaction models that were able to map the ligand in their bioactive conformation and to retrieve selectively a 78% fraction of the known factor Xa inhibitors within a small subset of the large Derwent World Drug Index library. A further application of the LIGANDSCOUT pharmacophore definitions covers the rationalized search for angiotensin converting enzyme (ACE)-2 inhibitors by virtual screening of approximately 3.8 million compounds from various commercial databases [53].

Figure 3
Concept for generating a structure-based pharmacophore model; Visualizing and calculation of chemical features using LigandScout [51]:

a. Protein (e.g., CDK2) complexed with a ligand {shown in ball-and-stick mode; N-methyl-{4-[2-(7-oxo-6,7-dihydro-8H-[1,3]thiazolo[5,4-E]indol-8-ylidene)-hydrazino]-phenyl}-methane-sulfonamide} in the active binding site (highlighted in the yellow cube)
b. Zoom up of the binding site with the ligand
c. Ligand with calculated distances to the interacting amino acid residues of the protein
d. Determination of interactions between the ligand and the target; evaluation and setting of chemical features (yellow sphere, hydrophobic feature; green arrow, hydrogen bond donor function; red arrow, hydrogen bond acceptor function)
e. Subtraction of the protein; the ligand, the chemical features and exclusion volumes (= grey spheres; representing areas not to be occupied by the ligand) are left
f. Subtraction of the ligand; the pharmacophore model remains comprising chemical features and exclusion volumes

Hit reduction and selection was achieved using a five feature hypothesis based on a recently resolved inhibitor-bound ACE2 crystal structure. Seventeen virtual hits were selected for their experimental validation in a bioassay; the concept was confirmed since all of them were revealed as ACE-2 inhibitors.

Barreca and co-authors developed a 3D structure-based pharmacophore model with LIGANDSCOUT for the discovery of new scaffolds acting as HIV-1 non-nucleoside reverse transcriptase inhibitors by virtual screening of large chemical databases. Six virtual hits were finally selected for determination of their inhibitory effects. Those belonging to the new scaffold class of the quinolin-2(1H)-one family exhibited reverse transcriptase inhibitory activity at sub-micromolar concentrations [54].

In a recently published work, Schuster et al. presented a so-called cytochrome P450 profiler [55]. Several structure-based (generated with LIGANDSCOUT) and ligand-based pharmacophore models (using CATALYST) for substrates and inhibitors of five cytochrome P450 isoenzymes (1A2, P450 2C9, P450 2C19, P450 2D6, and P450 3A4) were created and validated by the authors' group. Their results showed that the models were suitable for fast pharmacokinetic profiling of large drug-like databases.

In this context the **parallel screening** is of particular interest. Whereas in usual virtual screening cycles interactions of *thousands* or even *millions* of 3D database entries are browsed against *one* pharmacophore model, it is contrary in the case of parallel screening; low-energetic conformers of *one* structure are screened for their potential interactions against *numerous* models. The basics of parallel screening have just recently been presented by Steindl and co-authors [56, 57]. Furthermore, the authors exemplified this strategy for the activity profiling using a set of HIV protease pharmacophore models [58]. This *in silico* concept is of particular interest to virtually scrutinize drug candidates for their preliminary activity profiling relevant to putative side effects and toxicity [40]. According to the obtained interactions to virtually screened antitargets (e.g., hERG, sigma-1, sigma-2, alpha-1A, alpha-1B, alpha-1D, alpha-2A, alpha-2B, alpha-2C, D2L, D3, D4.2, 5-HT_{1A}, 5-HT_{2A}, 5-HT_7, H1, I2, A2A, A2B, cytochrome P 450) a first insight to potentially risky affinities is provided before time and cost intensive toxicological studies are performed.

The virtual screening approach using a structure-based pharmacophore model has revealed some first application examples in NP research: Niko-

lovska-Coleska and co-authors successfully pursued this *in silico* strategy in the area of X-linked inhibitors of apoptosis (XIAP) [59]. A high resolution 3D structure of the XIAP BIR3 domain complexed with the N-terminal end of the Smac/Diablo protein [60], which is an endogenous ligand of the respective XIAP binding pocket, was used as the starting point to virtually screen an in-house 3D-NP database. Embelin (**6**) from the Japanese Ardisia herb emerged as virtual small molecule weight hit, which was found to be a fairly potent inhibitor of XIAP using a fluorescence polarization binding assay.

In our group, we previously focused on acetylcholinesterase (AChE) [61]; according to the cholinergic hypothesis of the pathogenesis of Alzheimer's disease, inhibitors of the AChE are successfully used as therapeutic strategy. Based on the co-crystal structure of AChE with its ligand galanthamine, a structure-based pharmacophore model was generated and used for an *in silico* screening of a multi-conformational database consisting of more than 110,000 NPs. From the obtained hit list, promising, virtually active candidates were selected, namely scopoletin (**7**) and its glucoside scopolin (**8**). Their AChE inhibitory effect was first verified from the crude extract of *Scopolia carniolica* roots using a bioautographic TLC assay. The isolated coumarins showed a significant and dose-dependent inhibition of the AChE in the microplate enzyme assay as well as in the *in vivo* test. The i.c.v. application of both coumarins on rats resulted in a long-lasting, pronounced and – in case of the glucoside – even in a two-fold higher increase of the neurotransmitter's concentration than the one caused by the positive control galanthamine.

7 Ligand-based pharmacophore model

Very often, however, lead discovery projects have reached a well-advanced stage before detailed structural data on the protein target has become available, even though it is well recognized that modern methods of molecular biology together with biophysics and computational approaches enhance the likelihood of successfully obtaining detailed atomic structure information. A possible consequence is that often scientists identify and develop novel compounds for a target using preliminary structure-activity information, together with theoretical models of interaction. Only responses that

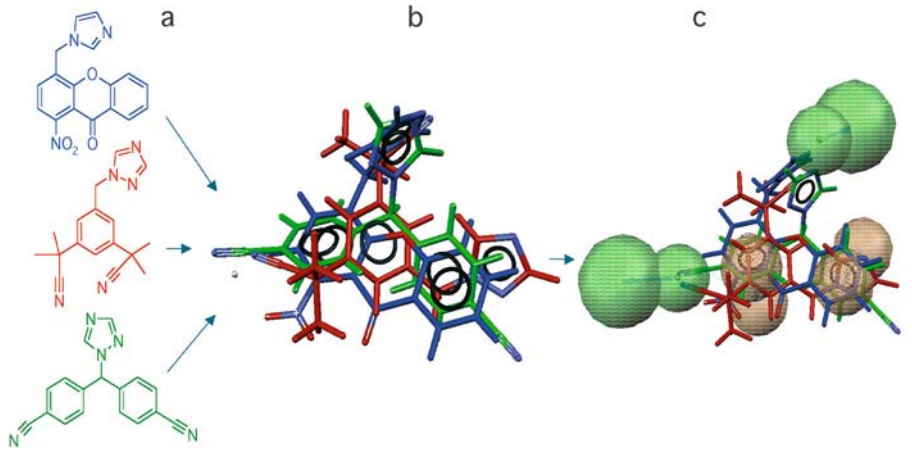


Figure 4.
Concept for generating a ligand-based pharmacophore model using the Catalyst program (Accelrys Inc., CA)
a. critical selection of active ligands
b. alignment of low-energetic conformers of the selected ligands
c. derivation and determination of common features

are consistent with the working hypotheses contribute to an evolution of the used models. Within this framework, the chemical feature-based pharmacophore approach has proven to be successful [62] allowing the perception and understanding of key interactions between a receptor and a ligand on a generalized level. A function-based pharmacophore represents the common ensemble of steric and electrostatic features of different compounds which are necessary for their interaction with a specific biological target structure (Fig. 4).

Such pharmacophore models together with large 3D structure databases originating either from in-house compound collections, from commercial vendors, or from natural products databases have proven to be extremely useful *in silico* screening experiments. When using ligand-based pharmacophore models as screening filters instead of protein 3D structures, affinity estimation is only based on geometric fit of compound atoms or groups to features of the model. In these cases, the values calculated are often far away from reality, however, still are useful for filtering possible

hits from non-binding molecules. Additionally, in pharmacophore fitting procedures, calculation demands are considerably lower than in docking algorithms allowing the number of compounds to be processed in the same time to be by far higher than even in high throughput docking.

Since in most of the studies no experimental information on either the biological conformation of the ligand or the target protein are currently available, the ligand-based chemical feature pharmacophore approach can provide essential information for medicinal chemists. Several successful applications within this subject have been performed using the CATALYST program, one of the leading software packages in chemical feature-based pharmacophore modeling. Schuster and co-workers succeeded in the identification of 11β-hydroxysteroid dehydrogenase type 1 inhibitors applying a common feature-based pharmacophore model for their virtual screening filtering experiments [63]. Similarly, the authors preceded by suggesting compounds with a proposed inhibition to the cytochrome P450 19 isoenzyme [64]. Several reviews covering successful applications of such feature-based methods have been published by Kurogi et al. [65], by Krovat et al. [28] and by Güner et al. [66]. They outline the theoretical background and describe several significant studies including 3D database search strategies.

In the field of NPs only a very limited number of studies report from the rationalized access to bioactive compounds *via* ligand-based virtual screening. For example, this method was pursued for the discovery of inhibitors of the COP9 signalosome (CNS) associated kinases CK2 and PKD [67]. Using NPs curcumin and emodin as lead structures, a virtual screening of an in-house database was carried out. Among the virtual hits seven NPs, e.g., anthraquinone (**9**) and piceatannol (**10**), were found to significantly induce apoptosis by inhibition of the CSN-associated kinases using *in vitro* and cell culture experiments. A further study has demonstrated the power of the ligand-based approach applied to pharmacophore modeling of sigma-1 ligands [68]. Therein, some reliable pharmacophore models could be extracted solely from ligand structure information. Compounds with potent affinities to the sigma-1 receptor known from literature were structurally aligned to derive distinct common features. Their 3D arrangement in combination with a spatial restriction was then used for the generation of a pharmacophore model, which was able to retrieve compounds with high affinity values, among them also NPs, like solanidine (**11**).

8 Discriminant analysis

Further ligand-based approaches use various forms of discriminant analysis, e.g., artificial **neural network** simulations. They are based on collections of mathematical models that are interconnected and organized in different layers. They are analogous to an adaptive human learning process and usually trained with learning sets applying one or more molecular descriptors in order to form clusters that enable to distinguish between different objects and their properties. The resulting models are then applied to make predictions on test sets, until the validated models may be used to derive a QSAR of chemically related structures or to mine larger datasets. One may distinguish between supervised and unsupervised (e.g., Kohonen network) learning methods as discussed in detail by Zupan and Gasteiger [69]. A successful application example within the field of NPs was published by Wagner et al. [70]. The authors used a dataset of 103 structurally diverse sesquiterpene lactones with known NF-κB inhibitory activity to derive a QSAR. By the application of multiple 3D structure representations as descriptors, a single model was achieved which provided detailed information on the structural influence of the investigated biological activity. Sangma and co-authors pursued a combination of two approaches to predict new inhibitors of the HIV-1 RT and HIV-1 PR from a NP database comprising metabolites from Thai medicinal plants. After a high throughput docking of the molecules into the target enzymes, self-organizing maps were generated to reduce the number of promising candidates to be tested [71].

A set of different *in silico* methodologies was previously applied by Cherkasov and co-authors to aid in the discovery of natural non-steroidal ligands for human sex hormone binding globulin [72]. Therein, a rigorously cross-validated neural network based QSAR model identified 105 prospective compounds from a structure collection of 23,836 commercial natural substances. This stringent QSAR ranking was combined with docking studies and pharmacophore-aided database search. The integrated computational methods resulted in a convincing predictive tool which identified a set of 29 structurally diverse NPs, of which every fourth compound was able to inhibit the target protein in a micromolar range.

Compounds of arbitrary structural diversity and with known activity against a target are particularly suitable not only for generating a ligand-

based pharmacophore model (as described before), but also for structure similarity studies using a **decision tree**. The object is to find as good a distinction as possible on the basis of a set of molecular descriptors, which identify molecular features shared by different subsets of active compounds and accordingly filter out compounds within the dataset in which these combinations are lacking. Using not only a simple logical description of one model, but an ensemble of decision trees tend to be the preferred option, since the consensus voting among trees give the approach higher predictive accuracy. One form of multiple decision trees well performed to virtually screen large 3D databases is Random Forest [73]. This chemoinformatic method was recently applied in a theoretical work performed by Ehrman and co-workers to predict ligands of multiple targets, like cyclooxygenase (COX), lipoxygenase (LOX), aldose reductase, HIV-1 enzymes etc., from a large dataset of Chinese herbs [74].

9 Databases

The advent of structure databases has provided a basis for the development and feasibility of automatic methods for the search of new lead structures. Conceptually, all the virtual screening concepts presented above have their origins in synthetic chemistry. Their application, however, is just as well adaptable to NPs' chemistry. Prior to the *in silico* filtering experiment, a 3D structure database requires an efficient generation of reasonable, energetically minimized conformations assumed to meet approximately those conformations that might be of biological relevance [75]. The underlying algorithms for 3D structure generation and conformation analysis are implemented in commercial software tools, e.g., in CORINA [21] or the CATALYST program (CATALYST, available from Accelrys Inc., San Diego, CA, USA; www.accelrys.com).

In the field of NPs the virtual screening application is mainly restricted due to the lack of searchable resources for structurally well defined natural compounds. In general, molecular databases with free access on the internet may comprise a high number of molecules, e.g., ChemBank (>1,100,000, http://chembank.broad.harvard.edu) or PubChem (>5,000,000; http://pubchem.ncbi.nlm.nih.gov); however, information about the number of contained natural molecules is rarely available. The library of the National

Cancer Institute (NCI) stores structural information of more than half a million compounds from both synthetic and natural origin that have been collected and tested by the NCI since 1955. About half of the synthetic compounds, which represent the large majority of the samples, may be used for free and are thus in the public domain. It is called the 'Open NCI Database' (Development Therapeutics Program NCI/NIH; http://dtp.nci.nih.gov/webdata.html). An interesting property prediction approach to the more than 250,000 compounds contained in this open database was provided by Poroikov and co-authors [76]. By use of the program PASS (Prediction of Activity Spectra for Substances) an *in silico* tool for complex searches of 565 different types of activities is provided; e.g., in the case of antineoplastic effects, the authors could demonstrate a substantial dataset enrichment over random selection by the use of PASS-predicted probabilities.

Libraries covering a major part of entities from nature (at least some thousands) or consisting of structural information exclusively from natural origin are not free of charge, e.g., the Traditional Chinese Medicinal Database (TCMD; http://tcm3d.com/services.htm [77]) or the Dictionary of Natural Product Database launched by Chapman & Hall (DNP; http://www.chemnetbase.com) providing chemical and physical data on some 200,000 natural compounds gathered from the world's chemical literature.

An excellent survey of public and commercial databases focusing on NPs has recently been published by Füllbeck and co-authors [78]. The authors provide information as to storing characteristics of the databases, web-addresses, total number of compounds and – if given – number of natural ones. In addition, a selection of suppliers and manufacturers of natural compounds and extracts are given. A new database is introduced by the authors (Super Natural Database [79]) storing information on available NPs, thus allowing the selection of compounds that can be purchased.

Moreover a number of non-commercial in-house created databases have been used from different groups for their virtual screening studies on NPs, e.g., a marine natural product database (MNDP [80]), a natural product database (NPD [61, 81]), a database based on the antique source '*de materia medica*' by Pedanius Dioscurides (DIOS [81]), or a database fed with metabolites of ethnopharmacologically known plants [82]. Recently, Ehrman and co-authors generated a 3D multiconformational database of

Chinese herbal constituents containing a total of more than 8,000 compounds from 230 Chinese herbs [83].

10 Integrated strategies for the discovery of bioactive NPs

The more or less accurate prediction of potentially active compounds by virtual screening has doubtlessly rationalized the early drug discovery process. These filtering experiments definitely assist in saving costly and time-intensive pharmacological assays, since the pool of predicted ligands (i.e., virtual hits) is usually drastically reduced compared to the initial amount of compounds (i.e., 3D-database). Demands to be made on a good model are selectivity and target-specificity on the one hand, but it is also seminal not to lose too many valuable ligands during the filtering process.

How far all of these demands can be fulfilled strongly depends on the quality of information used as the basis for generating the model and the algorithm underlying the virtual screening process. In medicinal chemistry, an activity prediction of 10–30% is usually regarded as satisfying enrichment. In NP research, however, this percentage may be too scarce. It is rarely found that a large set of natural compounds can be acquired so easily. Only a minority of secondary metabolites are commercially available – usually at incredibly high prices. Thus, extraordinary charges and efforts are typically necessary before a virtual hit from nature is available for pharmacological testing. This process embraces the acquisition of the natural material described to contain the desired metabolite to the point of phytochemical analysis and isolation. Though advanced separation techniques, analytical instrumentation, and innovative tools for structure identification are at the phytochemists' disposal, it remains a complex and sometimes uncertain endeavor. This is why the results obtained from *in silico* predictions may nevertheless be too vague for a NP researcher.

Methods are asked to further increase the probability of following the straight tip. There is the possibility to hyphenate sundry computational approaches, e.g., pharmacophore-based virtual screening combined with docking of the resulting virtual hits, or to consider only the consensus hits applying two or more screening concepts. Nevertheless all these strate-

gies remain virtual and speculative. The combination of two approaches, which are completely divergent in nature, like a computational and an empirical one may however offer a more deepened access to bioactive NPs and may sometimes help to avoid a distorted view.

Thus, the computer-aided molecular selection is best combined with further discovery methods, labeled as integrated approaches, to increase the probability in finding a real hit. In traditional pharmacognosy there are some well established methods in targeting this aim starting from a holistic level. These include (i) hints from ethnopharmacology, (ii) phenomenological effects registered after application of naturally derived preparations, (iii) guidance of chemotaxonomy, (iv) phylogenetic selection criteria, or (v) simply information gathered from a high/medium throughput screening of extracts. In a recently published review from our group, different strategies in the field of NPs have been presented with special emphasis on anti-inflammatory NPs interacting within the arachidonic cascade [84]. Integrated computational strategies for the discovery of natural bioactive compounds have been introduced elsewhere concentrating on their scope, strengths and limits [85].

Some strategies and examples from literature combining virtual screening approaches and classical methods for activity exploitation are outlined below.

10.1 Strategy A (Fig. 5)

As soon as a sensitive data-mining tool has been developed and has proved itself by more or less selectively finding the active compounds within a test set, it can be applied for screening a 3D multi-conformational database. The subsequent procedure consists of the evaluation of the virtual hits considering physicochemical properties, toxicity and pharmacokinetics. In this stage additional virtual filtering tools for the profiling of ADME parameters [86] might have an invaluable impact to aid a refined selection of compounds. Then, a sensible choice of natural materials known to contain the focused metabolites and worth investigating in detail is a crucial step. It requires a comprehensive study in literature considering the hit content in the natural source, its availability and maybe hints from ethnopharmacology.

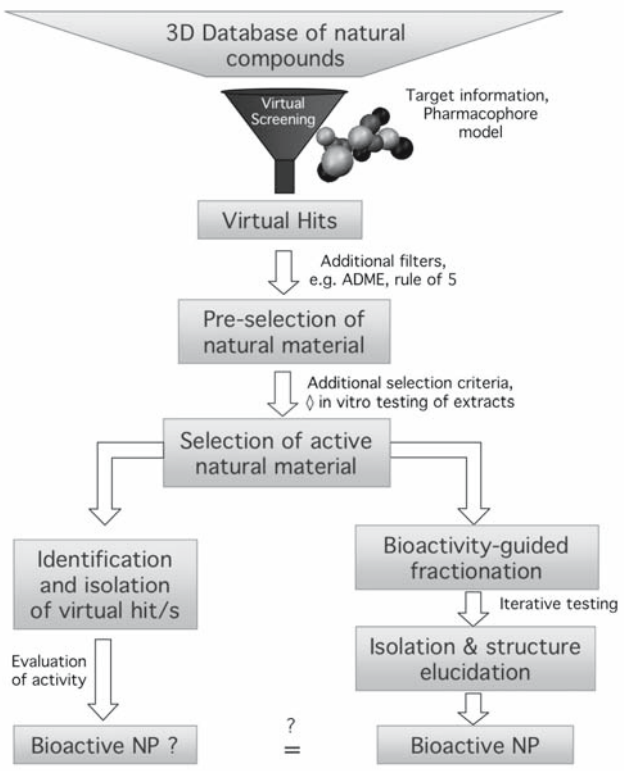


Figure 5.
Strategy A for the discovery of bioactive NPs using an integrated virtual screening approach

Once some natural materials are selected, it is advisable to perform a preliminary assay with those crude extracts and fractions assumed to contain the promising metabolite/s. Though being aware that in case of small hit amounts present in the natural material the activity may be overseen. Therefore, it is advisable to first identify the promising constituent and to possibly enrich it in the extract to be tested. Those samples that scored well are then subjected to phytochemical investigations. In this way, the tricky selection of the natural material turns from a bold venture to a more rationalized endeavor. As soon as a promising (i.e., active) starting material is found, there are in principle two possible strategies to embark on: The first one relies more on the *in silico* approach and focuses directly on

the identification of the initially obtained virtual hits within the natural matrix applying analytical tools, like LC-MS or LC-NMR, GC-MS etc. In a straightforward manner the hits are isolated using different chromatographic separation steps. After structural confirmation the compounds are then tested to hopefully verify the predicted activities. This strategy is very goal-oriented, since only pharmacological assays for the finally isolated virtual hits are necessary. On the other side, one may run the risk of ignoring further active NPs not necessarily fitting into the pharmacophore model.

The second strategy focuses on a bioactivity-guided fractionation irrespective of the virtual hits used for the selection of the starting material. Following the concept, the finally isolated active ingredients should correspond to the predicted virtual hits. This approach is usually associated with higher phytochemical efforts and costs, because it requires an iterative testing of all arising fractions and sub-fractions. For the evaluation of all the bioactive constituents in detail and for the discovery of possibly unknown metabolites this procedure is however indispensable.

The decision, which of the presented ways is the more appropriate for the investigation at hand, strongly depends on the reliability and selectivity of the used pharmacophore model, and the costliness of the used assay.

The strategy schematized in Figure 5 was recently applied to a medicinal plant with anti-inflammatory potential known from ethnopharmacological sources [87]. From the pharmacophore based virtual screening filtering experiment a number of secondary metabolites known from the mulberry tree complied with all the models' requirements, thus revealed as virtual hits. Indeed, *in vitro* tests attested extracts of *Morus* root bark a distinct COX inhibitory potential. The objective was to find the active principles from this plant material applying both different methods for their discovery. First, the computer-aided approach was used to identify the virtually active compounds able to interact with the pharmacophore models for COX-1 and -2. Second, the bioactivity-guided fractionation was conducted for the isolation of the COX-inhibiting constituents. This resulted in the isolation of nine compounds belonging to the chemical classes of sanggenons and moracins. In the enzyme assay, all the isolates showed moderate to potent inhibitory effects on COX-1 and -2. When comparing the hits of the virtual screening with the experimental data, a good correlation between predictions provided by the computer assisted method and *in vitro* data

could be obtained in the case of the isolated sanggenons (e.g., sanggenon C; **12**). However, this agreement could not be achieved with the moracins (e.g., moracin M; **13**). In any case the virtual screening was particularly helpful for the decision regarding which plant material is worth extensive study. Furthermore, the disclosed interactions of the sanggenons with the pharmacophore model – miming the binding site of the target – provided us with some essential information about the molecular requirements of COX-ligands.

10.2 Strategy B

A different integrated procedure is schematized in Figure 6. Applying this approach, the pre-selection of the natural material is not guided by virtual prediction; but a number of extracts is roughly screened with a bioassay to identify the active ones. A similar strategy is to collect information about the traditional application of natural preparations in the field of the focused pharmacological target. A 3D database is then generated consisting of all the metabolites known from literature to be included in that extract/s that came off well. Likewise, ethnopharmacological knowledge about useful preparations from nature may guide the selection of NPs. The resulting biased database is virtually screened with an established pharmacophore model of the aiming target.

The impact of ethnopharmacology has been analyzed in a previous study from our group; there we investigated the statistical evidence considering hints from folk medicine for the discovery of anti-inflammatory NPs utilizing pharmacophore-based virtual screening techniques [81]. COX-1 and -2 were used as preferential targets, since they are key enzymes in the inflammation process. Dioscorides' *de materia medica*, which was written in the 1st Century AD, was used as the ethno-pharmacological source. Secondary metabolites of those medicinal plants, which Dioscorides described as active against fever, rheumatism, pain and pus were stored in a multi-conformational 3D database. This was virtually screened against the validated pharmacophore models. The resulted hit list was analyzed and compared with those obtained by screening unbiased databases of natural as well as of synthetic origin. The effectiveness of an ethnopharmacological approach could be statistically demonstrated by obtaining a significantly

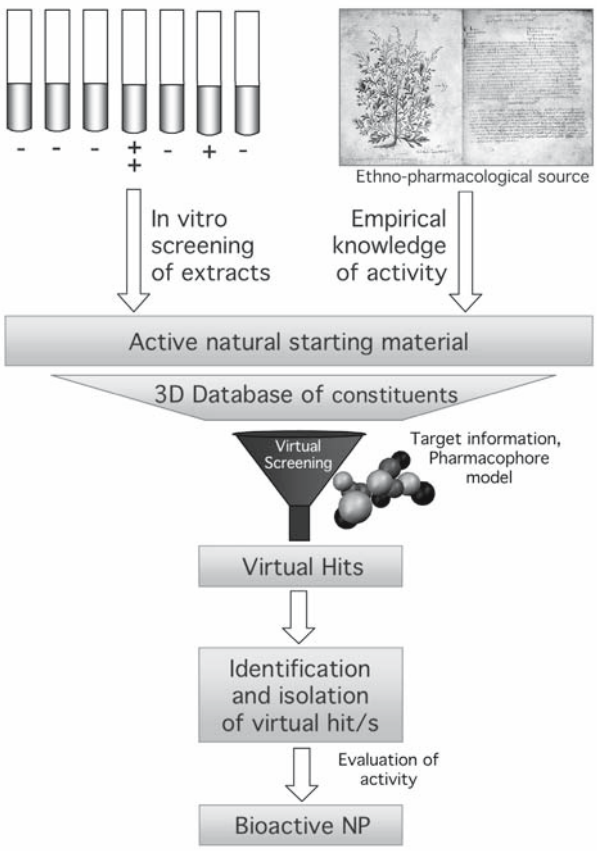


Figure 6.
Strategy B for the discovery of bioactive NPs using an integrated virtual screening approach

higher hit rate compared to the hit rates of the unbiased natural as well as synthetic databases.

Following this strategy the putative hits may then be identified by modern analytical tools like LC-MS or LC-NMR to isolate them from the natural matrix in a target-oriented way for pharmacological testing. This approach is especially helpful for intricate pharmacological assays, which would turn a bioguided fractionation into an unrealistic endeavor.

A combination of an ethnopharmacologically based pre-selection of plant material and a computational approach was reported by Bernard

and co-workers, who used this strategy to rationalize a phytochemical lead discovery [88]. Starting with an *in vitro* screening on phospholipase A_2 performed with traditionally used anti-inflammatory plant extracts, a focused structural database was generated and virtually screened on an established ligand-based pharmacophore model for human non-pancreatic phospholipase A_2. The combination of experimental data with database exploitation and molecular modeling resulted in the efficient identification of betulin (**14**) and betulinic acid (**15**) as extract ingredients with distinct anti-inflammatory *in vitro* effects.

The combination of the two different, but complementary strategies consisting of *in vitro* screens and *in silico* assessment has recently been described by van de Waterbeemd [89]. He labeled this method as '*in combo*' approach and used it for the straight forward access of various ADME properties. The application of the '*in combo*' approach for the discovery of NPs has recently been tested in our group by the search of natural acetylcholinesterase inhibitors [90]. In a medium-sized throughput screening about 100 plant extracts were investigated using an acetylcholinesterase enzyme test. From the sample showing the best inhibitory activity, all the known secondary metabolites were fed into a small 3D multiconformational database and subsequently subjected to a virtual screening on a generated pharmacophore model. The efficacy of this procedure could be confirmed by the isolation of the obtained virtual hits, i.e., 8-deoxylactucin (**16**) and lactucopicrine (**17**). They showed a significant and dose-dependent inhibitory effect in the enzyme assay.

Methods and expectations of this integrated virtual screening concept have previously been discussed in detail by J. Bajorath [91, 92] with the author's final statement that '*a meaningful integration of virtual and experimental screening programs, together with lessons to be learned from structural genomics, holds great promise for more rapid and consistent identification of high quality hits or leads across divers classes of therapeutic targets*'. Though this conclusion was not particularly coined to NPs, it comes especially true in the rich world of secondary metabolites.

Further hybridized computational strategies are quite sensible to get an improved understanding of ligand-target interactions. In the following two examples docking protocols helped enlighten the molecular mechanism of bioactive natural compounds. Chimenti and co-authors isolated quercetin (**18**) among other secondary metabolites from the Mediterra-

nean shrub *Hypericum hircinum* and identified this flavonol as selective inhibitor of the MAO-A with an activity in the nanomolar range (IC_{50} = 10 nM) [93]. For a more comprehensive understanding of the underlying molecular selectivity, conformation analysis and docking simulations were performed using the most recent crystallographic structures of both human isoforms MAO-A and MAO-B. This enabled the authors to identify the most important interactions between the residues and the cofactor within the enzymatic cleft. The estimated free energies of complexation were in agreement with experimental data and confirmed the distinct preference for the MAO-A cleft with more intermolecular hydrogen bonds and π-π interactions.

The goal of a recent in-house study was to rationalize the binding interaction of the protoalkaloid taspine (**19**) within acetylcholinesterase. Taspine was isolated in a bioactivity-guided manner from *Magnolia x soulangiana* and revealed as selective inhibitor of acetylcholinesterase with a significantly higher effect than the positive control galanthamine (**20**; IC_{50} = 0.33 ± 0.07 µM). Extensive molecular docking studies were performed with human and Torpedo californica-acetylcholinesterase employing Gold software (Vers. 3.1; www.ccdc.cam.ac.uk/products/life_sciences/gold/). The results suggested taspine to bind in an alternative binding orientation than galanthamine [94]. While this is located in close vicinity to the catalytic amino acid triad, taspine was found to be mainly stabilized by sandwich-like π-stacking interactions in the aromatic gorge of the enzyme.

In both case studies the active natural compound was already identified. Thus, the *in silico* tool was not employed for data mining, but to elicit the putative binding mode in the macromolecular target. Docking simulations turned out to be excellent tools to get an idea about the assumed molecular ligand target interaction.

10.3 Strategy C

Another approach capitalizes exactly on the just-mentioned observation that computational predictions may reveal an idea about the interaction to a specific target's binding site. Thus, it is possible to start with one compound of unknown activity and to mine it against a number of structurally disclosed targets in terms of elaborated pharmacophore models (Fig. 7), i.e.,

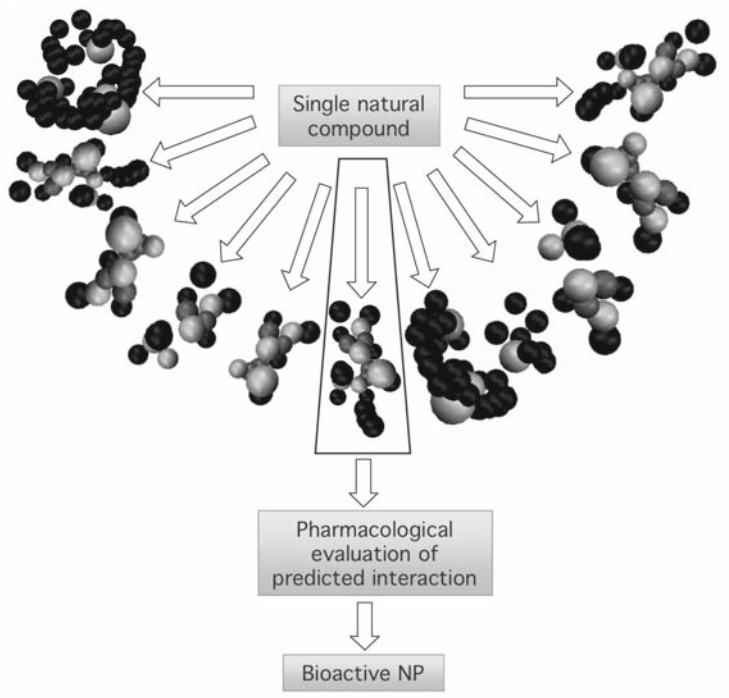


Figure 7.
Parallel screening for the discovery of bioactive NPs

parallel *in silico* screening (see previous). As soon as the orphaned molecule is able to comply with all the requirements and restrictions imposed by any model, it can be assessed as rational hint. Consequently, the focused compound will be subjected to a pharmacological testing on the predicted target/s. In this way, the parallel screening is not only helpful to estimate the interactions of a drug candidate with diverse antitargets; or to canvass its interactions to related targets as is performed for an activity profiling. In this approach, the parallel screening is a computational tool for *target fishing* to get a rational idea about any potential target interaction and to prioritize a few targets for experimental evaluation by applying simple ligand-based or target-based queries. The potential of virtual screening of target libraries was recently discussed by Didier Rognan [95]. In his group a structure-based method for target screening was pursued applying inverse

docking [96]. The authors used 2,148 structurally well-defined PDB entries to build a 3D protein library. The virtual screening of this protein library with four unrelated ligands was suitable for recovering the true targets of specific ligands and may as well be used for virtual selectivity profiling of any ligand of interest.

Nettles and co-authors performed the *target fishing* approach using a ligand-based procedure [97]. The potential of both 2D and 3D chemical descriptors were compared as tools for predicting the biological targets of ligand probes on the basis of their similarity to reference molecules in a chemical database comprising 46,000 biologically annotated compounds. The ligand-based 3D tool FEPOPS (FEature POint PharmacophoreS), which provides pharmacophoric alignment of the small molecules' chemical features consistent with those seen in experimental ligand/receptor complexes, was used for scaffold jumping within the screened database. Using ATP the authors were able to identify the natural compound balanol (21) as ligand of CDK2.

The highest effort applying this strategy is the availability of a representative amount of reliable pharmacophore models covering a wide range of relevant targets (Fig. 7). Thus, it may be of particular interest to focus on one pathological syndrome, e.g., obesity, inflammation, apoptosis etc., where a phenomenological activity of a NP is already evident. Applying this approach the disposition of pharmacophore models for targets involved in the respective pathological complex is easier to manage. In this way, a goal-oriented strategy may help to bridge the gap between a phenomenological effect and the underlying molecular mode of action.

11 Conclusion

Pertaining to the drug discovery from nature we are facing two facts: (i) statistics show that the myriad of structurally diverse natural compounds are the most favored source of new drugs for clinical use [5]; (ii) the drug discovery process has moved towards more rational concepts based on the increasing understanding of the molecular principles of protein–ligand interactions. Spurred on by economic interest fundamental advances have been made in research applying data mining strategies, like virtual screening.

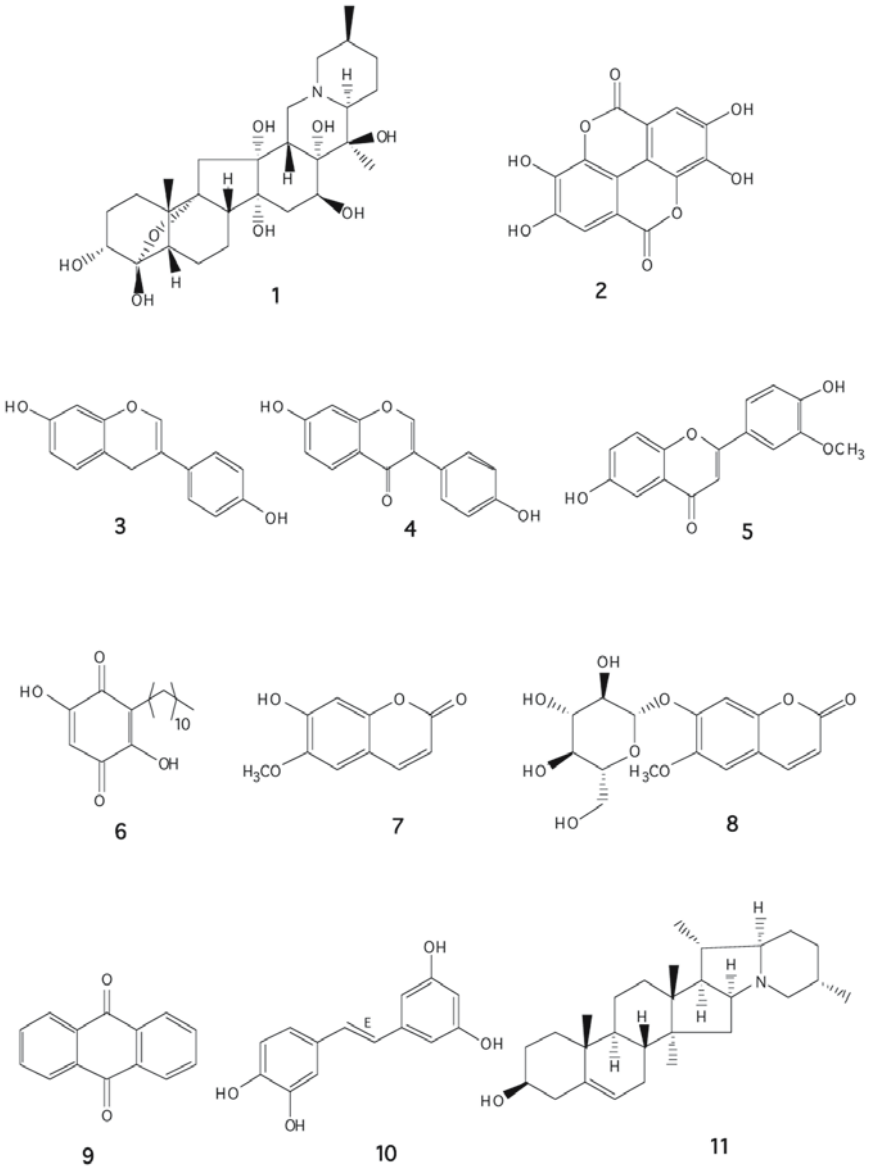
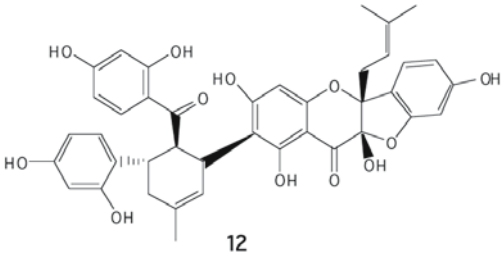
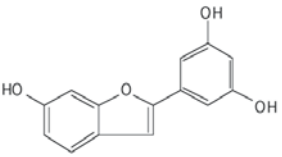


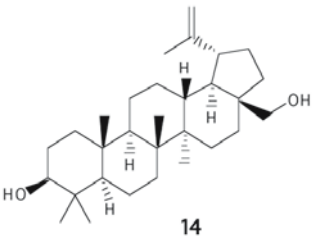
Chart 1.
Structures 1–11



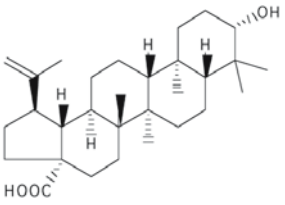
12



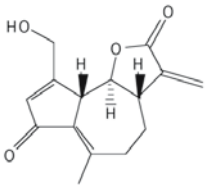
13



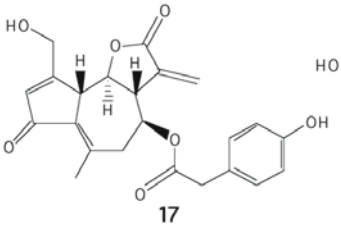
14



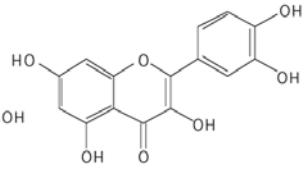
15



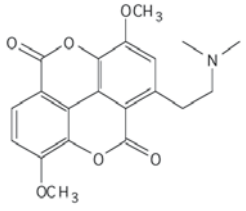
16



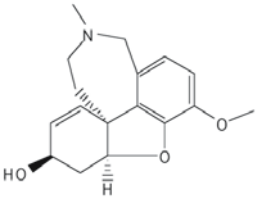
17



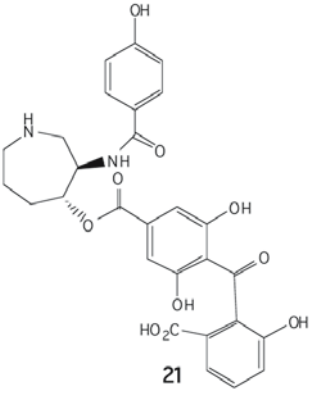
18



19



20



21

Chart 2.
Structures 12–21

Though being aware of both potentials, their combined benefit could only rudimentary be savored. Only limited attempts applying innovative *in silico* tools in NP research are pursued so far, because the search for bioactive compounds is a complex and multidisciplinary challenge. Thus, a sensible adaptation of computational strategies is in demand to profit in an economic way from the unique chemical and biological diversity associated with NPs. Virtual screening techniques, however, must not be used exclusively as activity-predicting tools, since the results provide merely an indication for a putative activity: it is only by the creation of interfaces between computational tools and well-established methods from pharmacognosy that a reasonable standard of success can be achieved. The search for the most effective strategy is best performed by a drug discovery process that involves the exploitation of all the information which can be gathered from bioactivity-guided fractionation, on-line analytical activity profiling, ethnopharmacological screening, chemoinformatics, virtual and *in vitro* screening studies. In the first instance it behoves modern pharmacognosy to skillfully exploit knowledge from all these fields because it is of paramount importance to sift through the enormous wealth of NPs.

Examples underlining the impact of virtual screening on the identification of active NPs have been presented in this survey. Though the full potential in this field is by far untapped, these early results indicate that the integrated virtual screening approaches are target-oriented and trendsetting strategies. However, as any computer-based technique, the successful use of virtual screening will entirely depend on the way it is utilized and the quality of its underlying experimental data. The advantages implemented to a virtual screening cycle compared to a conventional *in vitro* screening are obvious: (i) higher capacity, (ii) no need for isolated compounds, (iii) less experimental efforts for testing; (iv) theoretically, interactions of all known NPs to all structurally defined targets can be calculated and predicted, (v) the quality of hit compounds can be increased by additional drug-like filters and virtually restricted ADME properties; thus diminishing failures in the early drug development.

Nevertheless experimental investigations are seminal, but can be focused in a more effective fashion. A cautious handling of virtual hits together with lessons learned from traditional pharmacognosy seems to be crucial for a successful exploitation of treasures from nature. In this area,

virtual screening will most likely play an essential role in accelerating the early stage of drug discovery by efficiently digging out lead compounds from nature.

References

1 Brown D, Superti-Furga G (2003) Rediscovering the sweet spot in drug discovery. *Drug Discov Today* 8: 1067–1077

2 Drews J (2003) Strategic trends in the drug industry. *Drug Discov Today* 8: 411–420

3 Smith A (2002) Screening for drug discovery: the leading question. *Nature* 418: 453–459

4 Strohl WR (2000) The role of natural products in a modern drug discovery program. *Drug Disc Today* 5: 39

5 Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J Nat Prod* 70: 461–477

6 Harvey A (2000) Strategies for discovering drugs from previously unexplored natural products. *Drug Disc Today* 5: 294–300

7 Clardy J, Walsh C (2004) Lessons from natural molecules. *Nature* 432: 829–837

8 Jones WP, Chin YW, Kinghorn AD (2006) The role of pharmacognosy in modern medicine and pharmacy. *Curr Drug Targets* 7: 247–264

9 Dictionary of Natural Products provided by Chapman &Hall/CRC. Available at: http://www.crcpress.com/shopping_cart/products/product_detail.asp?sku=C9150 (accessed in January 2007)

10 Tulp M, Bohlin L (2005) Rediscovery of known natural compounds: Nuisance or goldmine? *Bioorg Med Chem* 13: 5274–5282

11 Sitte P, Weiler EW, Kadereit JW, Bresinsky A, Körner C (2002) *Strasburger – Lehrbuch der Botanik*, 35. Auflage, Spektrum Akademischer Verlag, Heidelberg, Berlin, 339–351

12 Balandrin M (1993) Plant-derived natural products in drug discovery and development: an overview. *Amer Chem Soc Symposium Series* No 534: 2–12

13 Cragg GM, Newman DJ, Snader KM (1997) Natural products in drug discovery and development. *J Nat Prod* 60: 52–60

14 Henkel T, Brunne RM, Muller H, Reichel F (1999) Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angew Chem Int Ed* 38: 643–647

15 Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar M, Doyle M, FitzHugh W et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860–921

16 Deisenhofer J, Smith JL (2001) Proteins. *Curr Opin Struc Biol* 11: 701–702

17 Maggio ET, Ramnarayan K (2001) Recent developments in computational proteomics. *Drug Disc Today* 6: 996–1004

18 Potterat O, Hamburger M (2006) Natural products in drug discovery – concepts and approaches for tracking bioactivity. *Curr Org Chem* 10: 899–920

19 Pottereat O (2006) Targeted approaches in natural product lead discovery. *Chimia* 60: 19–22

20 Morrell J (1996) Mining information from databases for drug discovery. *Book Abstr 211th ACS National Meeting* CINF–052

21 Gasteiger J, Teckentrup A, Terfloth L, Spycher S (2003) Neural networks as data mining tools in drug design. *J Phys Org Chem* 16: 232–245

22 Xu J, Hagler A (2002) Chemoinformatics and drug discovery. *Molecules* 7: 566–600

23 Alvarez AJ, Shoichet B (2005) *Virtual screening in drug discovery*. Taylor & Francis, CRC-Press, Boca Raton

24 Langer T, Hoffmann RD (2006) Pharmacophores and pharmacophore searches. In: *Methods and principles in medicinal chemistry* Vol. 32, Wiley-VCH, Weinheim

25 Manly CJ, Louise-May S, Hammer JD (2001) The impact of informatics and computational chemistry on synthesis and screening. *Drug Discov Today* 6: 1101–1110

26 Langer T, Hoffmann RD (2001) Virtual screening: an effective tool for lead structure discovery? *Curr Pharm Des* 7: 509–527

27 Böhm HJ, Schneider G (2000) *Virtual screening for bioactive molecules*. Wiley, New York

28 Krovat EM, Steindl T, Langer T (2005) Recent advances in docking and scoring. *Curr Comput-Aided Drug Des* 1: 93–102

29 Abagyan R, Totrov M (2001) High-throughput docking for lead generation. *Curr Opin Chem Biol* 5: 375–382

30 Schneider G, Böhm HJ (2002) Virtual screening and fast automated docking methods. *Drug Disc Today* 7: 64–70

31 Shoichet BK, McGovern SL, Wei B, Irwin JJ (2002) Lead discovery using molecular docking. *Curr Opin Chem Biol* 6: 439–446

32 Stahl M, Rarey M (2001) Detailed analysis of scoring functions for virtual screening. *J Med Chem* 44: 1035–1042

33 Charifson PS, Corkery JJ, Murcko MA, Walters WP (1999) Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J Med Chem* 42: 5100–5109

34 Wang R, Wang S (2001) How does consensus scoring work for virtual library screening? An idealized computer experiment. *J Chem Inf Comput Sci* 41: 1422–1426

35 Liu B, Zhou J (2005) SARS-CoV protease inhibitors design using virtual screening method from natural products libraries. *J Comp Chem* 26: 484–490

36 Toney JH, Navas-Martin S, Weiss SR, Koeller A (2004) Sabadinine: a potential non-peptide anti-severe acute-respiratory-syndrome agent identified using structure-aided design. *J Med Chem* 47: 1079–1080

37 Cozza G, Bonvini P, Zorzi E, Poletto G, Pagano MA, Sarno S, Donella-Deana A, Zagotto G, Rosolen A, Pinna LA et al (2006) Identification of ellagic acid as potent inhibitor of protein kinase CK2: a successful example of a virtual screening application. *J Med Chem* 49: 2363–2366

38 Zhao L, Brinton RD (2005) Structure-based virtual screening for plant-based ERβ-selective ligands as potential preventative therapy against age-related neuro-degenerative diseases. *J Med Chem* 48: 3463–3466

39 Liu H, Li Y, Song M, Tan X, Cheng F, Zheng S, Shen J, Luo X, Ji R, Yue J et al (2003) Structure-based discovery of potassium channel blockers from natural products virtual screening and electrophysiological assay testing. *Chem Biol* 10: 1103–1113

40 Langer T, Hoffmann RD (2006) Pharmacophore modelling: applications in drug discovery. *Exp Opin Drug Discov* 1: 261–267

41 Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, Stallings WC, Connolly DT, Shoichet BK (2002) Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J Med Chem* 45: 2213–2221
42 Samiulla DS, Vaidyanathan VV, Arun PC, Balan G, Blaze M, Bondre S, Chandrasekhar G, Gadakh A, Kumar R, Kharvi G et al (2005) Rational selection of structurally diverse natural product scaffolds with favorable ADME properties for drug discovery. *Mol Divers* 9: 131–139
43 Testa B (1984) Drugs? Drug research? Advances in drug research? Musings of a medicinal chemist. *Adv Drug Res* 13: 1–58
44 Smith DA, van de Waterbeemd H, Walker DK (2006) *Pharmacokinetics and metabolism in drug design*. 2nd Ed., Wiley-VCH, Weinheim
45 Avdeef A, Testa B (2002) Physicochemical profiling in drug research: a brief survey of the state-of-the-art of experimental techniques. *Cell Mol Life Sci* 59: 1681–1689
46 Feher M, Schmidt JM (2003) Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci* 43: 218–227
47 Stahura F, Godden JW, Ling X, Bajorath J (2002) Distinguishing between natural products and synthetic molecules by descriptor Shannon entropy analysis and binary QSAR calculations. *J Chem Inf Comput Sci* 40: 1254–1252
48 Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 23: 3–25
49 Congreve M, Murray CW, Blundell TL (2005) Structural biology and drug discovery. *Drug Disc Today* 10: 895–907
50 Berman H, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov I, Bourne P (2000) The protein data bank. *Nucleic Acids Res* 28: 235–242
51 Wolber G, Langer T (2000) LigandScout: 3D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model* 45: 160–169
52 Krovat EM, Frühwirth KH, Langer T (2005) Pharmacophore identification, *in silico* screening, and virtual library design for inhibitors of the human factor Xa. *J Chem Inf Model* 45: 146–159
53 Rella M, Rushworth C, Guy JL, Turner AJ, Langer T, Jackson RM (2006) Structure-based pharmacophore design and virtual screening for novel angiotensin converting enzyme 2 inhibitors. *J Chem Inf Model* 46: 708–716
54 Barreca ML, De Luca L, Iraci N, Rao A, Ferro S, Maga G, Chimirri A (2007) Structure-based pharmacophore identification of new chemical scaffolds as non-nucleoside reverse transcriptase inhibitors. *J Chem Inf Model* 47: 557–562
55 Schuster D, Laggner C, Steindl TM, Langer T (2006) Development and validation of an *in silico* P450 profiler based on pharmacophore models. *Curr Drug Discov Technol* 3: 1–48
56 Steindl TM, Schuster D, Laggner C, Langer T (2006) Parallel screening: a novel concept in pharmacophore modelling and virtual screening. *J Chem Inf Model* 46: 2146–2157
57 Steindl TM, Schuster D, Wolber G, Laggner C, Langer T (2007) High throughput structure-based pharmacophore modeling as a basis for successful parallel virtual screening. *J Comput-aided Mol Des* ASAP doi 10.1007/s10822–006–9066–y
58 Steindl TM, Schuster D, Laggner C, Chuang K, Hoffmann R, Langer T (2007) Parallel screening and activity profiling with HIV protease inhibitor pharmacophore models. *J Chem Inf Model* 47: 563–571

59 Nikolovska-Coleska Z, Xu L, Hu Z, Tomita Y, Li P, Roller PP, Wang R, Fang X, Guo R, Zhang M et al (2004) Discovery of embelin as a cell-permeable, small-molecular weight inhibitor of XIAP through structure-based computational screening of a traditional herbal medicine three-dimensional structure database. *J Med Chem* 47: 2430–2440

60 Wu G, Chai J, Suber TL, Wu JW, Du C, Wang X, Shi Y (2000) Structural basis of IAP recognition by Smac/DIABLO. *Nature* 408: 1008–1012

61 Rollinger JM, Hornick A, Langer T, Stuppner H, Prast H (2004) Acetylcholinesterase inhibitory activity of scopolin and scopoletin discovered by virtual screening of natural products. *J Med Chem* 47: 6248–6254

62 Langer T, Krovat EM (2003) Chemical feature-based pharmacophores and virtual library screening for discovery of new leads. *Curr Opin Drug Discov Dev* 6: 370–376

63 Schuster D, Maurer E, Laggner C, Nashev L, Wilckens T, Langer T, Odermatt A (2006) The discovery of new 11β-hydroxysteroid dehydrogenase Type 1 inhibitors by common feature pharmacophore modeling and virtual screening. *J Med Chem* 49: 3454–3466

64 Schuster D, Laggner C, Steindl TM, Palusczak A, Hartmann RW, Langer T (2006) Pharmacophore modeling and *in silico* screening for new P450 19 (aromatase) inhibitors. *J Chem Inf Model* 46: 1301–1311

65 Kurogi Y, Güner OF (2001) Pharmacophore modeling and three-dimensional database searching for drug design using Catalyst. *Curr Med Chem* 8: 1035–1055

66 Güner O, Clement O, Kurogi Y (2004) Pharmacophore modeling and three dimensional database searching for drug design using CATALYST: Recent advances. *Curr Med Chem* 11: 763–771

67 Füllbeck M, Huang X, Dumdey R, Frommel C, Dubiel W, Preissner R (2005) Novel curcumin- and emodin-related compounds identified by *in silico* 2D/3D conformer screening induce apoptosis in tumor cells. *BMC Cancer* 5: 97

68 Laggner C, Schieferer C, Fiechtner B, Poles G, Hoffmann RD, Glossmann H, Langer T, Moebius F (2005) Feature based pharmacophore models for sigma1 receptor, ERG2 and EBP. *J Med Chem* 48: 4754 –4764

69 Zupan J, Gasteiger J (1999) *Neural networks in chemistry and drug design*, 2nd Ed. Wiley-VCH, Weinheim

70 Wagner S, Hofmann A, Siedle B, Terfloth L, Merfort I, Gasteiger J (2006) Development of a structural model for NF-κB inhibition of sesquiterpene lactones using self-organizing neural networks. *J Med Chem* 49: 2241–2252

71 Sangma C, Chuakheaw D, Jongkon N, Saenbandit K, Nunrium P, Uthayopas P, Hannongbua S (2005) Virtual screening for anti-HIV-1 RT and anti-HIV-1 PR inhibitors from the Thai Medicinal Plants Database: A combined docking with neural networks approach. *Comb Chem HTS* 8: 417–429

72 Cherkasov A, Shi Z, Fallahi M, Hammond GL (2005) Successful *in silico* discovery of novel nonsteroidal ligands for human sex hormone binding globulin. *J Med Chem* 48: 3203–3213

73 Svetnik V, Liaw A, Tong D, Culberson C, Sheridan RP, Feuston BP (2003) Random Forest: a classification and regression tool for compound classification and QSAR Modeling. *J Chem Inf Comput Sci* 43: 1947–1958

74 Ehrman TM, Barlow DJ, Hylands PJ (2007) Virtual screening of Chinese herbs with random forest. *J Chem Inf Model* ASAP 10.1021/ci600289v

75 Kirchmair J, Laggner C, Wolber G, Langer T (2005) Comparative analysis of protein-

bound ligand conformations with respect to catalyst's conformational space subsampling algorithms. *J Chem Inf Model* 45: 422–430

76 Poroikov VV, Filimonov DM, Ihlenfeldt WD, Gloriozova TA, Lagunin AA, Borodina YV, Stepanchikova AV, Nicklaus MC (2003) PASS Biological activity predictions in the enhanced open NCI database browser. *J Chem Inf Comput Sci* 43: 228–236

77 Lu A, Liu B, Liu H, Zhou J, Xie G (2004) A traditional Chinese medicine plant-compound database aid its application for searching. *Int Electron J Mol Des* 3: 672–683

78 Füllbeck M, Michalsky E, Dunkel M, Preissner R (2006) Natural products: sources and databases. *Nat Prod Rep* 23: 347–356

79 Dunkel M, Füllbeck M, Neumann S, Preissner R (2006) SuperNatural: a searchable database of available natural compounds. *Nucleic Acid Res* 34: D678–683

80 Lei J, Zhou J (2002) A marine natural product database. *J Chem Inf Comp Sci* 42: 742–748

81 Rollinger JM, Haupt S, Stuppner H, Langer T (2004) Combining ethnopharmacology and virtual screening for lead structure discovery: COX-inhibitors as application example. *J Chem Inf Comp Sci* 44: 480–488

82 Bernard P, Berton JY, Chrétien JR (1999) Computer-aided molecular selection and design of natural bioactive molecules. *Curr Opin Drug Disc Dev* 2: 213–223

83 Ehrman TM, Barlow DJ, Hylands PJ (2007) Phytochemical databases of Chinese herbal constituents and bioactive plant compounds with known target specifities. *J Chem Inf Model* ASAP 10.1021/ci600288m

84 Rollinger JM, Langer T, Stuppner H (2006) Strategies for efficient lead structure discovery from natural products. *Curr Med Chem* 13: 1491–1507

85 Rollinger JM, Langer T, Stuppner H (2006) Integrated *in silico* tools to exploit the natural products' bioactivity. *Planta Med* 72: 671–678

86 Van de Waterbeemd H, Gifford E (2003) ADMET *in silico* modelling: towards prediction paradise? *Nature Rev Drug Disc* 2: 192–204

87 Rollinger JM, Bodensieck A, Seger C, Ellmerer EP, Bauer R, Langer T, Stuppner H (2005) Discovering COX-inhibiting constituents of *Morus* root bark: Activity-guided *versus* computer-aided methods. *Planta Med* 71: 399–405

88 Bernard P, Scior T, Didier B, Hibert M, Berthon JY (2001) Ethnopharmacology and bioinformatic combination for leads discovery: application to phospholipase A2 inhibitors. *Phytochemistry* 58: 865–874

89 Van de Waterbeemd H (2005) Which *in vitro* screens guide the prediction of oral absorption and volume of distribution? *Bas Clin Pharmacol Toxicol* 96: 162–166

90 Rollinger JM, Mock P, Zidorn C, Ellmerer EP, Langer T, Stuppner H (2005) Application of the *in combo* screening approach for the discovery of non-alkaloid acetylcholinesterase inhibitors from *Cichorium intybus*. *Curr Drug Discov Techn* 2: 185–193; Erratum (2006) 3: 89

91 Bajorath J (2002) Virtual screening in drug discovery: methods, expectations and reality. *Curr Drug Disc* March: 24–28

92 Bajorath J (2002) Integration of virtual and high-throughput screening. *Nature Rev Drug Discovery* 1 Nov: 882–894

93 Chimenti F, Cottiglia F, Bonsignore L, Casu L, Casu M, Floris C, Secci D, Bolasco A, Cimenti P, Granese A et al (2006) Quercetin as the active principle of *Hypericum hircinum*

exerts a selective inhibitory activity against MAO-A: extraction, biological analysis, and computational study. *J Nat Prod* 69: 945–949

94 Rollinger JM, Schuster D, Baier E, Ellmerer EP, Langer T, Stuppner H (2006) Taspine: Bioactivity-guided isolation and molecular ligand-target insight of a potent acetylcholinesterase inhibitor from *Magnolia x soulangiana*. *J Nat Prod* 69: 1341–1346

95 Rognan D (2006) *In silico* screening of the protein structure repertoire and of protein families. *Chemogenomics* 109–131

96 Paul N, Kellenberger E, Bret G, Mueller P, Rognan D (2004) Recovering the true targets of specific ligands by virtual screening of the Protein Data Bank. Proteins 54 : 671–680

97 Nettles JH, Jenkins JL, Bender A, Deng Z, Davies JW, Glick M (2006) Bridging chemical and biological space: "target fishing" using 2D and 3D molecular descriptors. *J Med Chem* 49: 6802–6810

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Strain improvement for production of pharmaceuticals and other microbial metabolites by fermentation

By Arnold L. Demain[1] and Jose L. Adrio[2]

[1]Research Institute for Scientists Emeriti (R.I.S.E.), HS-330, Drew University, Madison, NJ 07940 USA
<ademain@drew.edu>

[2]Industrial Biotechnology Unit, NEURONBioPharma, S.A., Edificio BIC, Lab-, Campus Tecnológico Ciencias de la Salud, 18100-Armilla, Granada, Spain
<jladrio@neuronbiopharma.com>

Abstract

Microbes have been good to us. They have given us thousands of valuable products with novel structures and activities. In nature, they only produce tiny amounts of these secondary metabolic products as a matter of survival. Thus, these metabolites are not overproduced in nature, but they must be overproduced in the pharmaceutical industry. Genetic manipulations are used in industry to obtain strains that produce hundreds or thousands of times more than that produced by the originally isolated strain. These strain improvement programs traditionally employ mutagenesis followed by screening or selection; this is known as 'brute-force' technology. Today, they are supplemented by modern strategic technologies developed *via* advances in molecular biology, recombinant DNA technology, and genetics. The progress in strain improvement has increased fermentation productivity and decreased costs tremendously. These genetic programs also serve other goals such as the elimination of undesirable products or analogs, discovery of new antibiotics, and deciphering of biosynthetic pathways.

1 Introduction

Microbes tend not to overproduce their metabolites and usually produce valuable products only in amounts that they need for their own benefit. In strain improvement programs of industry, a strain producing a high titer is usually the desired goal. Genetics has contributed to the production of microbial products for many years. The tremendous increases in fermentation productivity and the resulting decreases in costs have come about mainly by mutagenesis and screening/selection for higher producing microbial strains and the application of recombinant DNA technology.

Strain improvement encompasses creation of strains with (i) efficient assimilation of inexpensive and complex raw materials; (ii) alteration of product ratios and elimination of byproducts; (iii) product excretion; (iv) tolerance to high product concentrations; (v) short fermentation times; and (vi) overproduction of native products or foreign products after genetic recombination [1].

The contributions of microbial genetics to industrial microbiology began in the 1940s when the fermentative production of penicillin became an international necessity. The early studies in basic genetics concentrated heavily on the production of mutants and their properties. The ease with which 'permanent' characteristics of microorganisms could be changed by mutation and the simplicity of the mutation techniques had tremendous

appeal to microbiologists. Mutation has been the major factor involved in the hundred to thousand-fold increases obtained in production of microbial metabolites. The ability to modify genetically a microbial culture to higher productivity has been the most important factor in keeping the fermentation industry in a healthy state.

2 Mutagenesis

Microorganisms generate new genetic characters ('genotypes') by two means: (i) mutation and (ii) genetic recombination techniques such as protoplast fusion, transformation, conjugation and recombinant DNA technology, including metabolic engineering.

In mutagenesis, a gene is modified either unintentionally ('spontaneous mutation') or intentionally ('induced mutation'). Although the change is usually detrimental and eliminated by selection, some mutations are beneficial to the microorganism. Even if not beneficial to the organism, but beneficial to humans, the mutation can be detected by screening and preserved indefinitely. Mutation has been mainly used to improve the productivity of industrial cultures [2, 3], although it has also been used to shift the proportion of metabolites produced in a fermentation broth to a more favorable distribution, elucidate the pathways of secondary metabolism, yield new compounds, and other functions.

The most useful mutagens include nitrosoguanidine (NTG), 4-nitroquinolone-1-oxide, methylmethane sulfonate (MMS), ethylmethane sulfonate (EMS), hydroxylamine (HA) and ultraviolet light (UV). The most common method used to obtain high yielding mutants is that of treating a population with a mutagenic agent until a certain 'desired' kill is obtained, plating out the survivors and testing each resulting colony or a randomly selected group of colonies for product formation in flasks. The optimum level of kill for increased production of antibiotics is thought to be in the range of 70–95% [4], although some industrial programs use much higher levels, e.g., up to 99.99%. Although a successful mutation and screening procedure decreases the average production ability of all the survivors, more importantly, it increases frequency of improved mutants. Although single cells or spores are preferred for mutagenesis

[5–8], non-spore-forming filamentous organisms can also be handled. Here mycelia are mutagenized, protoplasts are formed and regeneration is carried out on solid medium [9]. Sonication is sometimes used to break up mycelia after mutagenesis and before screening for improved mutants [10–12].

2.1 Increasing metabolite production

Genetics has led to tremendous increases in fermentation productivity and decreased costs mainly by mutagenesis and screening for higher producing microbial strains. Overproduction of microbial metabolites is effected by (i) increasing precursor pools, (ii) adding, modifying or deleting regulatory genes, (iii) altering promoter, terminator and/or regulatory sequences, (iv) increasing copy number of genes encoding enzymes catalyzing bottleneck reactions, or (v) removing competing unnecessary pathways [13].

The first superior penicillin-producing mutant, *Penicillium chrysogenum* X-1612, was isolated afer X-ray mutagenesis in the mid-20th Century. This heralded the beginning of a long and successful relationship between mutational genetics and industrial microbiology [14]. Improvement of penicillin production by conventional strain improvement resulted both from enhanced gene expression and from gene amplification [15, 16]. Increased levels of mRNA corresponding to the three enzymes of penicillin G biosynthesis have been found in high-penicillin producing strains of *P. chrysogenum* as compared to wild-type strains [17]. High-producing strains contain an amplified region which is at least 35 kb. A 106 kb region amplified 5 to 6 times as tandem repeats was detected in a high-producing strain whereas Fleming's original strain of *Penicillium notatum* contained only a single copy [18].

Impressive titers of industrial metabolites have been achieved in all mutagenesis programs. The production titer of tetracycline as far back as 1979 was over 20 g/l [19], mainly due to strain improvement. Later, titers of 30–35 g/l were reached for chlortetracycline and tetracycline. Penicillin production titer is 70 g/l and that of cephalosporin C over 30 g/l [20]. Tylosin titer has been reported to be over 15 g/l [21] and that of salinomycin 60 g/l [22].

2.1.1 Mutants with altered morphology or color

Randomly obtained mutants as well as morphological mutants are often improved in production ability but the mechanisms are relatively unknown. Probably many of these mutations involve regulatory genes, especially since regulatory mutants obtained in basic genetic studies are sometimes found to be altered in colonial morphology. Morphological mutants were crucial in the strain improvement of *Streptomycees coeruleorubidus*, the producer of the anthracycline antitumor agent daunomycin (now daunorubicin) [23]. Screening for 'bald' mutants (suppressed in aerial mycelium development) followed by screening for white mutants (asporogenous aerial mycelium) led to a 15-fold improvement. Morphological variation was also useful in improving production of another anthracycline, beromycin [24]. The parent culture, producing about 100 mg/l, was streaked on three different agar media of which only one revealed different types of colonies. The best type had reddish orange colonies, no aerial mycelia, orange submerged mycelia, no melanin and produced between 400 mg/l and 900 mg/l. Further streaking out from submerged cultures and testing colonies of the last-mentioned type increased production to 1,000–1,100 mg/l. Additional work on medium development increased production to 3,350 mg/l [25]. An improved teicoplanin-producing strain of *Actinoplanes teichomyceticus* was obtained which produced pink instead of brown mycelia [26].

2.1.2 Nutritional ('auxotrophic') mutants

The end product of a biosynthetic pathway to a metabolite often exercises strict control over the amount of an intermediate accumulated by an auxotrophic mutant of that pathway. Only at a growth-limiting concentration of the end product would a large accumulation of the substrate of the deficient enzyme occur. This principle of decreasing the concentration of an inhibitory or repressive end product to bypass feedback inhibition or repression is best accomplished by the use of auxotrophic mutants. Production of secondary products is markedly affected by auxotrophic mutation, even when the auxotrophs are grown in nutritionally complete or even complex media. Although the change in product formation is usu-

ally in the negative direction, higher-producing auxotrophs are obtained from producers of antibiotics. For example, some leucine auxotrophs made more bacitracin than their prototrophic parent [27]. When several metabolites are produced by a single branched pathway, mutation in one branch of the pathway often leads to overproduction of the product of the other branch. In the case of branched pathways leading to a primary metabolite and a secondary metabolite, auxotrophic mutants requiring the primary metabolite sometimes overproduce the secondary metabolite [28].

Reversion of an auxotroph to prototrophy sometimes leads to new prototrophs possessing higher biosynthetic enzyme activity than present in the original 'grandparent' prototroph. Such increased enzyme activity is probably the result of a structural gene mutation producing a more active enzyme or an enzyme less subject to feedback inhibition. A revertant of an aspartate auxotroph of *Streptomyces fradiae* was found to be an improved producer of tylosin [29]. The auxotroph, lacking aspartate aminotransferase, produced no tylosin whereas the revertant produced 4-fold more than the grandparent strain. Similar examples have been described by Polsinelli and colleagues [30] dealing with actinomycin overproduction and by Saburova and co-workers [31] with respect to daunorubicin.

2.1.3 Revertants of non-producing mutants

Non-producing mutants sometimes produce increased amounts of secondary metabolites when mutated back to production. This has been observed in the cases of chlortetracycline [32] and aurodox [33].

2.1.4 Mutants resistant to antimetabolites

It is possible to select regulatory mutants, which overproduce end products of pathways, by the use of toxic metabolite analogs. Such antimetabolite-resistant mutants often possess enzymes that are insensitive to feedback inhibition, or enzyme-forming systems resistant to feedback repression.

Mutants of *Streptomyces clavuligerus* resistant to thialysine were found to have an aspartokinase resistant to covalent feedback inhibition by threonine and lysine, and 20% of these were hyperproducers (1.5 to 4-fold) of

cephalosporins [34]. Production of desferrioxamine, a commercial siderophore used for diseases associated with iron overload and pathological iron deposition, by *Streptomyces pilosus*, was markedly increased by mutagenesis and selection for resistance to thialysine [35]. Monensin A and B production by *Streptomyces cinnamonensis* was increased over 7-fold *via* deregulation of valine feedback inhibition by mutation to resistance to 2-ketobutyrate in the presence of valine or isoleucine [36]. Valine is the precursor of the fatty acid moiety of teicoplanin $A_{2\text{-}2}$ and its addition to the medium increases both total teicoplanin titer and the proportion of the desirable $A_{2\text{-}2}$ produced [37]. A mutant of *A. teichomyceticus* resistant to valine hydroxamate produced 50% more total teicoplanins and 50% more $A_{2\text{-}2}$ [38].

A variation of the antimetabolite selection techniques is possible when a precursor is toxic to the producing organism. The principle here is that the mutant most capable of detoxifying the precursor by incorporating it into the antibiotic will be the best grower in the presence of the precursor. Barrios-Gonzalez and co-workers [39] obtained phenylacetate-resistant cultures of *P. chrysogenum* which showed enhanced production of penicillin. Rifamycin B production was increased from 6 g/l to 19 g/l by sequential mutation to resistance to tryptophan (feedback inhibitor), *p*-hydroxybenzoate (analog of precursor) and propionate (precursor) [40]. When the produced secondary metabolite is itself a growth inhibitor of the producing culture, as in the case of certain antibiotics, the metabolite can sometimes be used to select resistant mutants which are improved producers. This has been reported with aurodox [33], nocardicin [41], leucomycin [42], oxytetracycline [43] and neomycin [43a].

Certain streptomycin resistance mutations cause increased production of unrelated antibiotics [44]. Selection for streptomycin-resistant mutants in wild-type *Streptomyces coelicolor* A3(2) led to 58% of the mutants producing 5- to 15-fold or more actinorhodin [45, 46]; similar results were obtained with *Streptomyces lividans*. Resistance to gentamicin and to paromomycin was reported to be even more effective than resistance to streptomycin [47]. Mutants resistant to combinations of two or all three of the antibiotics showed increased production up to 48-fold over wild-type [48]. The frequency of improved mutants varied between 5% and 18% of the total number of resistant mutants obtained. Selection of antibiotic-resistant mutants has been successfully used to increase titer of *Streptomyces* cultures even with high producers [49]. Starting with a strain of *Streptomy-*

ces albus producing 10 g/l of salinomycin, spontaneous Str^r, Gen^r or Rif^r mutants gave improved producers at a high rate (8–12%). The best strains made 12–15 g/l when 40 of each type were examined. Double mutants were generated by making spontaneous Gen^r mutants from one of the Str^r mutants; 7% were improved, the best producing 18 g/l. Triple mutants were made with rifamycin using the double mutant as parent; 8% were improved. The best made 23 g/l. It grew as well as the original parent and produced more abundant aerial mycelia and spores.

Nutritional repression can often be decreased by mutation to antimetabolite resistance. Examples of selective agents are 2-deoxyglucose (2-DOG) for enzymes and pathways controlled by carbon source regulation, methylammonium for those regulated by nitrogen source repression, and arsenate for phosphate regulation. Mutants of *P. chrysogenum* resistant to 2-DOG were found to be improved producers of penicillin [50]. A mutant of *Streptomyces aureofaciens* which used phosphate less efficiently for growth showed improved (60%) tetracycline production [51]. Mutants of *Streptomyces griseus* that make candicidin in the presence of levels of phosphate (that normally inhibit secondary metabolism) were found to be hyperproducers of the antibiotic even under low phosphate conditions [52].

2.1.5 Agar zone mutants

Fermentation performance on an agar plate is often related to production in submerged liquid culture and the method has application as a means of detecting superior mutants. 'Zone mutants' have proven useful for improved production of cephalosporin C [53], chlortetracycline [22, 54], mycobacillin [55] and penicillin [56]. A widely used modification involves the production of antibiotics by confluent growth atop separate plugs of agar followed by placement of these plugs on a seeded assay plate and measurement of the resultant clear zones. Use of this 'agar piece method' resulted in improvement of kasugamycin production from 0.5 g/l to 8 g/l [57], a 500-fold increase in bialaphos production [10], and improved cephalosporin C production [58]. Agar-piece screening of antibiotic production in the presence of inhibitory levels of phosphate (15 mM) led to isolation of six markedly improved and stable *Streptomyces hygroscopicus* strains producing the macrolide antifungal complex '165' [59].

2.2 Elimination of undesirable products or analogs

Since many organisms produce secondary metabolites as mixtures of a chemical family or of several chemical families, mutation has been used to eliminate undesirable products in such fermentations. As an example, streaking out of a natural single colony isolate from *S. aureofaciens* (producing the polyether narasin and the broad-spectrum antibiotic enteromycin) on galactose led to yellow and white sectoring [60]. The effect was specific for galactose. Of the four colony types obtained, one produced only narasin and two produced only enteromycin.

S. griseus subsp. *cryophilus* makes four R_3-sulfated and four R_3-unsulfated carbapenems. To completely eliminate the R_3 sulfated forms, that are less active than the unsulfated forms, sulfate transport mutants were obtained. These were of two types: (i) auxotrophs for thiosulfate or cysteine; and (ii) selenate-resistant mutants. Each type produced completely unsulfated forms and titers were equivalent to the total titer of the parent [61].

Eight avermectins are produced by *Streptomyces avermitilis* of which only a small number are desirable. A non-methylating mutant produced only four of the compounds and a mutant who failed to make the 25-isopropyl substituent (from valine) produced a different mixture of components. By protoplast fusion, a hybrid strain was obtained which made only two components, B2a and B1a [62]. Random PCR mutagenesis into gene *aveC* yielded a mutant which produced an avermectin B1:B2 ratio of 2.5, much improved over the 0.6 ratio of the parent *S. avermitilis* strain [63]. Gene shuffling further improved the ratio to about 15 [64].

Mutation was used to eliminate the undesirable polyketides sulochrin and asterric acid from broths of the lovastatin producer, *Aspergillus terreus* [65]. Mutants have also been employed to eliminate undesirable coproducts from the monensin fermentation [66].

2.3 Production of novel antibiotics

The medically useful demethyltetracycline and doxorubicin were discovered by simple mutation of the cultures producing tetracycline and daunomycin, respectively. Later, the technique of mutational biosynthesis ('mutasynthesis') was devised [67]. In this process, a mutant blocked in

secondary metabolism is fed analogs of the moiety whose biosynthesis is blocked. If successful, the mutant (called an 'idiotroph') produces a new antibiotic derivative [68]. Since then, mutational biosynthesis has been used for the discovery of many new secondary metabolites [69–71]. The most well-known is the commercial antihelmintic agent doramectin, production of which employed a mutant of the avermectin producer *S. avermitilis* [72].

New anthracyclines and aglycones have been isolated from blocked mutants of the daunorubicin and doxorubicin producers [73, 74]. By adding carminomycinone or 13-dihydrocarminomycinone to an idiotroph of *Streptomyces galilaeus* (the producer of aclacinomycin), the aglycones were glycosylated to form a new trisaccharide anthracycline, trisarubicionol, by mutational biosynthesis [75]. New macrolide antibiotics have been produced from blocked mutants of the tylosin-producer, *S. fradiae* [76]. Four new hybrid macrolide antibiotics were obtained by feeding erythronolide B to a blocked mutant of the oleandomycin producer, *Streptomyces antibioticus* [77]. A blocked-mutant of the mycinamicin producer, *Micromonospora polytrota*, was fed various rosaramicin precursors and converted them into new rosaramicins [78].

3 Deciphering biosynthetic pathways

The use of mutants for the elucidation of metabolic pathways been exploited for the biosynthesis of tetracyclines [79], novobiocin [80], erythromycin [81, 82], neomycin [83], tylosin [84], other aminoglycosides [85–89], rosaramicin [90], daunorubicin [73], other anthracyclines [91, 92], actinomycin [93], carbapenems [94, 95], ansamycins [96, 97], patulin [98] and phenazines [99].

4 Recombination

In contrast to the extensive use of mutagenesis and screening/selection in industry, employment of genetic recombination was meagre for many years. This lack of interest was prevalent, despite early claims of success [100, 101], mainly due to the total absence or the extremely low frequency

of genetic recombination in industrial microorganisms (in streptomycetes, it was usually 10^{-6} or even less). Other problems were evident with the β-lactam producing fungi. Although *Aspergillus* had sexual and parasexual reproduction, the most interesting genera from the commercial sense, *Cephalosporium* and *Penicillium*, only had parasexual reproduction that rarely resulted in recombination.

Recombination was erroneously looked upon as an alternative to mutation instead of a method which would complement mutagenesis programs. The most balanced and efficient strain development strategy would not emphasize one to the exclusion of the other; it would contain both mutagenesis-screening and recombination-screening components. In such a program, strains at different stages of a mutational line, or from lines developed from different ancestors, would be recombined. Such strains would no doubt differ in many genes and by crossing them, genotypes could be generated which would never occur as strictly mutational descendants of either parent.

Recombination has also been of importance in the mapping of production genes. Studies on the genetic maps of overproducing organisms such as actinomycetes are rather recent. The model for such investigations was the genetic map of *S. coelicolor* [102] which was found to be very similar to those of other *Streptomyces* species, such as *Streptomyces bikiniensis*, *Streptomyces olivaceous*, *Streptomyces glaucescens* and *Streptomyces rimosus*. The problem of low frequency of recombination was mainly solved by use of protoplast fusion. After 1980, there was a heightened interest in the application of genetic recombination to the production of important microbial products such as antibiotics. Today, frequencies of recombination have increased to even greater than 10^{-1} in some cases [103], and strain improvement programs routinely include protoplast fusion between different mutant lines. Conventional strain improvement increased clavulanate production 10-fold over the wild-type *S. clavuligerus* NRRL 3585. Then, protoplast fusion of arginine and cysteine auxotrophs yielded a fusant (CKD 1386) producing 30-fold more clavulanic acid than the wild-type [104]. Recombination is especially useful when combined with conventional mutation programs to solve the problem of 'sickly' organisms produced as a result of accumulated genetic damage over a series of mutagenized generations. For example, a cross *via* protoplast fusion was carried out with strains of *Acremonium chrysogenum*

(formerly *Cephalosporium acremonium*) from a commercial strain improvement program. A low-titer, rapidly-growing, spore-forming strain which required methionine to optimally produce cephalosporin C was crossed with a high-titer, slow-growing, asporogenous strain which could use the less expensive inorganic sulfate. The progeny included a recombinant which grew rapidly, sporulated, produced cephalosporin C from sulfate and made 40% more antibiotic than the high-titer parent [105]. Protoplast fusion was used to modify the characteristics of an improved penicillin-producing strain of *P. chrysogenum* which showed poor sporulation and poor seed growth. Backcrossing with a low-producing (12 g/l) strain yielded a high-producing (18 g/l) strain with better sporulation and better growth in seed medium [106]. Another application of protoplast fusion is the recombination of improved producers from a single mutagenesis treatment. By recombination, one could combine the yield-increase mutations and obtain an even more superior producer before carrying out further mutagenesis. Two improved cephamycin-C producing strains from *Nocardia* were fused and among the recombinants were two cultures which produced 10–15% more antibiotic than the best parent [107]. Genetic recombination allows the discovery of new antibiotics by fusing producers of different or even the same antibiotics. Protoplast fusion between non-producing mutants of a streptomycin producer (*S. griseus*) and an istamycin producer (*Streptomyces tenjimariensis*) led to a hybrid strain producing a new antibiotic [108]. A recombinant obtained from two different rifamycin-producing strains of *Nocardia mediterranei* produced two new rifamycins (16,17-dihydrorifamycin S and 16,17-dihydro-17-hydroxy-rifamycin S) [109]. However, according to Hopwood [110], these examples may reflect the different expression of genes from parent A in the cytoplasm of parent B rather than the formation of hybrid antibiotics. Interspecific protoplast fusion between S. griseus and five other species (*Streptomyces cyaneus*, *Streptomyce exfoliatus*, *Streptomyces griseoruber*, *Streptomyces purpureus* and *Streptomyces rochei*) yielded recombinants of which 60% produced no antibiotics and 24% produced antibiotics different from the parent strains [111]. New antibiotics can also be created by changing the order of the genes of an individual pathway in its native host [112]. A new antibiotic, indolizomycin, was produced by protoplast fusion between non-antibiotic producing mutants of *S. griseus* and *S. tenjimariensis* [113].

4.1 Transformation and transposition

Virtually all antibiotic-producing species of *Streptomyces* contain plasmid DNA. Some are sex plasmids and constitute an essential part of the sexual recombination process and others contain either structural genes or genes somehow influencing the expression of the chromosomal structural genes of antibiotic biosynthesis. Very few antibiotic biosynthesis processes are encoded by plasmid-borne genes. However, production of methylenomycin A is encoded by genes present on plasmid SCP1 in *S. coelicolor*. For many years, plasmid SCP1 was never observed nor isolated as a circular DNA molecule. The reason was that it was a giant linear plasmid. It was initially difficult to separate such giant linear plasmids from chromosomal DNA but this was later accomplished by pulsed field gel electrophoresis or orthogonal field alteration gel electrophoresis (OFAGE) [114]. When the plasmid was transferred to other streptomycetes, the recipients produced the antibiotic.

Cloning a 34 kb fragment from *S. rimosus via* a cosmid into *S. lividans* and *S. albus* resulted in oxytetracycline production by the recipients [115]. Contrary to earlier reports, all the oxytetracycline genes were clustered together on the *S. rimosus* chromosomal map. The biosynthetic genes were flanked by two resistance genes [116].

Plasmids have also been used to devise bioconconversions. 3-0-Acetyl-4"-*0*-isovaleryltylosin (AIV) is useful in veterinary medicine against tylosin-resistant *Staphylococcus aureus*. It is made by first producing tylosin with *S. fradiae* and then using *Streptomyces thermotolerans* (producer of carbomycin) to bioconvert tylosin into AIV. A new strain capable of carrying out a direct fermentation to AIV was constructed by transforming *S. fradiae* with *S. thermotolerans* plasmids containing acyl transferase genes [117].

Some products of unicellular bacteria are plasmid-encoded. These include aerobactin, a hydroxamate siderophore and virulence factor produced by *E. coli* [118] and other Gram-negative bacteria (*Enterobacter aerogenes*, *Enterobacter cloacae*, *Vibrio mimicus*, *Klebsiella*, *Salmonella*, and *Shigella*). Aerobactin is synthesized by a plasmid-borne five-gene cluster which is negatively regulated by iron [119], and also *via* chromosomal genes [120]. A microcin, an antimetabolite of methionine which is produced by *E. coli* and acts as a competitive inhibitor of homoserine-O-tran-

succinylase, is encoded by a plasmid which occurs at 20 copies per genome equivalent [121].

Instability in *Streptomyces* is brought about by environmentally stimulated macrolesions, e.g., deletions, transpositions, rearrangements and DNA amplification. They occur spontaneously or are induced by environmental stresses such as intercalating dyes, protoplast formation and regeneration, and interspecific protoplast fusion. Streptomycetes are the only prokaryotes known to be subject to spontaneous DNA amplification, sometimes amounting to several hundred tandem copies, accounting for over 10% of total DNA, in the absence of selection. Amplification seems to be coupled to DNA deletion and may involve insertion sequence (IS)-like elements [5]. Ethidium bromide cures plasmids in streptomycetes but also increases the frequency of deletion mutations especially in areas of the chromosome which are already unstable [122].

Transposable elements, i.e., DNA sequences encoding a transposase enzyme [123], that move from one replicon to another without host recombination functions or extensive homology with the site of integration, have been extremely useful for the following reasons: (i) they usually provide stable, nonreverting mutants; (ii) they can be used to determine the order of genes in an operon; (iii) it is easy to select for mutants since transposons contain antibiotic- or mercury-resistance markers; (iv) they provide portable regions of homology for chromosomal mobilization; (v) they provide markers for non-selectable genes and allow the cloning of such genes which can then be used as hybridization probes to fish out the wild type gene from a genomic library; and (vi) they often have unique restriction sites, thus are good markers for isolating defined deletion derivatives or locating the precise position of a gene by heteroduplex mapping.

In the daptomycin producer *Streptomyces roseosporus*, some Tn *5099* transposition mutants produced 57–66% more daptomycin than the parent whereas others produced less or the same as the parent [124, 125]. Transposition increased the rate-limiting step of tylosin biosynthesis in *S. fradiae*, i.e., the conversion of macrocin to tylosin. Transposing a second copy of *tylF* into a neutral site on the *S. fradiae* chromosome increased its gene product, macrocin O-methyltransferase, and tylosin production, while decreasing the concentration of the final intermediate (macrocin). Tylosin production was increased by up to 60% and total macrolide titer

was unchanged [126]. Transposon mutagenesis eliminated production of the troublesome toxic oligomycin by the avermectin-producing *S. avermitilis* [127].

5 Genetic engineering

One of the first indications that recombinant DNA technology could be applied to production of antibiotics and other secondary metabolites was that it could be carried out in streptomycetes [128]. Plasmids were constructed from plasmid SLP 1.2 of *S. lividans* and plasmid SCP2* from *S. coelicolor*. In mating of plasmid-negative *S. lividans*, 'pocks' (circular zones of sporulation inhibition associated with plasmid transfer in the lawn of streptomycete growth arising from a regenerated protoplast population) were seen. This was due to looping out of a piece of *S. coelicolor* DNA which became a series of small *S. lividans* plasmids (SLP 1.1 to 1.6) which were good cloning vehicles. Genetic engineering of actinomycetes was limited for a number of years by (i) restriction barriers hindering DNA introduction and (ii) inhibition of secondary metabolism by self-replicating plasmid-cloning vectors [129], but these problems were mainly overcome. Early reviews on cloning and expressing antibiotic production genes in *Streptomyces* were by Martin and Gil [130] and Liras [131].

An interesting possibility is the transfer of operons from one streptomycete to another in the hope that the structural genes might be better able to express themselves in another species. Clustering facilitates transfer of an entire pathway in a single manipulation. Studies revealed that many antibiotic biosynthesis genes were arranged in clusters including undecylprodigiosin, actinorhodin, chloramphenicol, rifamycin, cephamycin, erythromycin, tetracyclines and tylosin. Thus, the entire undecylprodigiosin pathway (*red* pathway) of *S. coelicolor* was transferred on a 37 kb fragment into *Streptomyces parvulus* and the antibiotic was produced [132]. Similarly, the entire cephamycin C pathway was cloned and expressed from a cephamycin-producing strain of *Streptomyces cattleya*. When the 29 kb DNA fragment was cloned into the non-β-lactam producer, *S. lividans*, one transformant (out of 30,000) made cephamycin [133]. When the fragment was introduced into another cephamycin producer, *Streptomyces lactamgens*, a 2 to 3-fold improvement in titer was obtained.

Cloning has been very important in understanding the biosynthesis of β-lactam antibiotics [134], its genetics, and improvement of the processes. The cloning of gene *pcbAB* from *P. chrysogenum* [135], *A. chrysogenum* [136], and *Nocardia lactamdurans* contributed greatly to the elucidation of the biosynthetic pathways. Gene *pcbC* (encoding isopenicillin N synthase) was cloned from *P. chrysogenum* [137], *A. chrysogenum*, *Aspergillus nidulans* [138], *S. clavuligerus* [139], *S. griseus* [140], *Streptomyces lipmannii* and *Streptomyces jumonjinensis* [138, 141] in order to provide pure enzyme for structural studies. Cloning multiple copies of this gene into *A. chrysogenum* yielded an improved cephalosporin C-producing strain [142].

Penicillin acyltransferase acts on IPN in *P. chrysogenum* to produce penicillin G. Its gene *penDE* (also known as *iat*, *aat* and *acyA* in *A. nidulans*) was cloned from *P. chrysogenum* into *A. chrysogenum* leading to production of penicillin G (in the presence of exogenous phenylacetic acid) along with cephalosporin C [136]. Without cloning, *A. chrysogenum* cannot produce penicillin G.

Transformation of early strain *P. chrysogenum* Wis54-1255 with individual genes, pairs of genes, and the entire three genes of the penicillin pathway showed that the major increases occurred when all three genes were overexpressed [143]. The best transformant contained three extra copies of *pcbAB*, one extra copy of *pcbC* and two extra copies of *penDE*, and produced 299% of control shake flask production and 276% of control productivity in continuous culture.

When an industrial production strain of *A. chrysogenum* was transformed with a plasmid containing genes *pcbC* and *cefEF* from an early strain of the *A. chrysogenum* mutant line, a transformant producing 50% more cephalosporin C than the production strain, as well as less penicillin N, was obtained. Production in pilot plant (150 liter) fermentors was further improved by 15% [144]. One copy of *cefEF* had been integrated into chromosome III whereas the native gene is on chromosome II.

An industrial strain improvement program based on genetic transformation showed that the best genes to increase cephalosporin C production in *A. chrysogenum* were *cefEF*, encoding expandase-hydroxylase, and *cefG*, encoding acetyltransferase [145]. The increased gene dosage raised production of cephalosporin C and decreased the concentration of intermediates deacetylcephalosporin and deacetoxycephalosporin C.

A. chrysogenum produces cephalosporin C but also excretes the intermediate DAOC at 1–2% of the cephalosporin C level. This undesirable situation was modified by genetically engineering the strain with two extra copies of the expandase-hydroxylase gene. The new strain excreted only half as much of this intermediate with no effect on cephalosporin C production [146].

Cloning of the benzylpenicillin acylase gene of *E. coli* on multicopy plasmids resulted in a 45-fold increase as compared to enzyme production by the uninduced wild-type. Interestingly, the cloned enzyme is constitutive [147]. Cloning of additional penicillin V amidase genes into wild-type *Fusarium oxysporium* increased enzyme titer by 130-fold [148]. These two enzymes are valuable for converting penicillin G into the valuable intermediate for making semi-synthetic penicillins, i.e., 6-aminopenicillanic acid (6-APA).

In another example, cephalosporin C was directly converted to 7-ACA by using an *E. coli* strain containing the D-amino acid oxidase gene from *Trigonopsis variabilis* and the glutaryl-7-aminocephalosporanic acid acylase gene from *Pseudomonas* sp. [149].

6 New strategies

In recent years, new techniques which markedly increase the options available to improve the production of microbial metabolites, have been added to the 'toolbox' including metabolic engineering, genomics, transcriptome analysis, proteomics, directed evolution, and whole genome shuffling, among others. These are described below.

'*Metabolic engineering*' is the directed improvement of product formation or cellular properties through the modification of specific biochemical reactions or introduction of new ones with the use of recombinant DNA technology [150, 151]. Its essence is the combination of analytical methods to quantify fluxes and the control of fluxes with molecular biological techniques to implement suggested genetic modifications. Metabolic control analysis has revealed that the overall flux through a metabolic pathway depends on several steps, not just a single rate-limiting reaction [152]. This strategy has been succesfully applied to antibiotic production [153–155], including complex polyketides [156, 157].

Progress in metabolic engineering to improve production of microbial metabolites has involved three major strategies: (i) manipulation of structural genes; (ii) manipulation of regulatory genes; and (iii) engineering of well-known pathways.

(i) *Manipulation of structural genes*. Amplification of an entire pathway resulted in a 2.3-fold increase in cephamycin C production [133]. Amplification of a pathway segment yielded a 7-fold increase in formation of daunorubicin formation [158], a 30% increase in tetracenomycin C production [159], and 3 to 4-fold more spinosyn [160]. Perturbing central metabolism by deleting either of the genes encoding the first two enzymes in the pentose phosphate pathway led to increased actinorhodin production in *S. lividans* [161].

(ii) *Manipulation of regulatory genes*. The gene *afsR2* is a global regulatory gene originally discovered in *S. lividans* with positive control action on actinorhodin and undecylprodigiosin formation; it is also present in *S. coelicolor* [162]. Incorporation of multiple copies of *afsR2* from *S. lividans* into *S. avermitilis* wildtype increased avermectin production by 2.3-fold [163]. Another regulatory gene appears to be an 8 kb DNA fragment of *S. avermitilis* which stimulates actinorhodin and undecylprodigiosin formation in *S. lividans*, as well as avermectin production in wild-type *S. avermitilis*, in an improved strain, and in a semi-industrial strain of *S. avermitilis* [164]. Amplification of other pathway-specific regulators resulted in a 5-fold enhancement of spiramycin titer in *Streptomyces ambofaciens* [165] and a 1.6-fold increase in mithramycin production in *Streptomyces argillaceus* [166].

Disruption of negative pathway specific regulatory genes improved production of avermectin by 1.5 to 3.5-fold [164], lovastatin by 7 to 10-fold [167], and also improved formation of methylenomycin [163] and mitomycin C [168]. Also, disruption of a negative global regulator resulted in production increases for actinorhodin and undecylprodigiosin [169]. Gene *nysF*, encoding a putative 4'-phosphopantetheinyl transferase in the nystatin producer, *Streptomyces noursei*, was first thought to be involved in a biosynthetic step but was actually found to be a negative regulatory gene [170]. Knocking out *nysF* increased production by 60%.

(iii) *Engineering of pathways*. By increasing the expression of rate-limiting enzymes, tylosin production was enhanced [171, 172] and so was penicillin. In this latter case, overexpression of the *acvA* gene by replacing

the normal promoter with the ethanol dehydrogenase promoter increased penicillin production in *A. nidulans* by 30-fold [173].

Eliminating the accumulation and excretion of pathway intermediate penicillin N raised cephalosporin C production by 15%. Weak acetyltransferase promoter activity appears to be the cause of the undesirable accumulation of intermediate DAC in broths of *A. chrysogenum*. Cloning of gene *cefG* (encoding DAC acetyltransferase) increased cephalosporin C titers in a dose-dependent manner [173a, 173b]. Cloning of the gene with its own promoter had no effect on the low level of DAC acetyltransferase normally observed in *A. chrysogenum* [174]. However, use of foreign promoters (the *gpd* promoter from *A. nidulans*, the *bla* promoter from *Aspergillus niger*, or the *pbcC* promoter from *P. chrysogenum*) had a major effect on the level of *cefG* transcripts, DAC acetyltransferase protein level and activity, and antibiotic production; cephalosporin C production rose 2- to 3-fold. Of the cephalosporins produced, the undesirable DAC decreased from 80% of the total down to 30–39%, whereas cephalosporin C increased by a similar amount.

Transformation of *P. chrysogenum* with the *S. lipmanii cefD* gene and the *S. clavuligerus cefE* gene allowed production of the intermediate deacetoxycephalosporin C (DAOC) [175] at titers of 2.5 g/l. DAOC is a valuable intermediate in the commercial production of semi-synthetic cephalosporins.

Metabolic engineering of *P. chrysogenum* allowed production of valuable cephalosporin intermediates (adipyl-7-ADCA and adipyl-7-ACA) [176]. Disruption and one-step replacement of the *cefEF* gene of an industrial cephalosporin C production strain of *A. chrysogenum* yielded strains accumulating up to 20 g/l of penicillin N. Cloning and expression of the *cefE* gene from *S. clavuligerus* into those high producing strains yielded recombinant strains producing high titers of DAOC [177]. Production levels were nearly equivalent (80%) to the total β-lactams biosynthesized by the parental strain.

Deletion of a gene leading to a side-product eliminated oligomycin production from the producer of avermectins [127]. Genetically increasing oxygen availability resulted in a 60% improvement of erythromycin titer [178, 179]. Enhancement of precursor uptake gave a 4-fold increase in production of deoxyerythonolide B and 8,8a-deoxyoleoandolide [180]).

'*Inverse Metabolic Engineering*' [181, 182] (also known as 'reverse engineering') involves construction or calculation of a desired phenotype, identifi-

cation of the molecular basis of that desirable property, and incorporation of that phenotype into another strain or other species by genetic manipulations and funtional genomics [183]. Once the differences are found, they can be introduced into other organisms or their dosages increased in the improved strain to further improve it. The techniques available are (i) genomic DNA sequencing; (ii) DNA sequencing of selected genes; (iii) transcriptional profiling; (iv) proteome analysis; (v) metabolite profiling; and (vi) comparative flux analysis. This strategy was used to increase erythromycin production by *Aeromicrobium erythreum* [184]. Plasmid insertion mutagenesis yielded 26 mutants (out of 3,049 isolates) with >50% increased production; seven were chosen for reverse engineering. Three were identical *mutB* mutants encoding the α subunit of methylmalonyl-CoA mutase. The other four mutants were in *cobA*, encoding Cob(I)alamin adenosyl transferase. The *mutB* mutant produced about 500 mg/l of erythromycin compared to ca 200 mg/l made by the parent. The *cobA* mutant produced 22% more than the parent. Feeding of nutrients such as leucine and propionate, which increase pools of propionyl-CoA and methylmalonyl-CoA, to the mutants increased production. Methionine also increased production by the *cobA* strain. The probable reason for improved production by methylmalonate-CoA mutase knockouts is that its substrate is a branchpoint intermediate and the favored direction is the production of succinyl-CoA, thus diverting methylmalonyl-CoA away from the erythromycin pathway. Since coenzyme B_{12} is its cofactor, *cobA* mutants would also favor flux into the erythromycin pathway. As expected, leucine, propionate and methionine, which are known precursors of methylmalonyl-CoA *via* propionyl-CoA, stimulated production in the mutants.

'*Genome-based strain reconstruction*' allows one to construct a strain superior to the production strain because it only contains mutations crucial to hyperproduction, but not other unknown mutations which accumulate by brute-force mutagenesis and screening. Although this approach has been done with a primary metabolite production process (lysine) [185], it should also be useful for secondary metabolites.

During the last few years, an expanded view of the cell has been possible due to the impressive advances in all the 'omics' techniques (genomics, proteomics, metabolomics) and high-throughput technologies for measuring different classes of key intracellular molecules. '*Systems Biology*' has recently emerged as a term and a scientific field to describe an approach

that considers genome-scale and cell-wide measurements in elucidating process and mechanisms [186].

Integrating transcriptional and metabolite profiles from 21 strains of *A. terreus* producing different levels of lovastatin and another 19 strains with altered (+)-geodin levels led to an improvement in lovastatin production of over 50% [187]. This approach, named '*Association Analysis*', served to reduce the complexity of profiling data sets in order to identify those genes whose expression is most tightly linked to metabolite production. Such applications are applicable to all industrially useful organisms for which genome data are limited.

Gene expression analysis of wild type and improved production strains of *Saccharopolyspora erythraea* and *S. fradiae* using microarrays of the sequenced *S. coelicolor* revealed that regulation of antibiotic biosynthetic enzymes as well as enzymes involved in precursor metabolism were altered in those mutated strains [188]. Comparison of *S. erythraea* wild type and an improved erythromycin producer was done and results revealed that the *S. erythraea* overproducer expressed the entire erythromycin gene cluster several days longer than the wild-type. It seems that the *eryA* gene and protein expression differences observed for the overproducer could account over 50% of the total erythromycin titer increase. A different situation was found with tylosin production. The overproducing *S. fradiae* strain produced tylosin for the same length of time but the rate of antibiotic production was 2-fold higher in the overproducer. The tylosin cluster was expressed over the same period of time; however, there were two genes: *aco* (encoding acyl-CoA dehydrogenase) and *icmA* (encoding isobutyryl-CoA mutase) that were expressed to a greater extent than in the wild type strain. The induction of these two genes could increase the flux of metabolites from fatty acids to tylosin precursors in the overproducer.

'*Directed Evolution*' (also known as applied molecular evolution or directed molecular evolution) is a fast and inexpensive way of finding variants of existing enzymes that work better than naturally occurring enzymes under specific conditions [189–191]. It has been used to improve the properties of rate-limiting enzymes of natural product biosynthesis. The process involves evolutionary design methods using random mutagenesis, gene recombination and high throughput screening [192]. A key limitation of these strategies is that they introduce random 'noise' mutations into the gene at every cycle and hence improvements are limited to

small steps. This strategy has been successfully used in various applications [193].

'*Molecular Breeding*™' techniques come closer to mimicking natural recombination by allowing *in vitro* homologous recombination [194]. DNA shuffling not only recombines DNA fragments but also introduces point mutations at a very low controlled rate [195, 196]. Unlike site directed mutagenesis, this method of pooling and recombining parts of similar genes from different species or strains has yielded remarkable improvements in enzymes in a very short amount of time [197]. A step forward in this technique was breeding a population with high genetic variability as a starting point to generate diversity (DNA Family Shuffling). Innovations that expand the formats for generating diversity by recombination include formats similar to DNA shuffling and others with few or no requirements for parental gene homology [198, 199]. These random redesign techniques are being currently used to increased biological activity of protein pharmaceuticals and biological molecules [197, 198] as well as novel vaccines [200, 201].

'*Whole genome shuffling*' is a novel technique that combines the advantage of multi-parental crossing allowed by DNA shuffling with the recombination of entire genomes. This method was successfully applied to improve tylosin production in *S. fradiae* [202]. Historically, twenty cycles of classical strain improvement at Eli Lilly and Co. carried out over 20 years employing about one million assays improved production 6-fold. In contrast, two rounds of this recursive recombination technique with seven early strains each were sufficient to achieve similar results in one year and involved only 24,000 assays.

'*Combinatorial biosynthesis*' is not used to increase titer but for discovery of new and modified drugs [203, 204]. In this strategy, recombinant DNA techniques are utilized to introduce genes coding for antibiotic synthases into producers of other antibiotics or into non-producing strains to obtain modified or hybrid antibiotics. In the first demonstration of this technology, gene transfer from a streptomycete strain producing the isochromanequinone antibiotic actinorhodin into strains producing granaticin, dihydrogranaticin and mederomycin (which are also isochromanequinones) led to the discovery of two new antibiotic derivatives, mederrhodin A and dihydrogranatirhodin [205]. Since then, hundreds of new polyketides have been made by combinatorial biosynthesis [206–228]. Some of these novel

polyketides contain sugars at normally unglycosylated positions [229] or as new sugar moieties [230, 231]. New anthracyclines [232–237] and peptide antibiotics [238] have been made by combinatorial biosynthesis. Manipulations include [239]: (i) deletion of one of the domains of a particular module; (ii) addition of a copy of the thioesterase domain to the end of an earlier module resulting in a shortened polyketide; (iii) replacement of an AT domain of a polyketide synthase (PKS) with an AT domain from another PKS, resulting in addition of a methyl group at a particular site or removal of a methyl group; (iv) addition of a reductive domain(s) to a particular module, thus changing a keto group to a double bond or to a methylene group; (v) use of synthetic diketides delivered as N-acetylcysteamine thioesters to load onto the active site of the ketosynthase (KS) in module 2 and to be taken all the way to a novel final product; (vi) replacement of the loading module of one PKS with the loading module of another PKS, thus changing the starter unit from propionate to acetate, for example; and (vii) replacement of the hydroxylase or glycosylase enzymes from one pathway to another, thus modifying the ring structure with respect to OH groups and/or sugars.

Progress in strain development will depend, not only on all the technologies mentioned above, but also on the development of mathematical methods that facilitate the elucidation of mechanisms and identification of genetic targets for modification. The availability of the complete genome sequences of *S. coelicolor*, producer of actinorhodin and three other antibiotics [240], and *S. avermitilis*, producer of avermectins [241], the published and ongoing sequencing projects involving hundreds of microbial genomes, as well as the ability to rapidly identify clusters of genes encoding biosynthesis of bioactive products and to predict their structures based on gene sequences [242] will contribute to an acceleration of strain improvement programs.

Closing comments

Microorganisms produce many compounds of industrial interest. These may be very large materials such as proteins, nucleic acids, carbohydrate polymers, or even cells, or they can be smaller molecules which can be essential for vegetative growth (primary metabolites) or inessential (sec-

ondary metabolites). The power of the microbial culture in the competitive world of commercial synthesis can be appreciated by the fact that even simple molecules are made by fermentation rather than by chemical synthesis. Most natural products are so complex that they probably will never be made commercially by chemical synthesis. Strains isolated from nature produce only tiny amounts of product. The reason is that they need small amounts of these compounds for their own competitive benefit (survival) in nature; they do not overproduce these metabolites. Regulatory mechanisms have evolved in microorganisms that enable a strain to avoid excessive production of its metabolites. Thus, strain improvement programs are absolutely required for commercial application. Their goal is to isolate cultures exhibiting desired phenotypes. Most commonly, the ability of a strain to improve titer is what is desired, although the spectrum of improvements can also include other traits. The tremendous increases in fermentation productivity and the resulting decreases in costs have come about mainly by mutagenesis and screening/selection. In recent years, recombinant DNA technology has contributed greatly. The promise of the future is *via* extensive use of new genetic techniques such as (i) metabolic engineering accomplishing quantification and control of metabolic fluxes and including inverse metabolic engineering, (ii) transcript expression analyses, (iii) directed evolution, (iv) molecular breeding, and (v) combinatorial biosynthesis. These efforts will facilitate not only the isolation of improved strains but also the elucidation and identification of new genetic targets to be used in strain improvement programs.

References

1 Parekh S (2000) Strain improvement. In: J Lederberg (ed): *Encyclopedia of microbiology*, vol. 4, 2nd ed. Academic Press, San Diego, 428–443

2 Vinci VA, Byng G (1999) Strain improvement by nonrecombinant methods. In: AL Demain, JE Davies (eds): *Manual of industrial microbiology and biotechnology*, 2nd ed. ASM Press, Washington, DC, 103–113

3 Parekh S, Vinci VA, Strobel RJ (2000) Improvement of microbial strains and fermentation processes. *Appl Microbiol Biotechnol* 54: 287–301

4 Simpson IN, Caten CE (1979) Induced quantitative variation for penicillin titre in clonal populations of *Aspergillus nidulans*. *J Gen Microbiol* 110: 1–12

5 Baltz RH (1986) Mutagenesis in *Streptomyces*. In: AL Demain, NA Solomon (eds): *Manual of industrial microbiology and biotechnology*. ASM, Washington, DC, 184–190

6 Baltz RH (1995) Gene expression in recombinant *Streptomyces*. In: A Smith (ed): *Gene expression in recombinant microorganisms*. Marcel Dekker, New York, 309–381

7 Baltz RH (1998) Genetic manipulation of antibiotic producing *Streptomyces*. *Trends Microbiol* 6: 76–83

8 Baltz RH (1999) Mutagenesis. In: MC Flickinger, SW Drew (eds): *Encyclopedia of bioprocessing technology: fermentation, biocatalysis, and separation*. Wiley, New York, 1819–1822

9 Keller U (1983) Highly efficient mutagenesis of *Claviceps purpurea* by using protoplasts. *Appl Environ Microbiol* 46: 580–584

10 Takebe H, Imai S, Ogawa H, Satoh A, Tanaka H (1989) Breeding of bialaphos producing strains from a biochemical engineering viewpoint. *J Ferm Bioeng* 67: 226–232

11 Kim KS, Cho NY, Pai HS, Ryu DDY (1983) Mutagenesis of *Micromonospora rosaria* by using protoplasts and mycelial fragments. *Appl Environ Microbiol* 46: 689–693

12 Kurzatkowski W, Kurylowicz W, Solecka J, Penyige A (1986) Improvement of *Streptomyces* strains by the regeneration of protoplasts. In: G Szabo, S Biro, M Goodfellow (eds): *Biological, biochemical and biomedical aspects of actinomycetes. Part A*. Academiai Kiado, Budapest, 289–292

13 Strohl WR (2001) Biochemical engineering of natural product biosynthesis pathways. *Metab Eng* 3: 4–14

14 Hersbach GJM, van der Beck CP, van Dijck PWM (1984) The penicillins: properties, biosynthesis, and fermentation. In: EJ Vandamme (ed): *Biotechnology of industrial antibiotics*. Marcel Dekker, New York, 45–140

15 Smith DJ, Bull JH, Edwards J, Turner G (1989) Amplification of the isopenicillin N synthetase gene in a strain of *Penicillium chrysogenum* producing high levels of penicillin. *Mol Gen Genet* 216: 492–497

16 Barredo JL, Diez B, Alvarez E, Martin JF (1989) Large amplification of a 35-kb DNA fragment carrying two penicillin biosynthetic genes in high penicillin producing strains of *P. chrysogenum*. *Curr Genet* 16: 453–459

17 Smith DJ, Burnham MKR, Bull JH, Hodgson JE, Ward JM, Browne P, Brown J, Barton B, Earl AJ, Turner G (1990) β-Lactam antibiotic biosynthesis genes have been conserved in clusters in prokaryotes and eukaryotes. *EMBO J* 9: 741–747

18 Fierro F, Barredo JL, Diez B, Gutierrez S, Fernandez FJ, Martin JF (1995) The penicillin gene cluster is amplified in tandem repeats linked by conserved hexanucleotide sequences. *Proc Natl Acad Sci USA* 92: 6200–6204

19 Podojil M, Blumauerova M, Culik K, Vanek Z (1984) The tetracycyclines: properties, biosynthesis and fermentation. In: EJ Vandamme (ed): *Biotechnology of industrial antibiotics*. Marcel Dekker, New York, 259–279

20 Demain AL, Elander RP (1999) The β-lactam antibiotics: past, present, and future. *Ant v Leeuwenhoek* 75: 5–19

21 Chen W-Q, Yu ZN, Zheng Y-H (2004) Expression of *Vitreoscilla* hemoglobin gene in *Streptomyces fradiae* and its effect on cell growth and synthesis of tylosin. *Chin J Antibiot* 29: 516–520

22 Liu C-M (1982) Microbial aspects of polyether antibiotics: activity, production and biosynthesis. In: JW Westley (ed): *Polyether antibiotics. Naturally occurring acid ionophores. Vol. 1. Biology*. Marcel Dekker, New York, 43–102

23 Blumauerova M, Pokorny V, Stastna J, Hostalek Z, Vanek Z (1978) Developmental mutants of *Streptomyces coeruleorubidis*, a producer of anthracyclines: isolation and preliminary characterization. *Folia Microbiol* 23: 177–182

24 Blumaerova M, Podojil M, Gauze G., Maksikmova TS, Panos J, Vanek Z (1980) Spontaneous variability of *Streptomyces glomeratus*, a producer of anthracycline antibiotics beromycins. *Folia Microbiol* 25: 207–212

25 Blumaerova M, Podojil M, Gauze GF, Maksikmova TS, Panos J, Vanek Z (1980) Effect of cultivation conditions on the activity of the beromycin producer *Streptomyces glomeratus* 3980 and its spontaneous variants. *Folia Microbiol* 25: 213–218

26 Lee JC, Park HR, Park DJ, Son KH, Yoon KH, Kim YB, Kim CJ (2003) Production of teicoplanin by a mutant of *Actinoplanes teicomyceticus*. *Biotechnol Lett* 25: 537–540

27 Haavik HI, Froyshov O (1982) On the role of L-leucine in the control of bacitracin formation by *Bacillus licheniformis*. In: H Kleinkauf, H von Doehren (eds): *Peptide antibiotics: biosynthesis and functions*. Walter de Gruyter, Berlin, 155–159

28 Godfrey O (1973) Isolation of regulatory mutants of the aspartic and pyruvic acid families and their effect on antibiotic production in *Streptomyces lipmanii*. *Antimicrob Agents Chemother* 4: 73–79

29 Lee SH, Lee KJ (1995) Threonine dehydratases in different strains of *Streptomyces fradiae*. *J Biotechnol* 43: 95–102

30 Polsinelli M, Albertini A, Cassani G, Ciferri O (1965) Relation of biochemical mutations to actinomycin synthesis in *Streptomyces antibioticus*. *J Gen Microbiol* 39: 239–246

31 Saburova TP, Lapchinskaya OA, Sinyagina OP, Trutneva EM, Lvova NA (1977) Auxotrophic mutants in selection of *Actinomyces coeruleorubidus*, an organism producing rubomycin. *Antibiotiki* 22: 1095–1100

32 Dulaney EL, Dulaney DD (1967) Mutant populations of *Streptomyces viridifaciens*. *Trans NY Acad Sci II* 29: 782–799

33 Unowsky J, Hoppe DC (1978) Increased production of the antibiotic aurodox (X–5108) by aurodox-resistant mutants. *J Antibiot* 31: 662–666

34 Mendelovitz S, Aharonowitz Y (1983) β-Lactam antibiotic production by *Streptomyces clavuligerus* mutants impaired in regulation of aspartokinase. *J Gen Microbiol* 129: 2063–2069

35 Smith A (1987) Enzyme regulation of desferrioxamine biosynthesis: a basis for a rational approach to process development. In: M Alacevic, D Hranueli, Z Toman (eds): *Fifth International Symposium on the Genetics of Industrial Microorganisms*, 1986. Pliva, Zagreb, 513–527

36 Pospisil S, Kopecky J, Prikrylova V, Spizek J (1999) Overproduction of 2-ketoisovalerate and monensin production by regulatory mutants of *Streptomyces cinnamonensis* resistant to 2-ketobutyrate and amino acids. *FEMS Microbiol Lett* 172: 197–204

37 Wang MR, Ding H, Hu YJ (1996) The action of arginine and valine in the biosynthesis of teicoplanin. *Chin J Antibiot* 21 (Suppl): 77–80

38 Jin ZH, Wang MR, Cen PL (2002a) Production of teichoplanin by valine-analogue resistant mutant strains of *Actinoplanes teichomyceticus*. *Appl Microbiol Biotechnol* 58: 63–66

39 Barrios-Gonzalez J, Montenegro E, Martin JF (1993) Penicillin production by mutants resistant to phenylacetic acid. *J Ferm Bioeng* 7: 455–458

40 Jin ZH, Lin JP, Xu ZN, Cen PL (2002b) Improvement of industry-applied rifamycin B-producing strain, *Amycolatopsis mediterranei*, by rational screening. *J Gen Appl Microbiol* 48: 329–334

41 Elander RP, Aoki H (1982) β-Lactam producing microorganisms – their biology and fermentation behavior. In: RB Morin, M Gorman (eds): *Chemistry and biology of* β*-lactam antibiotics*, vol. 3. Academic Press, New York, 83–153

42 Higashide E (1984) The macrolides; properties, biosynthesis, and fermentation. In: EJ Vandamme (ed): *Biotechnology of industrial antibiotics*. Marcel Dekker, New York, 451–509

43 Gravius B, Glocker D, Pigac J, Pandza K, Hranueli D, Cullum J (1994) The 387 kb linear plasmid pPZG101 of *Streptomyces rimosus* and its interactions with the chromosome. *Microbiology* 140: 2271–2277

43a Crameri R, Davies JE (1986) Increased production of aminoglycosides associated with amplified antibiotic resistance genes. *J Antibiot* 39: 128–135

44 Hosoya Y, Okamoto S, Muramatsu H, Ochi K (1998) Acquisition of certain streptomycin resistant (*str*) mutations enhances antibiotic production in bacteria. *Antimicrob Agents Chemother* 42: 2041–2047

45 Hesketh A, Ochi K (1997) A novel method for improving *Streptomyces coelicolor* A3(2) for production of actinorhodin by introduction of rpsL (encoding ribosomal protein S12) mutations conferring resistance to streptomycin. *J Antibiot* 50: 532–535

46 Okamoto S, Lezhava A, Hosaka T, Okamoto-Hosoya Y, Ochi K (2003) Enhanced expression of S-adenosylmethionine synthetase causes overproduction of actinorhodin in *Streptomyces coelicolor* A3(2). *J Bacteriol* 185: 601–609

47 Okamoto-Hosoya Y, Sato T, Ochi K (2000) Resistance to paromomycin is conferred by rpsL mutations, accompanied by an enhanced antibiotic production in *Streptomyces coelicolor* A3(2). *J Antibiot* 53: 1424–1427

48 Hu H, Ochi K (2001) Novel approach for improving the production of antibiotic-producing strains by inducing combined resistant mutations. *Appl Environ Microbiol* 67: 1885–1892

49 Tamehiro N, Hosaka T, Xu J, Hu H, Otake N, Ochi K (2003) Innovative approach for improvement of an antibiotic-overproducing industrial strain of *Strepromyces albus*. *Appl Environ Microbiol* 69: 6412–6417

50 Chang LT, McGrory EL, Elander RP (1980) Penicillin production by glucose-derepressed mutants of *Penicillium chrysogenum*. *J Indust Microbiol* 6: 165–169

51 Colombo AL, Crespi-Pevellino N, Grein A, Minghetti A, Spalla CJ (1981) Metabolic and genetic aspects of the relationship between growth and tetracycline production in *Streptomyces* aureofaciens. *Biotechnol Lett* 3: 71–76

52 Martin JF, Naharro G, Liras P, Villanueva JR (1979) Isolation of mutants deregulated in phosphate control of candicidin biosynthesis. *J Antibiot* 32: 600–606

53 Elander RP (1969) Applications of microbial genetics to industrial fermentations. In: D Perlman (ed): *Fermentation Advances*. Academic Press, New York, 89–114

54 Katagiri I (1954) Study on chlortetracycline. Improvement of chlortetracycline-producing strain by several kinds of methods. *J Antibiot* 7: 45–52

55 Bannerjee AB, Bose SK (1964) A rapid method for isolating mutants of *Bacillus subtilis* producing increased or decreased amounts of the antibiotic, mycobacillin. *J Appl Bacteriol* 27: 93–95

56 Ditchburn P, Giddings B, MacDonald KD (1974) Rapid screening for the isolation of mutants of *Aspergillus nidulans* with increased penicillin yields. *J Appl Bacteriol* 37: 515–523

57 Ichikawa T, Date M, Ishikura T, Ozaki A (1971) Improvement of kasugamycin-producing strain by agar piece method and the prototroph method. *Folia Microbiol* 16: 218–224

58 Chang LT, Elander RP (1979) Rational selection for improved cephalosporin C productivity in strains of *Acremonium chrysogenum*. *Devel Indust Microbiol* 20: 367–379

59 Gesheva V (1994) Isolation of spontaneous *Streptomyces hygroscopicus* 111-81 phosphate-deregulated mutants hyperproducing its antibiotic complex. *Biotechnol Lett* 16: 443–448

60 Nakatsukasa WM, Mabe JA (1978) Galactose induced colonial dissociation in *Streptomyces aureofaciens*. *J Antibiot* 31: 805–808

61 Kitano K, Nozaki Y, Imada A (1985) Selective accumulation of unsulfated carbapenem antibiotics by sulfate transport-negative mutants of *Streptomyces griseus* subsp. *cryophilus* C-19393. *Agric Biol Chem* 49: 677–684

62 Omura S, Ikeda H, Tanaka H (1991) Selective production of specific components of avermectins in *Streptomyces avermitilis*. *J Antibiot* 44: 560–563

63 Stutzman-Engwall K, Conlon S, Fedechko R, Kaczmarek F, McArthur H, Krebber A, Chen Y, Minshull J, Raillard SA, Gustafsson C (2003). Engineering the *aveC* gene to enhance the ratio of doramectin to its CHC-B2 analogue produced in *Streptomyces avermitilis*. *Biotechnol Bioeng* 82: 359–369

64 del Cardayre SB (2003) Delivering industrial microorganisms through gene, pathway, and genome shuffling. *Abstr S75, SIM Ann Mtg Prog & Abstr* 73

65 Vinci VA, Hoerner TD, Coffman AD, Schimmel TG, Dabora RL, Kirpekar AC, Ruby CL, Stieber RW (1991) Mutants of a lovastatin-hyperproducing *Aspergillus terreus* deficient in the production of sulochrin. *J Indust Microbiol* 8: 113–120

66 Pospisil S, Peterkova M, Krumphanzl V, Vanek Z (1984) Regulatory mutants of *Streptomyces cinnamonensis* producing monensin A. *FEMS Microbiol Lett* 24: 209–213

67 Shier WT, Rinehart KL Jr, Gottlieb D (1969) Preparation of four new antibiotics from a mutant of *Streptomyces fradiae*. *Proc Natl Acad Sci USA* 63: 198–204

68 Nagaoka K, Demain AL (1975) Mutational biosynthesis of a new antibiotic, streptomutin A, by an idiotroph of *Streptomyces griseus*. *J Antibiot* 28: 627–635

69 Lemke JR, Demain AL (1976) Preliminary studies on streptomutin A. *Eur J Appl Microbiol* 2: 91–94

70 Daum SJ, Lemke JR (1979) Mutational biosynthesis of new antibiotics. *Ann Rev Microbiol* 33: 241–265

71 Kitamura S, Kase H, Odakura Y, Iida T, Shirahata K, Nakayama K (1982) 2-Hydroxysagamicin: A new antibiotic produced by mutational biosynthesis of *Micromonospora sagamiensis*. *J Antibiot* 35: 94–97

72 Cropp TA, Wilson DJ, Reynolds KA (2000) Identification of a cyclohexylcarbonyl CoA biosynthetic gene cluster and application in the production of doramectin. *Nature Biotechnol* 18: 980–983

73 McGuire J, Thomas MC, Pandey RC, Toussaint M, White RJ (1981) Biosynthesis of daunorubicin glycosides: analysis with blocked mutants. In: M Moo-Young (ed): *Advances in biotechnology. Vol. III. Fermentation products*. Pergammon Press, New York, 117–122

74 Cassinelli G, Di Matteo F, Forenza S, Ripamonti M, Rivola G, Arcamone F, Di Marco A, Cassaza AM, Soranzo C, Pratesi G (1980) New anthracycline glycosides from *Micromonospora*. II. Isolation, characterization and biological properties. *J Antibiot* 33: 1468–1473

75 Yoshimoto A, Matsuzawa Y, Matsuhashi Y, Oki T, Takeuchi T, Umezawa H (1981) Trisarubicinol, new antitumor anthracycline antibiotic. *J Antibiot* 34: 1492–1494

76 Kirst HA, Wild GM, Baltz RH, Seno ET, Hamill RL, Paschal JW, Dorman DE (1983) Elucidation of structure of novel macrolide antibiotics produced by mutant strains of *Streptomyces fradiae*. *J Antibiot* 36: 376–382

77 Spagnoli R, Cappalletti L, Toscano L (1983) Biological conversion of erythronolide B,

an intermediate of erythromycin biogenesis, into new 'hybrid' macrolide antibiotics. *J Antibiot* 36: 365–375

78 Lee BK, Puar MS, Patel M, Bartner P, Lotvin J, Munayyer H, Waitz JA (1983) Multistep bioconversion of 20-deoxo-20-dihydro-12, 13-deepoxy-12,13-dehydrorosaranolide to 22-hydroxy-23-o-mycinosyl-20-deoxo-20-dihydro-12,13-deepoxy-rosaramicin. *J Antibiot* 36: 742–743

79 McCormick JRD (1965) Biosynthesis of the tetracyclines. In: Z Vanek, Z Hostalek (eds): *Biogenesis of antibiotic substances*. Publishing House of the Czechoslovak Academy of Science, Prague, 73–91

80 Kominek LA (1972) Biosynthesis of novobiocin by *Streptomyces niveus*. *Antimicrob Agents Chemother* 1: 123–134

81 Martin JR, Perun TJ, Girolami RL (1966) Studies on the biosynthesis of erythromycins. I. Isolation of an intermediate glycoside, 3-alpha-L-mycarosylerythronolide B. *Biochemistry* 5: 2852–2856

82 Martin JR, Rosenbrook WR (1967) Studies on the biosynthesis of erythromycins. II. Isolation and structure of a biosynthetic intermediate, 6-deoxyerythronolide B. *Biochemistry* 6: 435–440

83 Pearce CJ, Akhtar M, Barnett JEG, Mercier D, Sepulchre AM, Gero S (1978) Sub-unit assembly in the biosynthesis of neomycin. The synthesis of 5-O-β-D-ribofuranosyl and 4-O-β-D-ribofuranosyl-2,6-dideoxystreptamines. *J Antibiot* 31: 74–81

84 Baltz RH, Seno ET, Stonesifer J, Wild GM (1983) Biosynthesis of the macrolide antibiotic tylosin: a preferred pathway from tylactone to tylosin. *J Antibiot* 36: 131–141

85 Takeda K, Aihara K, Furumai T, Ito Y (1978) An approach to the biosynthetic pathway of butirosins and the related antibiotics. *J Antibiot* 31: 250–253

86 Fujiwara T, Takahashi Y, Matsumoto K, Kondo E (1980) Isolation of an intermediate of 2-deoxystreptamine biosynthesis from a mutant of *Bacillus circulans*. *J Antibiot* 33: 824–829

87 Kase H, Odakura Y, Nakayama K (1982) Sagamycin and the related aminoglycosider: fermentation and biosynthesis. I. Biosynthetic studies with the blocked mutants of *Micromonospora sagamiensis*. *J Antibiot* 35: 1–9

88 Hanssen R, Kirby R (1983) The induction by N-methyl-N'-nitro-nitrosoguanidine of multiple closely linked mutations in *Streptomyces bikiniensis* ISP5235 affecting streptomycin resistance and streptomycin biosynthesis. *FEMS Microbiol Lett* 17: 317–320

89 Penzikova GA, Levitov MM (1970) A study on transamidinase activity with respect to streptomycin biosynthesis. *Biologiya* 39: 337–342

90 Vaughn RV, Lotvin J, Puar MS, Patel M, Kershner A, Kalyanpur MG, Marquez J, Waitz JA (1982) Isolation and characterization of two 16-membered lactones: 20-deoxorosaramicin and 20-deoxo-12,13-desepoxy-12,13-dehydrorosaramicin aglycones, from a mutant strain of *Micromonospora rosaria*. *J Antibiot* 35: 251–253

91 Motamedi H, Wendt-Pientowski E, Hutchinson CR (1986) Isolation of tetracenomycin C non-producing *Streptomyces glaucescens* mutants. *J Bacteriol* 167: 575–580

92 Yue S, Motamedi H, Wendt-Pienskowski E, Hutchinson CR (1986) Anthracycline metabolites of tetracenomycin-non-producing *Streptomyces glaucescens*. *J Bacteriol* 167: 581–586

93 Troost T, Katz E (1979) Phenoxazinone biosynthesis: accumulation of a precursor, 4-methyl-3-hydroxyanthranilic acid, by mutants of *Streptomyces parvulus*. *J Gen Microbiol* 1: 121–132

94 Nozaki Y, Kitano K, Imada I (1984) Blocked mutants in the biosynthesis of carbapenem antibiotics from *Streptomyces griseus* subsp. *cryophilus*. *Agric Biol Chem* 48: 37–44

95 Kojima I, Fukagawa Y, Okabe M, Ishikura T, Shibamoto N (1988) Mutagenesis of OA-6129 carbapenem-producing blocked mutants and the biosynthesis of carbapenems. *J Antibiot* 41: 899–907

96 Kibby JJ, McDonald IA, Rickards RW (1980) 3-Amino-5-hydroxybenzoic acid as a key intermediate in ansamycin and maytansinoid biosynthesis. *J Chem Soc Chem Comm* 768–769

97 Ghisalba O, Fuhrer H, Richter W, Moss S (1981) A genetic approach to the biosynthesis of the rifamycin-chromophore in *Nocardia mediterranei*. III. Isolation and identification of an early ansamycin-precursor containing the seven-carbon amino starter-unit and three initial acetate/propionate-units of the ansa chain. *J Antibiot* 34: 58–63

98 Gaucher GM, Lam KS, Grootwassink JWD, Neway J, Deo YM (1981) The initiation and longevity of patulin biosynthesis. *Devel Indust Microbiol* 22: 219–232

99 Byng GS, Eustice DC, Jensen RA (1979) Biosynthesis of phenazine pigments in mutant and wild-type cultures of *Pseudomonas aeruginosa*. *J Bacteriol* 138: 846–852

100 Jarai M (1961) Genetic recombination in *Streptomyces aureofaciens*. *Acta Microbiol Acad Sci Hung* 8: 73–79

101 Mindlin SZ (1969) Genetic recombination in the actinomycete breeding. In: G Sermonti, M Alacevic (eds): *Genetics and breeding of Streptomyces*. Yugoslav Academy of Sciences and Arts, Zagreb, 147–159

102 Kieser HM, Kieser T, Hopwood DA (1992) A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. *J Bacteriol* 174: 5496–5507

103 Ryu DDY, Kim KS, Cho NY, Pai HS (1983) Genetic recombination in *Micromonospora rosaria* by protoplast fusion. *Appl Environ Microbiol* 45: 1854–1858

104 Choi KP, Kim KH, Kim JW (1997) Strain improvement of clavulanic acid producing *Streptomyces clavuligerus*. *Abstr. 12P9, 10th Internat Symp Biol Actinomycetes (Beijing)*

105 Hamlyn PF, Ball C (1979) Recombination studies with *Cephalosporium acremonium*. In: OK Sebek, AI Laskin (eds): *Genetics of industrial microorganisms*. ASM, Washington, DC, 185–191

106 Lein J (1986) The Panlabs penicillin strain improvement program. In: Z Vanek, Z Hostalek (eds): *Overproduction of microbial metabolites; strain improvement and process control strategies*. Butterworth Publishers, Boston, 105–139

107 Wesseling AC, Lago B (1981) Strain improvement by genetic recombination of cephamycin producers, *Nocardia lactamdurans* and *Streptomyces griseus*. *Devel Indust Microbiol* 22: 641–651

108 Yamashita F, Hotta K, Kurasawa S, Okami Y, Umezawa H (1982) Antibiotic formation by interspecific protoplast fusion in streptomycetes and emergence of drug resistance by protoplast regeneration. *Abstr P-II-20, 4th Internatl Symp Genet Indust Microorgs, (Kyoto):* 108

109 Traxler P, Schupp T, Wehrli W (1982) 16,17-dihydrorifamycin S and 16,17-dihydro-17-hydroxyrifamycin S, two novel rifamycins from a recombinant strain C5/42 of *Nocardia mediterranei*. *J Antibiot* 35: 594–601

110 Hopwood DA (1983) Actinomycete genetics and antibiotic production. In: LC Vining (ed): *Biochemistry and genetic regulation of commercially important antibiotics*. Addison Wesley, Reading, MA, 1–23

111 Okanishi M, Suzuki N, Furuta T (1996) Variety of hybrid characters among recombinants

obtained by interspecific protoplast fusion in streptomycetes. *Biosci Biotechnol Biochem* 60: 1233–1238

112 Hershberger CL (1996) Metabolic engineering of polyketide biosynthesis. *Curr Opin Biotechnol* 7: 560–562

113 Gomi S, Ikeda D, Nakamura H, Naganawa H, Yamashita F, Hotta K, Kondo S, Okami Y, Umezawa H, Iitaka Y (1984) Isolation and structure of a new antibiotic, indolizomycin, produced by a strain SK2-52 obtained by interspecies fusion treatment. *J Antibiot* 37: 1491–1494

114 Kinashi H, Shimaji M (1987) Detection of giant linear plasmids in antibiotic producing strains of *Streptomyces* by the OFAGE technique. *J Antibiot* 40: 913–916

115 Binnie C, Warren M, Butler MJ (1989) Cloning and heterologous expression in *Streptomyces lividans* of *Streptomyces rimosus* genes involved in oxytetracycline biosynthesis. *J Bacteriol* 171: 887–895

116 Butler MJ, Friend EJ, Hunter IS, Kaczmarek FS, Sugden DA, Warren M (1989) Molecular cloning of resistance genes and architecture of a linked gene cluster involved in biosynthesis of oxytetracycline by *Streptomyces rimosus*. *Molec Gen Genet* 215: 231–238

117 Arisawa A, Kawamura N, Narita T, Kojima I, Okamura K, Tsunekawa H, Yoshioka T, Okamoto R (1996) Direct fermentative production of acyltylosins by genetically-engineered strains of *Streptomyces fradiae*. *J Antibiot* 49: 349–354

118 McDougall S, Neilands JB (1984) Plasmid- and chromosome-coded aerobactin synthesis in enteric bacteria: insertion sequences flank operon in plasmid-mediated systems. *J Bacteriol* 159: 300–305

119 Roberts M, Leavitt RW, Carbonetti NH, Ford S, Cooper RA, Williams PH (1986) RNA-DNA hybridization analysis of transcription of the plasmid ColV-K30 aerobactin gene cluster. *J Bacteriol* 167: 467–472

120 Moon YH, Tanabe T, Funahashi T, Shiuchi K, Nakao H, Yamamoto S (2004) Identification and characterization of two contiguous operons required for aerobactin transport and biosynthesis in *Vibrio mimicus*. *Microbiol Immunol* 48: 389–398

121 Perez-Diaz JC, Clowes RC (1980) Physical characterization of plasmids determining synthesis of a microcin which inhibits methionine synthesis in *Escherichia coli*. *J Bacteriol* 141: 1015–1023

122 Crameri R, Davies JE, Huetter R (1986) Plasmid curing and generation of mutations induced with ethidium bromide in Streptomycetes. *J Gen Microbiol* 132: 819–824

123 Berg DE, Berg CM (1983) The prokaryotic transposable element Tn5. *Bio/Technology* 1: 417–435

124 McHenney MA, Baltz RH (1996) Gene transfer and transposition mutagenesis in *Streptomyces roseosporus*: mapping of insertions that influence daptomycin or pigment production. *Microbiology* 142: 2363–2373

125 Baltz RH, McHenney MA, Cantwell CA, Queener SW, Solenberg PJ (1997) Applications of transposition mutagenesis in antibiotic producing streptomyces. *Ant v Leeuwenhoek* 71: 179–187

126 Solenberg PJ, Cantwell CA, Tietz AJ, McGilvray D, Queener SW, Baltz RH (1996) Transposition mutagenesis in *Streptomyces fradiae*: identification of a neutral site for the stable insertion of DNA by transposon exchange. *Gene* 16: 67–72

127 Ikeda H, Takada Y, Pang CH, Tanaka H, Omura S (1993) Transposon mutagenesis by Tn4560 and applications with avermectin-producing *Streptomyces avermitilis*. *J Bacteriol* 175: 2077–2082

128 Thompson CJ, Ward JM, Hopwood DH (1982) Cloning of antibiotic resistance and nutritional genes in streptomycetes. *J Bacteriol* 151: 668–677
129 Baltz RH, Hosted TJ (1996) Molecular genetic methods for improving secondary-metabolite production in actinomycetes. *Trends Biotechnol* 14: 245–250
130 Martin JF, Gil JA (l984) Cloning and expression of antibiotic production genes. *Bio/Technology* 2: 63–72
131 Liras P (1988) Cloning of antibiotic biosynthetic genes. In: JA Thompson (ed): *Use of recombinant DNA techniques for improvement of fermentation organisms*. CRC Press, Boca Raton, 217–253
132 Coco EA, Narva KE, Feitelson JS (1991) New classes of *Streptomyces coelicolor* A3(2) mutants blocked in undecylprodigiosin (Red) biosynthesis. *Mol Gen Genet* 227: 28–32
133 Chen CW, Lin HF, Kuo CL, Tsai HL, Tsai JFY (1988) Cloning and expression of a DNA sequence conferring cephamycin C production. *Bio/Technology* 6: 1222–1224
134 Demain AL, Elander RP (1999) The β-lactam antibiotics: past, present, and future. *Ant v Leeuwenhoek* 75: 5–19
135 Diez B, Gutierrez S, Barredo JL, van Solinger P, van der Voort LHM, Martin JF (1990) The cluster of penicillin biosynthetic genes. Identification and characterization of the *pcbAB* gene encoding the γ-aminoadipyl-cysteinyl-valine synthetase and linkage to the *pcbC* and *pcbDE* genes. *J Biol Chem* 265: 16358–16365
136 Gutierrez S, Diez B, Alvarez E, Barredo JL, Martin JF (1991) Expression of the *pen*DE gene of *Penicillium chrysogenum* encoding isopenicillin N acyltransferase in *Cephalosporium acremonium*: production of benzyl penicillin by the transformants. *Mol Gen Genet* 225: 56–64
137 Carr LG, Skatrud PL, Sheetz ME, Queener SW, Ingolia TD (1986) Cloning and expression of the isopenicillin N synthetase gene from *Penicillium chrysogenum*. *Gene* 48: 257–266
138 Weigel BJ, Burgett SG, Chen VJ, Skatrud PL, Frolik CA, Queener SW, Ingolia TD (1988) Cloning and expression in *Escherichia coli* of isopenicillin N synthetase genes from *Streptomyces lipmanii* and *Aspergillus nidulans*. *J Bacteriol* 170: 3817–3826
139 Leskiw BK, Aharonowitz Y, Mevarech M, Wolfe S, Vining LC, Westlake DWS, Jensen SE (1988) Cloning and nucleotide sequence determination of the isopenicillin N synthetase gene from *Streptomyces clavuligerus*. *Gene* 62: 187–196
140 Garcia-Dominguez M, Liras P, Martin JF (1991) Cloning and characterization of the isopenicillin N synthase gene of *Streptomyces griseus* NRRL 3851 and studies of expression and complementation of the cephamycin pathway in *Streptomyces clavuligerus*. *Antimicrob Agents Chemother* 35: 44–52
141 Shiffman D, Mevarech M, Jensen SE, Cohen G, Aharonowitz Y (1988) Cloning and comparative analysis of the gene encoding isopenicillin N synthase in *Streptomyces*. *Mol Gen Genet* 214: 562–569
142 Skatrud PL, Fisher DL, Ingolia TD, Queener SW (1987) Improved transformation of *Cephalosporium acremonium*. In: M Alacevic, D Hranueli (eds): *Genetics of industrial microorganisms, part B*. Pliva, Zagreb, 111–119
143 Theilgaard HA, van den Berg MA, Mulder CA, Bovenberg RAL, Nielsen J (2001) Quantitative analysis of *Penicillium chrysogenum* Wis54-1255 transformants overexpressing the penicillin biosynthetic genes. *Biotechnol Bioeng* 72: 379–388
144 Skatrud PL, Tietz AJ, Ingolia TD, Cantwell CA, Fisher DL, Chapman JL, Queener SW (1989) Use of recombinant DNA to improve production of cephalosporin C by *Cephalosporium acremonium*. *Bio/Technology* 7: 477–485

145 Rodriguez-Saiz M, Lembo M, Bertetti L, Muraca R, Velasc, J, Malcangi A, Luis de la Fuente J, Barredo JL (2004) Strain improvement for cephalosporin production by *Acremonium chrysogenum* using geneticin as a suitable transformation marker. *FEMS Microbiol Lett* 235: 43–49
146 Basch J, Chiang S-JD (1998) Genetic engineering approach to reduce undesirable byproducts in cephalosporin C fermentation. *J Indust Microbiol Biotechnol* 20: 344–353
147 Mayer H, Collins J, Wagner F (1980) Cloning of the penicillin G-acylase gene of *Escherichia coli* ATCC 11105 on multicopy plasmids. In: HH Weetall, GP Royer (eds): *Enzyme engineering*, vol. 5. Plenum, New York, 61–69
148 Elander RP (1995) Genetic engineering applications in the development of selected industrial enzymes and therapeutic proteins. In: R Sankaran, KS Manja (eds): *Microbes for better living*. Defense Food Research Laboratory, Mysore, India, 619–628
149 Luo H, Yu H, Li Q, Shen Z (2004) Cloning and co-expression of D-amino acid oxidase and glutaryl-7-aminocephalosporanic acid acylase genes in *Escherichia coli*. *Enzyme Microb Technol* 35: 514–580
150 Stephanopoulos G (1999) Metabolic fluxes and metabolic engineering. *Metab Eng* 1: 1–11
151 Nielsen J (2001) Metabolic engineering. *Appl Microbiol Biotechnol* 55: 263–283
152 Kacser H, Acerenza L (1993) A universal method for achieving increases in metabolite production. *Eur J Biochem* 216: 361–367
153 Khetan A, Hu WS (1999) Metabolic engineering of antibiotic biosynthetic pathways. In: AL Demain, JE Davies (eds): *Manual of industrial microbiology and biotechnology*. ASM Press, Washington, DC, 717–724
154 Khetan A, Hu W-S (1999) Metabolic engineering of antibiotic biosynthesis for process improvement. In: SY Lee, ET Papoutsakis (eds): *Metabolic engineering*. Marcel Dekker, New York, 177–202
155 Thykaer J, Nielsen J (2003) Metabolic engineering of β-lactam production. *Metab Eng* 5: 56–69
156 Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, Khosla C (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *Escherichia coli*. *Science* 291: 1790–1792
157 Khosla C, Keasling JD (2003) Metabolic engineering for drug discovery and development. *Nature Rev/Drug Disc* 2: 1019–1025
158 Otten SL, Stutzman-Engwall J, Hutchinson CR (1990) Cloning and expression of daunorubicin biosynthesis genes from *Streptomyces peucetius* and *S. peucetius* subsp. *caisius*. *J Bacteriol* 172: 3427–3434
159 Decker H, Summers RG, Hutchinson CR (1994) Overproduction of the acyl carrier protein component of a type II polyketide synthase stimulates production of tetracenomycin biosynthetic intermediates in *Streptomyces glaucescens*. *J Antibiot* 47: 54–63
160 Madduri K, Waldron C, Matsushima P, Broughton MC, Crawford K, Merlo DJ, Baltz RH (2001) Genes for the biosynthesis of spinosyns: applications for yield improvement in *Saccharopolyspora spinosa*. *J Indust Microbiol Biotechnol* 27: 399–402
161 Butler MJ, Bruheim P, Jovetic S, Marinelli F, Postma PW, Bibb MJ (2002) Engineering of primary carbon metabolism for improved antibiotic production in *Streptomyces lividans*. *Appl Environ Microbiol* 68: 4731–4739
162 Voegtli M, Chang PC, Cohen SN (1994) *afsR2*: a previously undetected gene encoding

a 63-amino acid protein that stimulates antibiotic production in *Streptomyces lividans*. *Mol Microbiol* 14: 643–653

163 Lee JY, Hwang YS, Kim SS, Kim ES, Choi CY (2000) Effect of a global regulatory gene, afsR2, from *Streptomyces lividans* on avermectin production in *Streptomyces avermitilis*. *J Biosci Bioeng* 89: 606–608

164 Hwang YS, Kim ES, Biro S, Choi CY (2003) Cloning and analysis of a DNA fragment stimulating avermectin production in various *Streptomyces avermitilis* strains. *Appl Environ Microbiol* 69: 1263–1269

165 Geistlich M, Losick R, Turner JR, Rao RN (1992) Characterization of a novel regulatory gene governing the expression of a polyketide synthase gene in *Streptomyces ambofaciens*. *Mol Microbiol* 6: 2019–2029

166 Lombo F, Brana AF, Mendez C, Salas JA (1999) The mithramycin gene cluster of *Streptomyces argillaceus* contains a positive regulatory gene and two repeated DNA sequences that are located at both ends of the cluster. *J Bacteriol* 181: 642–647

167 Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC, Hutchinson CR (1999) Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* 284: 1368–1372

168 Mao Y, Varoglu M, Sherman DH (1999) Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* NRRL 2564. *Chem Biol* 6: 251–263

169 Brian P, Riggle PJ, Santos RA, Champness WC (1996) Global negative regulation of *Streptomyces coelicolor* antibiotic synthesis mediated by an *absA*-encoded putative transduction system. *J Bacteriol* 178: 3221–3231

170 Volokhan O, Skletta H, Sekurova ON, Ellingsen TE, Zotchev SB (2005) An unexpected role for the putative 4′-phosphopantetheinyl transferase-encoding gene *nysF* in the regulation of nystatin biosynthesis in *Streptomyces noursei* ATCC 11455. *FEMS Microbiol Lett* 249: 57–64

171 Cox KL, Fishman SE, Larson JL, Stanzak R, Reynolds PA, Yeh WK, Van Frank RM, Birmingham VA, Hershberger CL, Seno ET (1987) Cloning and characterization of genes involved in tylosin biosynthesis. In: M Alacevic, D Hranueli, Z Toman (eds): *Genetics of industrial microorganisms, part B*. Pliva, Zagreb, 337–346

172 Fishman SE, Cox K, Larson JL, Reynolds PA, Seno ET, Yeh WK, Van Frank R, Hershberger CL (1987) Cloning genes for the biosynthesis of a macrolide antibiotic. *Proc Natl Acad Sci USA* 84: 8248–8252

173 Kennedy J, Turner G (1996) δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase is a rate limiting enzyme for penicillin production in *Aspergillus nidulans*. *Mol Gen Genet* 253: 189–197

173a Matsuda A, Sugiura H, Matsuyama K, Matsumoto H, Ichikawa S, Komatsu KI (1992) Molecular cloning of acetyl coenzyme A: deacetylcephalosporin C O-acetyltransferase cDNA from *Acremonium chrysogenum*: sequence and expression of catalytic activity in yeast. *Biochem Biophys Res Commun* 182: 995–1001

173b Mathison L, Soliday C, Stepan T, Aldrich T, Rambosek J (1993) Cloning, characterization, and use in strain improvement of the *Cephalosporium acremonium* gene cefG encoding acetyl transferase. *Curr Genet* 23: 33–41

174 Gutierrez S, Velasco J, Marcos AT, Fernandez FJ, Fierro F, Barredo JL, Diez B, Martin JF (1997) Expression of the *cefG* gene is limiting for cephalosporin biosynthesis in *Acremonium chrysogenum*. *Appl Microbiol Biotechnol* 48: 606–614

175 Cantwell C, Beckmann R, Whiteman P, Queener SW, Abraham EP (1992) Isolation of deacetoxycephalosporin C from fermentation broths of *Penicillium chrysogenum* transformants: construction of a new fungal biosynthetic pathway. *Proc R Soc Lond (Biol)* 248: 283–289

176 Crawford L, Stepan AM, Mcada PC, Rambosek JA, Conder MJ, Vinci VA, Reeves CD (1995) Production of cephalosporin intermediates by feeding adipic acid to recombinant *Penicillium chrysogenum* strains expressing ring expansion activity. *Bio/Technology* 13: 58–62

177 Velasco J, Adrio JL, Moreno MA, Diez B, Soler G, Barredo JL (2000) Environmentally safe production of 7-aminodeacetoxycephalosporanic acid (7-ADCA) using recombinant strains of *Acremonium chrysogenum*. *Nature Biotechnol* 18: 857–861

178 Brunker P, Minas W, Kallio PT, Bailey JE (1998). Genetic engineering of an industrial strain of *Saccharopolyspora erythrea* for stable expression of the *Vitreoscilla* hemoglobin gene (*vhb*). *Microbiology* 144: 2441–2448

179 Minas W, Brunker P, Kallio PT, Bailey JE (1998) Improved erythromycin production in a genetically engineered industrial strain of *Saccharopolyspora erythraea*. *Biotechnol Prog* 1: 561–566

180 Lombo F, Pfeifer B, Leaf T, Ou S, Kim YS, Cane DE, Licari P, Khosla C (2001) Enhancing the atom economy of polyketide biosynthetic processes through metabolic engineering. *Biotechnol Prog* 17: 612–617

181 Bailey JE, Sburlati A, Hatzimanikatis V, Lee K, Renner WA, Tsai PS (1996) Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnol Bioeng* 52: 109–121

182 Khosla C, Bailey JE (1988) Heterologous expression of a bacterial haemoglobin improves the growth properties of recombinant *Escherichia coli*. *Nature* 331: 633–635

183 Bro C, Nielsen J (2004) Impact of 'ome' analyses on inverse metabolic engineering. *Metab Eng* 6: 204–211

184 Reeves AR, Cernota WH, Brikun IA, Wesley RK, Weber JM (2004) Engineering precursor flow for increased erythromycin production in *Aeromicrobium erythreum*. *Metab Eng* 6: 300–312

185 Ohnishi J, Mitsuhashi S, Hayashi M, Ando S, Yokoi H, Ochiai K, Ikeda M (2002) A novel methodology employing *Corynebacterium glutamicum* genome information to generate a new L-lysine-producing mutant. *Appl Microbiol Biotechnol* 58: 217–223

186 Stephanopoulos G, Alper H, Moxley J (2004) Exploiting biological complexity for strain improvement through systems biology. *Nature Biotechnol* 22: 1261–1267

187 Askenazi M, Driggers EM, Holtzman DA, Norman TC, Iverson S, Zimmer DP, Boers ME, Blomquist PR, Martinez EJ, Monreal AW et al (2003) Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nature Biotechnol* 21: 150–156

188 Lum AM, Huang J, Hutchinson CR, Kao CM (2004) Reverse engineering of industrial pharmaceutical-producing actinomycete strains using DNA microarrays. *Metab Eng* 6: 186–196

189 Skandalis A, Encell LP, Loeb LA (1997) Creating novel enzymes by applied molecular evolution. *Chem Biol* 4: 889–898

190 Kuchner O, Arnold FH (1997) Directed evolution of enzyme catalysis. *Trends Biotechnol* 15: 523–530

191 Arnold FH (1998) Design by directed evolution. *Acc Chem Res* 31: 125–131

192 Arnold FH (2001) Combinatorial and computational challenges for biocatalyst design. *Nature* 409: 253–257
193 Zhao H, Chockalingam K, Chen Z (2002) Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr Opin Biotechnol* 13: 104–110
194 Ness JE, Del Cardayre SB, Minshull J, Stemmer WP (2000) Molecular breeding: the natural approach to protein design. *Adv Prot Chem* 55: 261–292
195 Stemmer WP (1994) Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* 370: 389–391
196 Zhao H, Arnold FH (1997) Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res* 25: 1307–1308
197 Patten PA, Howard RJ, Stemmer W (1997) Applications of DNA shuffling to pharmaceuticals and vaccines. *Curr Opin Biotechnol* 8: 724–733
198 Kurtzman AL, Govindarajan S, Vahle K, Jones JT, Heinrichs V, Patten PA (2001) Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. *Curr Opin Biotechnol* 12: 361–370
199 Lutz S, Ostermeier M, Moore GL, Maranas CD, Benkovic SP (2001) Creating multiple-crossover DNA libraries independent of sequence identity. *Proc Natl Acad Sci USA* 98: 11248–11253
200 Marshall SH (2002) DNA shuffling: induced molecular breeding to produce new generation long-lasting vaccines. *Biotechnol Adv* 20: 229–238
201 Locher CP, Soong NW, Whalen RG, Punnonen J (2004) Development of novel vaccines using DNA shuffling and screening strategies. *Curr Opin Mol Ther* 6: 34–39
202 Zhang YX, Perry K, Vinci VA, Powell K, Stemmer WPC, del Cardayre SB (2002) Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* 415: 644–646
203 Hutchinson C (1998) Combinatorial biosynthesis for new drug discovery. *Curr Opin Microbiol* 1: 319–329
204 Reeves CD (2003) The enzymology of combinatorial biosynthesis. *Crit Rev Biotechnol* 23: 95–147
205 Hopwood DA, Malpartida F, Kieser HM, Ikeda H, Duncan J, Fujii I, Rudd BAM, Floss, HG, Omura S (1985) Production of 'hybrid' antibiotics by genetic engineering. *Nature* 314: 642–644
206 Rodriguez E, McDaniel R (2001) Combinatorial biosynthesis of antimicrobials and other natural products. *Curr Opin Microbiol* 4: 526–534
207 Donadio S, Sosio M (2003) Strategies for combinatorial biosynthesis with modular polyketide synthases. *Comb Chem High Throughput Screen* 6: 489–500
208 Kantola J, Kunnari T, Mantsala P, Ylihonko K (2003) Expanding the scope of aromatic polyketides by combinatorial biosynthesis. *Comb Chem High Throughput Screen* 6: 501–512
209 McAlpine JB, Tuan JS, Brown DP, Grebner KB, Whittern DN, Buko A, Katz L (1987) New antibiotics from genetically engineered actinomycetes. I. 2-Norerythromycins, isolation and structural determinations. *J Antibiot* 40: 1115–1122
210 Epp JK, Huber MLB, Turner JR, Goodson T, Schoner BE (1989) Production of hybrid macrolide antibiotic in *Streptomyces ambofaciens* and *Streptomyces lividans* by introduction of a cloned carbomycin biosynthetic gene from *Streptomyces thermotolerans*. *Gene* 85: 293–301
211 Weber JM, Leung JO, Swanson SJ, Idler KB, McAlpine JB (1991) An erythromycin deriva-

tive produced by targeted gene disruption in *Saccharopolyspora erythraea*. *Science* 252: 114–117

212 Donadio S, Staver MJ. McAlpine JB, Swanson SJ, Katz L (1991) Modular organization of genes required for complex polyketide biosynthesis. *Science* 252: 675–679

213 Donadio S, McAlpine JB, Sheldon PA, Jackson MA, Katz L (1993) An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc Natl Acad Sci USA* 90: 7119–7123

214 Hara O, Hutchinson CR (1992) A macrolide 3-O-acyltransferase gene from the mide-camycin-producing species *Streptomyces mycarofaciens*. *J Bacteriol* 174: 5141–5144

215 Katz L, Donadio S (1993) Polyketide synthesis: prospects for hybrid antibiotics. *Annu Rev Microbiol* 47: 875–912

216 Hopwood DA (1993) Genetic engineering of *Streptomyces* to create hybrid antibiotics. *Curr Opin Biotechnol* 4: 531–537

217 McDaniel R, Ebert-Khosla S, Hopwood D, Khosla C (1993) Engineered biosynthesis of novel polyketides. *Science* 262: 1546–1550

218 McDaniel R, Ebert-Khosla S, Hopwood D, Khosla C (1993) Engineered biosynthesis of novel polyketides: manipulation and analysis of an aromatic polyketide synthase with unproven catalytic specificities. *J Am Chem Soc* 115: 11671–11675

219 McDaniel R, Thamchaipenet A, Gustafsson C, Fu H, Betlach M, Betlach M, Ashley G (1999) Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel 'unnatural' natural products. *Proc Natl Acad Sci USA* 96: 1846–1851

220 Khosla C, McDaniel R, Ebert-Khosla S, Torres R. Sherman DH, Bibb MJ, Hopwood DA (1993) Genetic construction and functional analysis of hybrid polyketide synthases containing heterologous acyl carrier proteins. *J Bacteriol* 175: 2197–2204

221 Decker H, Hutchinson CR (1993) Transcriptional analysis of the *Streptomyces glaucescens* tetracenomycin biosynthesis gene cluster. *J Bacteriol* 175: 3887–3892

222 Hutchinson CR, Fujii I (1995) Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Annu Rev Microbiol* 49: 201–238

223 Tsoi CJ, Khosla C (1995) Combinatorial biosynthesis of unnatural natural products – the polyketide example. *Chem Biol* 2: 355–362

224 Kao CM, Luo GL, Katz L, Cane DE, Khosla C (1995) Manipulation of macrolide ring size by directed mutagenesis of a modular polyketide synthase. *J Am Chem Soc* 117: 9105–9106

225 Pacey MS, Dirlam JP, Geldart RW, Leadlay PF, McArthur HA, McCormick EL, Monday RA, O'Connell TN, Staunton J, Winchester TJ (1998) Novel erythromycins from a recombinant *Saccharopolyspora erythraea* strain NRRL 2338pIG1. I. Fermentation, isolation and biological activity. *J Antibiot* 51: 1029–1034

226 Wohlert SE, Blanco G, Lombo F, Fernandez E, Brana AF, Reich S, Udvarnoki G, Mendez C, Decker H, Frevert J et al (1998) Novel hybrid tetracenomycins through combinatorial biosynthesis using a glycosyltransferase encoded by the elm-genes in cosmid 16F4 and which shows a very broad sugar substrate specificity. *J Amer Chem Soc* 120: 10596–10601

227 Xue Q, Hutchinson CR, Santi DV (1999) A multi-plasmid approach to preparing large libraries of polyketides. *Proc Natl Acad Sci USA* 96: 11740–11745

228 Pfeifer BA, Khosla C (2001) Biosynthesis of polyketides in heterologous hosts. *Microbiol Molec Biol Rev* 65: 106–118

229 Trefzer A, Blanco G, Remsing L, Kunzel E, Rix U, Lipata F, Brana AF, Mendez C, Rohr J, Bechtold A et al (2002) Rationally designed glycosylated premithramycins: hybrid aromatic polyketides using genes fom three different biosynthetic pathways. *J AmChem Soc* 124: 6056–6062

230 Zhao L, Ahlert J, Xue Y, Thorson JS, Sherman DH, Liu H-W (1999) Engineering a methymycin/pikromycin-calicheamicin hybrid: construction of two new macrolides carrying a designed sugar moiety. *J Am Chem Soc* 121: 9881–9882

231 Mendez C, Salas JA (2001) Altering the glycosylation pattern of bioactive compounds. *Trends Biotechnol* 19: 449–456

232 Bartel PL, Zhu CB, Lampel JS, Dosch DC, Connors NC, Strohl WR, Beale JM Jr, Floss HG (1990) Biosynthesis of anthraquinones by interspecies cloning of actinorhodin biosynthesis genes in streptomycetes; clarification of actinorhodin gene functions. *J Bacteriol* 172: 4816–4826

233 Strohl WR, Bartel PL, Li Y, Connors NC, Woodman RH (1991) Expression of polyketide biosynthesis and regulatory genes in heterologous streptomycetes. *J Indust Microbiol* 7: 163–174

234 Hwang CK, Kim HS, Hong YS, Kim YH, Hong SK, Kim SJ, Lee JJ (1995) Expresssion of *Streptomyces* peucetius genes for doxorubicin resistance and aklavinone 11-hydroxylase in *Streptomyces galilaeus* ATCC 31133 and production of a hybrid aclacinomycin. *Antimicrob Agents Chemother* 39: 1616–1620

235 Niemi J, Mantsala P (1995) Nucleotide sequences and expression of genes from *Streptomyces purpurescens* that cause the production of new anthracyclines in *Streptomyces galilaeus*. *J Bacteriol* 177: 2942–2945

236 Ylihonko K, Hakala J, Kunari T, Mantsala P (1996) Production of hybrid anthracycline antibiotics by heterologous expression of *Streptomyces nogalater* nogalamycin biosynthesis genes. *Microbiology* 142: 1965–1972

237 Kim HS, Hong YS, Kim YH, Yoo OJ, Lee JJ (1996) New anthracycline metabolites produced by the aklavinone 11-hydroxylase gene in *Streptomyces galilaeus* ATCC 31133. *J Antibiot* 49: 355–360

238 Stachelhaus T, Schneider A, Marahiel MA (1995) Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* 269: 69–72

239 Staunton J (1998) Combinatorial biosynthesis of erythromycin and complex polyketides. *Curr Opin Chem Biol* 2: 339–345

240 Hopwood DA (2003) The *Streptomyces* genome – be prepared! *Nature Biotechnol* 21: 505–506

241 Ikeda H, Ishikawa J, Hanamoto A, Shinose K, Kikuchi H, Shiba T, Sakaki Y, Hattori M, Omura S (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nature Biotechnol* 21: 526–531

242 Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, Ahlert J, Thorson JS, Shen B, Farnet CM (2003) A genomics-guided approach for discovering and expressing cryptic metabolic pathways. *Nature Biotechnol* 21: 187–190

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Nutritional and engineering aspects of microbial process development

By Prakash S. Masurekar

Department of Plant Biology and Pathology, School of Environmental and Biological Sciences,
Rutgers University, 59 Dudley Road, New Brunswick, NJ 08901-8525, USA
<prakash_masurekar@verizon.net>

Abstract

Today we use many drugs produced by microorganisms. However, when these drugs were discovered it was found that the yields were low and a substantial effort had to be put in to develop commercially viable processes. A key part of this endeavor was the studies of the nutritional and the engineering parameters. In this chapter, the basic principles of optimizing the nutritional and engineering aspect of the production process are described with appropriate examples. It was found that two critical components of nutritional medium, carbon and nitrogen source regulated the synthesis of the compounds of interest. Rapidly utilizable carbon source such as glucose supported the growth but led to catabolite repression and alternative carbon sources or methods of addition had to be devised. Inorganic nitrogen sources led to undesirable changes in pH of the medium. Organic nitrogen sources could influence the yields positively or negatively and had to be chosen carefully. Essential nutrients like phosphates often inhibited the synthesis and its concentration had to be maintained below the inhibitory levels. On many occasions, trace nutrients like metal ions and vitamins were found to be critical for good production. Temperature and pH were important environmental variables and their optimum values had to be determined. The media were designed and optimized initially with 'one variable at a time' approach and later with experimental design based on statistics. The latter approach is preferred because it is economical, considers interactions between medium components and allows rapid optimization of the process. The engineering aspects like aeration, agitation, medium sterilization, heat transfer, process monitoring and control, become critical as the process is scaled-up to the production size. Aeration and agitation are probably the most important variables. In many processes dissolved oxygen concentration had to be maintained above a critical value to obtain the best yields. The rheological properties of fermentation broth significantly affect the aeration and mixing efficiency. The removal of heat from the large fermentors can be difficult under certain conditions. However, new designs of impellers, availability of sensors to monitor important physiological and process variables and advent of computers have facilitated successful scale-up of fermentation processes.

1 Introduction

For millennia humans have utilized microorganisms for many purposes, ranging from preparation of food and beverages to medicine. This review will focus on the preparation of medicines with the use of microbial processes. However, it will not cover biotransformation processes as some of the underlying principles applicable to the development of these vary from those that apply to the development of fermentation processes.

Drugs of microbial origin are often prepared from compounds referred to as 'secondary metabolites'. The reason for calling these compounds secondary metabolites is that they have no obvious function. One of the hypotheses to explain their production suggests that they provide com-

petitive advantage in nature. The difficulty in proving this suggestion has been that it has not been possible to detect these compounds in natural habitats such as soil, because in almost all cases the desired metabolites are produced in very low quantities. A substantial effort, therefore, is needed to develop an economically viable production process with the original isolate from nature. Nutritional and engineering aspects of this endeavor are discussed in this chapter.

1.1 Nutritional aspects

The nutritional requirements for the growth of the producer microorganism are different from those needed for the production of the desired secondary metabolite. As a result the medium used for substantial growth is often referred to as, 'seed medium' and that used during the production phase is called 'production medium'. Both seed and production media contain carbon source(s), nitrogen source(s) and sources of trace nutrients such as magnesium, sodium, potassium, sulfur, copper, zinc etc., vitamins and amino acids. Often they also contain some buffers. The carbon and nitrogen sources used may be the same in both the seed and the production media; however their concentrations are often different. Generally, different carbon and nitrogen sources are used in these two types of media. In the early days of fermentation process developments our understanding of the microbial metabolism, especially that of secondary metabolism was limited and as a result the media contained complex materials of plant and animal origin. Such media are designated as 'complex' media. The advantages of complex media are; i) they are usually inexpensive, ii) they provide 'unknown critical' ingredient, iii) they support good growth and iv) their development is rapid. Their disadvantages are; i) due to the presence of the crude complex ingredients the performance is variable, ii) they pose difficulties in scale-up and iii) these may interfere in downstream processing. However, the increase in our knowledge of these processes and the availability of fermentors capable of monitoring and controlling various physiological parameters have resulted in desire for media with 'defined' ingredients [1]. The advantages of defined media are; i) the performance is consistent, ii) ease of scale-up, iii) amenable to control and iv) fewer problems in downstream processing. The disadvantages of these media are; i)

cost and ii) reproducing the high yields obtained with the complex media. The following discussion will try to accommodate both types of media.

1.1.1 Carbon source

The carbon sources provide energy and building blocks for the microorganisms. Commonly used carbon sources are monosaccharides like glucose and fructose, disaccharides such as lactose (whey contains 65% lactose) and molasses, polysaccharides like dextrins, starch and cellulose, polyols like glycerol; oils and alcohols [1]. Glucose is an attractive carbon source. It is available in crude form as cerelose and corn syrup, both of which are very inexpensive. It is utilized by all microorganisms. However, glucose can cause 'catabolite repression' [2] of critical enzymes in the biosynthetic pathway of the desired product. For example two critical enzymes in the biosynthesis of cephamycin C in *Nocardia lactamdurans*, (d(α-aminoadipyl)-cysteinyl-valine) synthetase and deacetoxycephalosporin C synthase, are repressed by glucose [3]. It has also been proposed that the higher rate of growth obtained with glucose is the cause of such reduction in yield [4]. Regardless of mechanism, where glucose inhibits the product formation, other carbon sources mentioned above can be used in place of glucose. Table 1 lists examples of replacement of glucose with other carbon sources [4]. Another approach that is often used is to use lower concentration of glucose in combination with another slowly utilized carbon source. In such situations glucose is used to support rapid growth and the second carbon source is used for the synthesis of the product, where slower growth rate is desirable. This approach was used to develop medium for the production of lovastatin [26].

Secondly, glucose is often metabolized rapidly through the glycolytic pathway, which results in a drop in pH. It has been noted that pH is one of the critical environmental variables. Addition of an inorganic buffering agent such as either sodium or potassium phosphate or organic buffers like 2-(N-morpholino)-ethanesulphonic acid (MES) or 3-(N-morpholino)-propanesulphonic acid (MOPS) can prevent the reduction in pH. In cases where buffers cannot be used, substitution of glucose with disaccharides, polysaccharides or polyols, can reduce the pH drop. Their slower utilization reduces both the drop in pH and the potential for catabolite repression. While studying the effect of carbon source on lovastatin production

Table 1.
Carbon catabolite regulation of antibiotic biosynthesis

Antibiotic	Interfering carbon source	Non-interfering carbon source	Refs
Penicillin	Glucose	Lactose	[5]
Actinomycin	Glucose	Galactose	[6]
Streptomycin	Glucose	Mannan	[7]
	Glucose	Slowly fed glucose	[8]
Siomycin	Glucose	Maltose	[9]
Indolmycin	Glucose	Fructose	[10]
Bacitracin	Glucose	Citrate	[11, 12]
Cephalosporin C	Glucose	Sucrose	[13]
Chloramphenicol	Glucose	Glycerol	[14]
Violacein	Glucose	Maltose	[15]
Prodigiosin	Glucose	Galactose	[16]
Mitomycin	Glucose	Low glucose	[17]
Neomycin	Glucose	Maltose	[18]
Kanamycin	Glucose	Galactose	[19]
Enniatin	Glucose	Lactose	[20]
Puromycin	Glucose	Glycerol	[21]
Novobiocin	Citrate	Glucose	[22]
Candidin	Glucose	Slow feed of glucose	[23]
Candihexin	Glucose	Slow feed of glucose	[23]
Butirosin	Glucose	Glycerol	[24]
Cephamycin	Glycerol	Aspargine, Starch	[25]

Table reproduced from [4] with the permission from the copyright holder, American Society for Microbiology

it was found that with glucose as carbon source the pH reduced to 4.5. This pH drop was avoided with the use of phosphate, MES or MOPS buffers and the yield was doubled. Similar results were obtained when glucose was replaced with glycerol [26].

In addition to the two approaches mentioned above, these problems can be solved at the fermentor scale, by automatic pH control and slow feeding of glucose [5].

1.1.2 Nitrogen source

Nitrogen sources are required for the synthesis of cell components such as proteins, nucleic acids, cell wall as well as metabolites both primary and secondary. Examples of commonly used nitrogen sources are ammonium salts, nitrates, urea, amino acids, protein hydrolysates and proteins. These crude protein sources can be of plant or animal origins. It should be noted that due to the concern about bovine spongiform encephalitis the use of animal proteins may cause difficulties with the regulatory agencies. The inorganic nitrogen sources are often used up rapidly leading to reduction in pH and adverse effect on the growth and/or production. Therefore, usually complex organic nitrogen sources are preferred. Amino acids are also used as nitrogen sources especially in defined media. Of all the amino acids utilized the most common is monosodium glutamate. Amino acids in addition to acting as precursors of the metabolites of interest can also play a regulatory role. An excellent example of the multiple roles amino acids play is described by Petersen et al. [27]. They showed that addition of L-proline to fermentation resulted in almost doubling of yield of pneumocandin B_0 and reduction in the yields of pneumocandins C_0 and D_0. On the other hand, addition of L-threonine and L-serine reduced the titer of pneumocandin B_0. The reduction in the production of penicillin G upon the addition of L-lysine is another example of amino acid regulation of synthesis of secondary metabolites [28].

1.1.3 Effect of phosphate and trace minerals

Phosphate is critical for the biosynthesis of nucleic acids and for energy metabolism. It also can be used as a buffering agent. However, it has been found to have a regulatory role in the biosynthesis of many antibiotics [29]. Phosphate concentrations above 10 mM were found to inhibit the production of antibiotics. The products regulated by phosphate come from different chemical groups such as peptide antibiotics, polyene, macrolides, tetracyclines and complex antibiotics [4]. The mechanism of phosphate regulation is not fully understood [30]. It has been suggested that the ATP concentration is the intracellular mediator of the phosphate regulation. Another suggestion is that adenylate energy charge, i.e., a ratio of ATP+

0.5 ADP concentration to that of total adenylate phosphates is involved in the regulation of synthesis. A recent paper describes the role of PhoR-PhoP in phosphate regulation [30]. Regardless of the mechanism of phosphate regulation, it is important to be aware of the regulatory role of phosphate while designing a production medium.

Trace minerals, which are important for microbial nutrition include magnesium, copper, iron, cobalt, molybdenum, manganese, calcium, boron, zinc, sulfate and chloride. Some of these can also affect the production of antibiotics. Cobalt salts had positive effect on the production of thienamycin by *Streptomyces cattleya* (Inamine E 1982. Abst Interscience Conf Antimicrob Agents Chemother). In the case of synthesis of pneumocandins, addition of Zn^{++}, Co^{++}. Cu^{++} and Ni^{++} generally reduced the production of pneumocandin B_0 and affected the product spectrum [27]. So care should be exercised in adding the trace minerals to medium.

1.1.4 Effect of pH and temperature

An example of the effect of pH was described above [26]. Since pH can affect the metabolism it is not surprising that it is an important variable. Similarly, temperature also has strong influence on the metabolism and therefore it too affects the production of the secondary metabolites. Since both these variables are often optimized in screening or single variable type of approach they are considered further in the subsequent section.

1.2 Medium development

In this section the approaches used to design an optimized production medium are described. The first step usually is to study the medium in which the compound of interest was discovered to look for the clues in designing a better medium. Next a number of carbon and nitrogen sources, temperatures and pH values are evaluated. This is normally done with a 'one variable at a time' approach; although sometimes statistical methods such as the Plackett-Burman technique are also used. Once the composition of the medium is identified it is further improved and optimized with the use of experiment design based on statistical methods. This approach

allows testing of multiple variables at the same time. There are many advantages of this approach, the most important being it saves time and labor. Secondly, the interaction between various medium components cannot be detected by the 'one variable at a time' approach. For the development of the optimum medium it is essential that these interactions are taken into account.

1.2.1 Screening and 'one variable at a time' approach

The initial selection of the carbon and nitrogen source tested is based either on the clues from the medium used to discover the compound of interest or by choosing one or two members of each of the various types mentioned earlier. These are tested one at a time. For example, to design a medium for the production of rifamycin by *Amycolatopsis mediterranei* a number of carbon and nitrogen sources were tested. These included monosaccharides, disaccharides, polysaccharides and oils, ammonium salts, soybean meal and peanut meal and other protein hydrolysates. It was found that glucose and lactose supported the highest yields. However, increase in their concentration above 25 g/L resulted in decrease in titer. Similarly, ammonium sulfate and ammonium carbonate were good inorganic nitrogen sources and soybean meal and peanut meal were good organic ones. From the pH and temperature values tried, a pH of 7.2 and temperature of 28°C were the optimum [31]. During the studies on the medium and temperature for the production of a novel antifungal agent ascosteroide by *Ascotricha amphitrica*, carbon sources tested were mono- and disaccharides, polyols, whey (crude lactose), choline and ethanolamine. The carbon sources that supported good yields were glucose, sorbose, glucosamine and inositol. Glucose was selected for further work due to its cost advantage. The best titers were seen at a glucose concentration of 30 g/L. Optimum temperatures were 16°C and 25°C for production in either shake flask or in fermentor, respectively [32]. Another example of this type of approach is the development of medium for the production of anthracycline by *Streptomyces peucetius*. Here too, a number of mono- and disaccharides along with soluble starch were examined in a defined medium. From these, fructose, maltose, mannose and soluble starch supported high production of anthracyclin. Glucose was a poor carbon source. From the nitrogen sources

tried aspargine was the best. Aspartic acid also gave good yields. However, inorganic nitrogen sources like ammonium salts or nitrites were poor. Surprisingly, nitrate was a good nitrogen source [33]. Production of anthracyclin was inhibited by inorganic phosphate. In the case if actinorhodin synthesis by *Streptomyces coelicolor*, starch up to concentration of 50 g/L, was found to support good yields. Glutamate was used as nitrogen source. Phosphate concentrations above 2.5 mM inhibited the synthesis of actinorhodin. A complex interaction between nitrogen source and phosphate was noted [34]. Although it has been observed that glucose usually is not an ideal carbon source for the production of the secondary metabolites, there are examples where it can be used. It was reported that production of novobiocin by *Streptomyces niveus* was not inhibited by glucose. Interestingly, citrate inhibited the synthesis. In the medium containing both glucose and citrate, the latter was used first and production began after it was exhausted. Glucose supported the production [35]. Aharonowitz and Demain did a careful analysis of the effect of nitrogen source on the production of cephalosporin by *Streptomyces clavuligerus* [36]. Ammonium salts, nitrates, amino acids and urea were studied for their effect on the growth and the antibiotic formation both in the presence and in the absence of glycerol. It was found that amino acids and urea supported good growth and antibiotic synthesis. The best yields were observed with L-aspargine as nitrogen source. Good growth was obtained with ammonium salts but no antibiotic production was seen. Antibiotic synthesis was inhibited by 75% when L-aspargine was supplemented with ammonium chloride. The examples given above demonstrate the use of screening and 'one variable at a time' approach for carbon and nitrogen sources and phosphate. Some more examples are listed in references [37–42].

As mentioned earlier, pH and temperature are also important variables to be considered. James et al. [43] reported on the process development for the production of granaticin by *Streptomyces thermoviolaceus*. This culture is a thermophile and grows in the temperature range 25°–57°C. The temperature and pH screens showed that 50°C and 7.0 were the optimum temperature and pH, respectively. Of all the carbon sources tested, it was found that those which fed into the TCA cycle were better than those which fed into glycolysis. Another example of optimization of these two variables was described by Palmer for the production of a phytotoxin, coronatine [44]. This compound is produced by *Pseudomonas syringe*. The

surprising finding was that the optimum temperature was 18°C. Increase in the temperature reduced the production and reached a minimum at 30°C. The synthesis of this compound was reduced at temperatures below 18°C. There was no significant effect on the growth or the yields in the pH range of 6.5–7.8. There was reduction of 65% in growth at pH 5.8 but the specific rate of synthesis was not affected. Osmolarity of the medium, varied by addition of NaCl had a strong negative effect on the yields but not on the growth. Interestingly, glucose was the best carbon source from the mono- and disaccharides and polyols tested. Nitrates were better in supporting production than ammonium salts and amino acids. Phosphate did not inhibit the production. Addition of ferrous salts increased the titer.

1.2.2 Experiment design based on statistics

The following discussion of the empirical techniques used is intended for understanding the basic statistical approach involved and not as a 'how to guide'. There are a number of software packages available for this purpose such as JMP^{R} by SAS Institute. These are relatively easy to use. If such software is not available and experimenter is not familiar with statistics, it is recommended that a statistician be consulted before using this methodology.

The most popular technique used for the purpose of studying the joint effects of several factors is called factorial design. The term, 'joint factor effects', includes both the main effects and the interactions. Since in a production medium there are interactions between various components and the exact mechanisms of these interactions are not understood this method is preferred.

In application of this approach the relationship between the variables tested and the final result can be written as $y = f(\varepsilon_1,\varepsilon_2,\ldots\ldots\varepsilon_k) + \varepsilon$. Usually function f is a first-order or second-order polynomial. A graphical representation of the solution of this polynomial in three dimension results in a surface lying above the plane that represents the factors under consideration. This graphical representation has been referred to as 'response surface' and this empirical model is called response surface model. If this surface is viewed from the top and all points, which have the same yield are connected to produce contour lines of constant response, the resulting figure is called

'contour plot'. The response surface and contour plots are extremely useful for the analysis of the results of a factorial experiment. Therefore, the terms response surface and factorial design are used to describe this statistical approach [45]. In the following discussion of these methods, general statistical terms are used since these techniques are applicable to a broad range of design and optimization problems. In context of medium design and optimization, medium components or environmental variables are denoted as 'factors' and their concentrations or values are represented as 'levels'. Factorial design of k factors each at two levels requires 2 x 2 x.....2 = 2^k observations, it is referred to as 2^k factorial design. To distinguish from the class of fractional factorial designs it is also called 2^k full factorial design [46]. A model matrix for 2^3 design is shown in Table 2. Two key characteristics of 2^k full factorial design are balance and orthogonality; balance indicates that each factor level appears in the same number of runs and orthogonality defines that for two factors to be considered orthogonal they must appear in all their level combinations in the same number of runs. A design is considered orthogonal if all pairs of factors are orthogonal. The polynomial generated from this design will show the main effects and the interaction effects of the variables. The high and the low values of a variable are denoted by + and – signs. For example, high value of variable A will be represented as A+ and the low value as A–. Let $\bar{z}$ be the average of observations at high and low values of A. Then the main effects (ME) of A are given by,

$$ME(A) = \bar{z}\,(A+) - \bar{z}\,(A-)$$

and the interaction effects of variables A and B [INT(A,B)] are given by,

$$INT(A,B) = \frac{1}{2}\{\bar{z}\,(B+|A+) - \bar{z}\,(B-|A+)\} - \frac{1}{2}\{\bar{z}\,(B+|A-) - \bar{z}\,(B-|A-)\}$$

This equation can be generalized for k factors as follows:

$$INT(A_1,A_2 \ldots\ldots A_k) = \frac{1}{2}\ INT(A_1,A_2 \ldots\ldots A_{k-1}|A_k+) - (A_1,A_2 \ldots\ldots A_{k-1}A_k-)$$

(The above relationships are from [46])

Fundamental principles for factorial effects are hierarchical, effect sparsity and effect heredity. Hierarchical ordering principle can be described as follows:

Table 2.
2^3 Factorial design including all interactions

	Factorial Design							
Treatment combination	**I**	**A**	**B**	**C**	**AB**	**AC**	**BC**	**ABC**
A	+	+	–	–	–	–	+	+
B	+	–	+	–	–	+	–	+
C	+	–	–	+	+	–	–	+
abc	+	+	+	+	+	+	+	+
Ab	+	+	+	–	+	–	–	–
Ac	+	+	–	+	–	+	–	–
Bc	+	–	+	+	–	–	+	–
(1)	+	–	–	–	+	+	+	–

Reproduced from [45] with the permission from the copyright holder, John Wiley & Sons, Inc.

1. Lower order effects are more likely to be important than higher order effects
2. Effects of the same order are equally likely to be important

Effect sparsity principle means the number of relatively important effects in a factorial experiment is small.

The last principle states that in order for an interaction to be significant, at least one of its parent factors should be significant. These principles are important to the analysis of the results of a factorial experiment.

In the days prior to the availability of the computational power of today, the effects of the factors had to be determined by hand. A number of methods were available to do that. An example described below demonstrates the use of regression and model matrix to compute the effects. For factor A in the Table 2 the eight rows can be designated as x_{11}, x_{21}, ……,x_{81} or xi1 where i = 1……8. Linear model is given by:

$$\bar{z}\,i = \sum_{j=1}^{7} \beta ixij + \varepsilon i$$

In the ith run i = 1,….,8 and for j = 1,…,7 the least square estimate of βj is

$$\beta j = 1/1-(-1)\{\bar{z}(xij=+1)-\bar{z}(xij=-1)\},$$

…which is one half of the factorial effect and therefore, each factorial effect can be obtained by doubling the corresponding regression estimate βj [46]. As mentioned earlier, this type of hand calculations are not needed today.

Two level experiments assume a linear relationship between the levels of the factor and their effect. Very often it is not true and in order to determine the curvilinear relation it is necessary to test three levels. This type of models will be considered later in the chapter.

Since the number of experiments needed to grow exponentially with addition of new factor to be studied, it is obvious that full factorial design becomes less feasible. For example with six factors the full factorial design will require 64 experiments. Only six of the 63 degrees of freedom are used to estimate the main effects, 15 degrees of freedom are used to estimate two-factor interactions. The remaining 42 degrees of freedom are associated with the three-factor and higher interactions. It has been generally observed that in medium design and optimization the three-factor and higher interactions are not important, therefore it is not necessary to use a full factorial design [45]. The information relevant for the purpose can be obtained by carrying out only a fraction of the full factorial design. These fractional factorial designs, because of their efficiency, are the most popular design type used in the industry. These are designated for k factors and two levels as 2^{k-p} designs. For example 2^{k-1} will be one-half factorial, 2^{k-2} will be one-quarter factorial and so on. There are two concepts, which are important to know about fractional factorial designs. These are called 'aliasing' and 'design resolution'. One consequence of choosing to run only a fraction of the full factorial design is that it is not possible to separate main effects from some of the interactions. Two or more effects that have this property are called 'aliases'. There are techniques available for resolving the aliases. The definition 'design resolution' is based on the extent of aliasing. These definitions are listed below.

Resolution III Design: No main effects aliased with each other but main effects are aliased with two-factor interactions and two-factor interactions may be aliased with each other.

Resolution IV Design: No main effects aliased with each other or with two-factor interactions, but two-factor interactions are aliased with each other.

Table 3.
2^{3-1} Factorial design (One-half factorial)

	2^{3-1}_{iii}, I = ABC			2^{3-1}_{iii}, I = –ABC		
Run	**A**	**B**	**C = +AB**	**A**	**B**	**C = –AB**
1	–	–	+	–	–	–
2	+	–	–	+	–	+
3	–	+	–	–	+	+
4	+	+	+	+	+	–

Reproduced from [45] with the permission from the copyright holder, John Wiley & Sons, Inc.

Resolution V Design: No main effects or two-factor interactions are aliased with any other main effect or two-factor interactions but two-factor interactions are aliased with three-factor interactions.

In general, the designs which have the highest level of resolution are preferred. The higher the resolution, the less restrictive the assumptions that are required regarding which interactions are negligible to obtain a unique interpretation of the data.

An example of one-half factorial is shown in Table 3. These one-half factorial designs were prepared by choosing the treatment combinations, which have pluses in ABC column. Thus ABC is called the 'generator' of this particular fraction. It is sometimes referred to as a 'word'. Furthermore, the identity column I is always plus therefore, I = ABC and which is called the defining relation of the design. One-half fraction, which is defined by I = ABC is called the 'principal fraction' and remainder is referred to as 'alternate' or 'complimentary' fraction [45].

As mentioned above both response surface and contour plots are important tools for analysis of the data obtained in a factorial experiment. The properties needed to fit a second-order response surface are:

i. at least three levels of each design variable
ii. at least $1 + 2k + k(k-1)/2$ distinct data points

All designs commonly used have these properties [45, 46].

1.2.3 Commonly used designs

1.2.3.1 Central Composite Design (CCD)

This design is by far the most popular second-order design used for process optimization [47]. In this type of design two level factorial or fractional factorial (resolution V) is combined with 2k axial or star points and n_c central points. The resolution V design contribute to estimation of linear terms and two-factor interactions; axial points contribute to estimation of quadratic terms and central points provide the estimation of pure error and contribute towards the estimation of quadratic terms. Selection of α, the axial distance, and n_c allow flexibility in this design. The axial distance is selected based on the nature of the design region, whether it is spherical or cuboidal. In spherical region α is $\sqrt{k}$ and in cuboid region it is 1. In spherical region it is most effective to use three to five center runs (n_c). On the other hand, in cuboidal region one to two center runs are adequate. In cuboidal design the axial points are on the center of the faces and hence it is often referred to as face-centered cube (FCC) [46]. For actual construction and analysis of this type of design see [45–47].

1.2.3.2 Simplex design

Regular simplex is a first order design, wherein minimum number of experimental points (in any number of dimensions) are called for; for example, in k dimensions it requires (k+1) points [48]. In two dimensions the regular simplex is an equilateral triangle; in three dimensions it is a tetrahedron and so on. The design is both efficient and 'rotatable'. It has been shown that this design without replication is optimal for the estimation of slope in the presence of error. For the present purpose of optimization however, the additional (and unique) attraction is that it is possible by adding just one further point to complete a new simplex on the face of the original simplex design. Thus to the question 'where to move', the answer is 'into an adjacent simplex'. Only one further observation is needed to complete this new simplex and enable the experimenter to ask that question once more. It may be shown that the direction of the steepest ascent estimated from observations at the vertices of regular simplex will proceed from the center of simplex through that face of the simplex, which is opposite to (does not contain) the point corresponding to the lowest observation. More generally, for a simplex S_0 with vertices v_1, v_2,……,v_{k+1} and center

c_0, it is possible to construct a new simplex S_i with center c_i, which has k vertices v_1, v_2,.....v_{i-1}, , v_{i+1},.....v_{k+1} in common with S_0 and is completed by adding v_i. To find any new coordinate v_c we take twice the average of the corresponding coordinates for the common vertices, v_1, v_2,.....v_{i-1}, , v_{i+1},.....v_{k+1} and subtract the coordinate v_i. In vector notation it is:

$$v_i^* = \frac{2}{k}(v_1, v_2, \ldots\ldots v_{i-1}, , v_{i+1}, \ldots\ldots v_{k+1}) - v_i$$

It is important to note that the path taken by a simplex design will turn around the optimum and no more improvement in the results can be seen. To define the optimum more precisely, a second series of simplex experiments can be undertaken where the size of the change in the level of a factor is reduced compared to the first series [1].

1.2.3.3 Plackett–Burman design

Often it is not possible to do the number of experimental runs as required even by fractional factorial design and at such times different types of designs are required. The factorial designs in which it is possible to estimate any two factorial effects independently of each other or they are fully aliased, is referred to as 'regular design'. The designs, which do not possess this property, are called 'non-regular designs'. As mentioned above these designs are used for the reasons of run size economy or flexibility. The non-regular designs not only require fewer runs but can also accommodate various combinations of factors with different number of levels [45]. These designs are built from orthogonal arrays. Plackett and Burman described a large collection of this type of array [49]. These are denoted as $OA(N,2^{n-1})$, where N = number of runs and k = N–1 variables. Other characteristics of N are that it is a multiple of 4 but not a power of 2. Therefore, the smallest design is N=12. Some of the other designs commonly used consist of N=20, 24 and 28 [46]. In the early phase of medium development a number of ingredients are tested or screened for their usefulness. Plackett–Burman designs are very suitable for this purpose. In a 12 run experiment, the first seven columns of the array represent seven medium components and the last four are dummy variables. The dummy variables are used to measure the error in the experiment, which is then used to determine the significance of the effect of any of the variables tested. Table 4 shows the first row of Plackett–Burman designs for 12, 20 and 28 runs. In this table + and

Table 4.
First row of the Plackett–Burman arrays

N	Vectors
12	+ + – + + + – – – + –
20	+ + – – + + + + – + – + – – – – + + –
24	+ + + + + – + – + + – – + + – – + – + – – – –

– signs are used to denote high and low levels of the variables, respectively. From this first row the rest of the rows are generated by moving the previous row to the right. For example to generate the second row the last entry in the first row is placed in the first column of the second row and all other entries in the first column are moved one position to the right. The third row is generated from the second row in the same manner [49]. This design has been very useful in rapidly screening a large number of variables.

1.2.3.4 Box–Behnken design

Medium development and optimization for the production of fermentation products involves studying quantitative variables and this design was developed specifically for such purpose [50]. It is a class of three level incomplete factorial designs useful for estimating the coefficients in a second degree polynomial. As described by the authors their aim was, where possible, to generate second order rotatable designs. To do so three levels taken by the variables $x_1, x_2, \ldots, x_k$ are coded –1, 0 and +1 and it is assumed that the second degree graduating polynomial is fitted by the method of least squares

$$\hat{y} = b_0 + \sum_{i=1}^{k} \beta i x i + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta i j x i x j$$

A second order rotatable design is such that the variance of $\hat{y}$ is constant for all points equidistant from the center of the design. A balanced incomplete block design for four variables and three levels in six blocks is given in Table 5. This balanced incomplete block design is then combined with 2 levels and 2 factors (2^2). Two asterisks in the rows of Table 5 are replaced with the two columns of 2^2 factorial design. The blank spaces are replaced with zeros. The design is completed with the addition of center points (0,

Table 5.
A balanced incomplete block design for four variables in six blocks

	x_1	x_2	x_3	x_4
1	*	*		
2			*	*
3	*			*
4		*	*	
5		*		*
6	*		*	

Reproduced from [50] with permission from copyright holder, American Statistical Association.

0, 0, 0). For the design described here three rows of center points will be desirable. In general, three to five center runs are recommended for this type of design.

One of the characteristics of this design is that it is spherical. It can be imagined that for three variables, all experimental points fall at the center of edges of a cube. The Box–Behnken design is comparable to the central composite design in the number of runs required. For example for three variables the former design will contain $12 + n_c$ runs and the latter will have $14 + n_c$ runs. Both the Box–Behnken and the central composite designs are popular with experimenters. One of the reasons for their popularity is the run size, which is large enough to provide a comfortable margin for lack of fit but not so large as to involve wasted degrees of freedom or unnecessary experimental expense [45].

The designs described here are not the only ones used but are those which are used more often. There are new approaches that are being tried such as 'fuzzy logic' and 'artificial neural network'. Those interested in these are recommended to refer to references [51–55].

1.3 Some examples

A complete factorial design was used to study the effect of glucose as the carbon source and monosodium glutamate as the nitrogen source on the production of the red pigments by *Monoascus ruber* in synthetic medium.

Optimum concentrations of glucose and monosodium glutamate were found to be 26 g/L and 5 g/L, respectively [56].

Roseiro et al. used Plackett–Burman design to screen the medium constituents in the studies on medium development for xanthan production [57]. Carbon source (glucose) and nitrogen source (ammonium sulfate) were found to be critical variables. Response surface design showed that the C/N ratio of 23 was optimum. In addition, calcium carbonate and potassium phosphate (monobasic) were found to be important factors, probably because of their role in pH modulation.

An interesting example of use of factorial design is the development of a medium based on sugarcane juice for increasing the production of astaxanthine [58]. In the first round of study, the variables were the concentrations of sugarcane juice, peptone and ammonium phosphate (monobasic). The design used was central composite. The yield was increased by 23% but the specific productivity was decreased by 16%. In the second round, oxygen transfer rate, agitation and pH were studied. This resulted in specific production of 418 μg astaxanthin/g of yeast and the overall increase in production to 2 mg/L.

The central composite design was used to develop an industrial medium for the improved production of carotenoids by a mutant strain of *Phaffia rhodozyma* [59]. The concentrations of corn starch hydrolysate and corn steep liquor were the variables tested. The optimized medium was supplemented with urea and potassium phosphate (monobasic). The use of supplemented optimized medium in fed-batch fermentation gave a yield of 52 mg/L.

Dasu and Pande [60] used central composite design for the optimization of microbiological parameters for enhanced production of griseofulvin. The culture used was *Penicillium griseofulvin* MTCC 1898 and *Penicillium griseofulvum* MTCC 2004. The variables tested were slant age, seed age and inoculum level. The optimum values of three variables studied for MTCC 1898, were 9 days, 4 days and 10%, respectively, and those for MTCC 2004 were 8 days, 3 days and 9%, respectively. The titer with both strains under optimum conditions was 0.4 g/L.

A similar strategy was also used to develop a medium for the production of riboflavin by *Ermothecium ashbyii*. For the design of the initial screening experiments, Plackett–Burman methodology was used. The results showed that magnesium sulfate and sodium chloride had a significant effect on

the yield. The production medium was optimized with the use of central composite design. The optimum medium contained molasses 50 g/L (glucose equivalent); sesame seed cake 50 g/L; yeast extract 2 g/L; potassium phosphate (monobasic) 2 g/L; magnesium sulfate 0.117 g/L and sodium chloride 1.13 g/L [61].

The results reported by Wang et al. [62] are noteworthy because they optimized a solid production medium with Box–Behnken design and response surface methodology. All the examples given so far involved the use of liquid media. They optimized concentrations of carbon source and nitrogen source for modulating the ratios between monocolin K, γ-aminobutyric acid and citrinin produced by *Monoascus purpureus*.

Another type of fractional factorial, Box–Wilson design, and response surface methodology was used to optimize the concentration of glucose and yeast extract for the production of compactin by *Penicillium* sp. [63].

These examples illustrate the use of statistical experimental design methods to develop media for the production of secondary metabolites. More examples of the use of these methodologies are given in reference [52].

2 Engineering aspects

So far the discussion of process development has been focused on the work done in small scale systems such as shake flasks, culture tubes or deep-well plates. At this scale the engineering aspects are not relevant. However, once the process is optimized, it has to be scaled up to larger sizes. Secondly, at the larger scale the process is run in fermentors. As a result, a number of new engineering variables have to be considered. These include aeration, agitation (mixing), rheological nature of the fermentation broth, heat transfer and sterilization of the medium. It is important to recognize that each of these variables affects the others.

2.1 Aeration and agitation

For an aerobic process, i.e., the producing microorganism requires oxygen for growth, it is necessary to supply adequate oxygen. The relationship

between the oxygen transfer rate (OTR) and the physical parameters is given by the following equation

$$OTR = K_L.a\ (C^* - C)$$

where K_L = liquid film mass transfer coefficient, a = interfacial area between gas phase and liquid phase, C* = saturation concentration of oxygen in liquid, C = Concentration of oxygen in liquid.

This equation indicates that in order to improve the transfer rate, it is necessary to increase the values of K_L, a and C*.

In fermentors, the air is sparged at the bottom of the tank and the air bubbles rise to the top. As they travel up, they are sheared by multiple impellers. This results in the generation of smaller bubbles. Thus it can be seen that aeration and agitation are critical for assuring adequate supply of oxygen to the growing culture. In addition to shearing air bubbles, agitation is necessary to keep the fermentation broth well mixed in order to avoid any local limitations of oxygen or any medium ingredient. Both these requirements have to be considered in impeller design. It has been recognized that biggest resistance to oxygen transfer from air to liquid medium is that of the liquid film. It was found that $K_L.a$ is affected by many operating variables. Earliest report on this was by Cooper and his co-workers [64]. They showed that for vaned disk impeller, volumetric mass transfer coefficient (K_v) was correlated with the unit power input (P_v) and superficial gas velocity (V_s) as shown by the following two equations.

$$K_v = k\ P_v^{0.95} \text{ and } K_v = k'V_s^{0.67}$$

where k and k' are constants.

In addition to the relationship shown above, subsequent studies have shown it to be proportional to impeller tip speed [65] and to apparent viscosity in non-Newtonian fermentation broths [66]. As indicated by these equations increase in the power input and in superficial velocity were necessary to obtain higher $K_L.a$.

However, there is a limit to the increase in power input and in superficial velocity. So the attention was turned to the redesign of the impellers and the fermentors. The most common impeller type was Rushton

turbine. Based on the flow pattern it is classified as radial flow turbine. It was capable of high shear but did not have the high pumping capacity needed. Multiple Rushton turbine impellers were the design of choice for fermentors from the early 1950s to the early 1980s [67]. Currently agitators with other design types are available. The biggest design change has been the development of axial flow hydrofoil impellers. These offer better performance over the Rushton impellers for large-scale viscous, mycelial fermentations. The specific advantages of this type of impellers are: improved oxygen transfer per unit power; lower maximum shear rates and improved bulk mixing resulting in elimination of compartmentalization of flow obtained by multiple radial flow impellers, better control of pH, elimination of local nutrient limitation and hot spots [67].

Consideration of geometry is also important to obtain better oxygen transfer. The relationship between the diameter of the tank and that of impeller (D/T) can modulates the efficiency of aeration. The optimum value of D/T is affected by the aeration rate and the mixer power input. Placement of impeller in the tank has relatively low impact on K_L.a [68].

With the advent of sterilizable dissolved oxygen (DO) probes, it became possible to measure that parameter during fermentation. It was noted in many processes that there was a critical concentration of (DO), below which the production was adversely affected. Therefore, it is imperative to maintain it above the critical level for the duration of the run. As mentioned earlier, due to the limitations of power input, the power per unit volume decreases with the scale and as a result the mixing time increases with the fermentor volume. The longer mixing times in large fermentors cause DO concentration gradients. Therefore, location of the DO probe has to be determined carefully.

As mentioned earlier there is a limit to increase in power input and in superficial velocity of air. In such cases, either the back-pressure can be increased or the air stream can be enriched with oxygen. Both methods increase C* and thereby enhance the driving force for OTR. It is a common practice to operate the fermentors with back-pressure.

Just as the concentration of dissolved oxygen is critical to obtain high yields so is that of carbon dioxide. It is highly soluble in aqueous media and has an effect on its pH. It can also inhibit the production. Therefore, the above-mentioned variables also have to be optimized/controlled to efficiently strip carbon dioxide from the medium.

There is a vast amount of literature on the topic of aeration and agitation and it is not possible to review it in depth in this chapter. However, the information provided in this section should help in appreciating their importance.

2.2 Rheology of fermentation broths

Microorganisms belonging to different classes such as unicellular bacteria, actinomycetes, yeast and molds have been used for the production of metabolites. However, the majority of the useful compounds are produced either by actinomycetes or by fungi. Both these types of organisms grow in mycelial form, i.e., the ratio of length to diameter is very high. The fermentation broths containing mycelium are, even at low solids concentration, of non-Newtonian characteristics [69]. The rheological properties of the fermentation broths significantly affect efficiency of aeration, agitation and mixing. In viscous broth significant gradients of DO, nutrients, pH and temperature are observed. These often adversely affect the yield of the desired product. Furthermore, the broth viscosity is affected by the culture morphology, which in turn changes with the age of the fermentation. Therefore, it is very important to study this parameter during the development of fermentation process. Excellent analysis of the effects of broth rheology was done by Roels et al. [69] and by Blanch and Bhavaraju [66].

Newton's law defining the viscosity μ, is shown below

$$\tau = -\mu \bar{\bar{\Delta}}$$

where μ is a constant, τ is shear stress and $\bar{\bar{\Delta}}$ is stretching tensor.

However, there are cases where the viscosity is not a constant, i.e., where the fluids do not obey Newton's law, the relationship is:

$$\tau = -\eta \bar{\bar{\Delta}}$$

where η is apparent viscometric viscosity. These type of fluids are called non-Newtonian in character. This class of fluids is divided into, 1) time-independent fluids, 2) time-dependent fluids and 3) viscoelastic fluids. The first sub-class is further divided into dilatant and pseudoplastic fluids.

Many fermentation broths exhibit pseudoplastic behavior [66]. A number of models have been proposed to correlate shear stress with shear rate. Some of these are the power law model, Bingham model and Casson model. A number of studies on the rheological behavior of mycelial broths have been reported. Deindorfer and Gaden showed that *P. chrysogenum* broth behavior could be described by the Bingham model [70]. Deindorfer and West noted that the data from the same fungus could be better fitted to power law equation [71]. Whereas, Roels et al. found that their results from *P. chrysogenum* fermentation could be better modeled by Casson equation [69]. The usefulness of these models is that they allow prediction of rheological behavior of the broths in the early stage of process development, which is important from an engineering point of view and assist in selection of strains more amenable to scale-up [66].

It was mentioned earlier that the morphology of microorganism affects the rheology of the broth. In general, shorter mycelial length results in lower viscosity. Fungi can grow as pellets. However, as in the case of mycelial growth, that in the form of pellets also has its drawbacks. These are – slower growth, lower reproducibility and potential for DO limitation. It is also difficult to get good samples because the pellets settle quickly [72]. In spite of these shortcomings, often the change over from mycelium to pellets is preferred since it results in lower viscosity and thus decreases the power required for good mixing and oxygen transfer. Pellet characteristics can be influenced by inoculum concentration and its morphological characteristics, trace elements in the medium, polymer additives to the medium and agitation (shear sensitivity of the culture). Those interested in more information are directed to Junker et al., which lists a number of references on this topic [72].

Sometimes it is not possible to reduce the broth viscosity by controlling the morphology of the culture. In such cases the broth is diluted with water. This simple and practical technique reduces the viscosity significantly. It was found that a 10% dilution resulted in 50% drop in viscosity [73].

2.3 Sterilization of medium

In the laboratory the medium is sterilized by autoclaving, usually at 121°C, which is a batch process. The time required to come up to the

sterilization temperature and cool down to operating temperature is relatively short and does not significantly affect yield. But at the higher scale these are significantly longer and lead to the destruction of temperature sensitive medium constituents and increased levels of undesirable reactions between the medium components, such as those between the reducing sugars and the free amino groups. To avoid these problems continuous sterilization also called high temperature short time sterilization (HTST), is preferred over the batch type. As the name suggests the medium is rapidly heated to the desired high temperature (140°C or above) by pumping the medium through a heat exchanger or by steam injection. If the latter approach is used it is important to correct for the dilution of the medium caused by the condensation of steam. The heated medium is held at the sterilization temperature for the desired time and cooled down quickly in a heat exchanger. To make the process economically efficient the heat is recovered in the cooling stage by preheating the incoming medium.

The rationale for the use of HTST process can be explained based on the Arrhenius equation given below.

$$\ln K = C - E/RT$$

where T is temperature, C is constant, E is activation energy and R is Boltzman's constant. The activation energy for killing of spores is in the range of 50–100 kcals/mole and that for the destruction temperature sensitive medium components is in the range of 10–30 kcals/mole. It can be seen from the application of Arrhenius equation that at high temperatures there will be less loss of sensitive medium components [74].

2.4 Heat transfer

Removal of heat generated by metabolic activity is a serious problem in large vessels. The present design of jacketed fermentors limit the surface area available for the transfer of heat from the medium to the cooling water. A number of other designs such as hollow baffles have been used to increase the surface area. In the geographic areas where cooling water temperatures in summer can be high, severe heat transfer problems are

encountered. Therefore, it is important to measure the heat load at the pilot plant level so adjustments can be made to the medium to reduce the peak rate of metabolism at the production scale.

2.5 Process control

Initially, the process variables, which were monitored, were temperature and pH. Later dissolved oxygen probe became available. Analysis of exhaust gases became possible with the development of mass spectrometers for this purpose. With the advent of computers it was possible to complete the loop and control the process based on the real time values of the variables to maintain them in the optimum range. It became possible to use different strategies such as cascade control to keep dissolved oxygen above the critical value. This can be done by controlling two variables, namely the rate of aeration and that of agitation. This is very useful when the producing culture is shear sensitive. The rate of agitation can be increased up to a point above which shear will damage the culture. If needed, further improvement in OTR can be obtained by increasing the airflow or increasing the back-pressure or both.

The pH is controlled by the addition of acid or base. It is important to pay attention to the location of the pH probe and the addition tube(s). The probe should be placed in the well-mixed region of the tank. If either the acid or the base is added on the top surface of the medium, where bulk mixing is the worst in the tank, they will be dispersed slowly. This will result in local pH gradients and may have deleterious effect on the product formation. The addition point should be near one of the impellers so the acid or base can be dispersed quickly [74]. Sometimes ammonia gas mixed with incoming air is used to control pH. It is important to note that it can change the C/N ratio of the medium and may have detrimental effect on the yields.

Exhaust gas analysis allows real time determination of the metabolism. The concentrations of carbon dioxide, oxygen and nitrogen in the exhaust gas can be accurately measured. This information can be directly fed to a computer to determine a number of critical parameters, such as carbon dioxide evolution rate, oxygen uptake rate, respiratory quotient, dry cell weight, and specific growth rate [75]. These parameters can be used to

initiate/control of nutrient feeding, adjustment of batch temperature and control aeration and agitation.

More information on process control and use of computers is available in the reference numbers [76–78].

2.6 Process models

The aim of modeling fermentation processes is to assist in process design and control. Microbial metabolism is complex and it is not possible to understand it fully at any given time in the fermentation process. These models try to mathematically correlate critical aspects of the metabolism with the desired final outcome. The approach used is to express fermentation kinetics in different system levels. They are, 1) molecular or enzyme level, 2) macromolecular or cellular component level, 3) cellular level and 4) population level [79]. Simplest of the models described the relationships between growth of the cell and limiting substrate concentration, as postulated by Monod.

$$\frac{dx}{dt} = \mu_{\max} X \frac{s}{Ks+S}$$

where, X is the cell concentration, S is the limiting substrate concentration, t is time, μ_{max} is the maximum growth rate and Ks is substrate coefficient.

The production of the compound of interest may be growth associated or non-growth associated or both. For growth associated product the relationship is shown below

$$\frac{dP}{dt} = \alpha \frac{dX}{dt}$$

where, P is concentration of the product. Analogous to specific growth rate designated by μ, specific production rate is designated by q and is defined as:

$$q = \frac{dP}{Xdt}$$

In the cases where the product formation is function of cell concentration the relationship can be expressed as:

$$\frac{dP}{dt} = \beta X$$

where β is specific rate constant and has units of gm P/gm X-time

Luedeking and Piret modeled lactic acid production by combining the two equations [80].

$$\frac{dP}{dt} = \alpha \frac{dX}{dt} + \beta X$$

The models got more sophisticated and complex as more computational power became available. For example Heijnen and co-workers described a model for penicillin fermentation where an elemental balancing method was used. It involved solving seven equations, which show balancing of carbon, hydrogen, nitrogen, oxygen, sulfur, phosphorus and enthalpy [81]. A kinetic model to describe the cephalosporin biosynthesis in *Streptomyces clavuligerus* was developed by Malmberg and Hu [82]. In this model rates of change of the intermediates catalyzed by five enzymes – ACV synthetase, cyclase, epimerase, expandase and hydrolase – were denoted by five differential equations. The equations correlated the rates of change in the concentrations of intermediates with rates of reaction catalyzed by the five enzymes. The rates of reactions, V_i (i = 1, 2,…5) were characterized by Michaelis–Menten equation. These rate equations are non-linear and solved numerically. The computer simulation showed a good agreement with the experimental results. The usefulness of the model to guide metabolic engineering to increase the yield was shown.

The new approaches to optimize the process such as artificial neural network and fuzzy logic are also being used to model and optimize fermentation processes [51–55, 83, 84].

2.7 Antifoaming agents

Foaming is not a serious problem in laboratory scale processes but in fermentors it is. To control it a number of antifoaming agents are used. These include polyglycols, e.g., P-2000 and silicone antifoams like Dow Corning 1520. In general, addition of antifoams reduced OTR. The reasons for this are considered to be reduction of K_L [85] by suppression of mobility of

Table 6.
List of antifoam agents

Antifoam agent	Producer/Provider
P-2000	Dow Chemicals, Midland, MI
SAG 471	Union Carbide, (Dow Chemicals) Midland, MI
Ucon LB625	Dow Chemicals, Midland MI
P-4000	Dow Chemicals, Midland MI
Mazu DF204	PPG Industries Pittsburgh, PA
P-1200	Dow Chemicals, Midland MI
Silicone oil	Dow Corning Midland, MI
Hodag FD62	Lambert Technologies Gurnee, IL
P-200	Dow Chemicals, Midland MI
Hodag FD82	Lambert Technologies Gurnee, IL
Chemax DF30	PCC Chemax Piedmont, S. C.

the surface, and increase in bubble coalescence leading to reduction of interfacial area a [86]. Lowering of mass transfer was also reported by Yagi and Yoshida [87]. However, the observed effect on the OTR was smaller, which could have been due to the lower concentration of antifoam used. While these data were obtained in bubble reactors, reduction in OTR was also noted in stirred tanks [88, 89]. Some commonly used antifoam agents are listed in Table 6.

Fermentation processes can be operated in batch, fed-batch or continuous manner, fed-batch mode being the most common. More information on this type of processes can be found in a review by Whitaker [90].

2.8 Some examples

The examples listed below illustrate some of the engineering aspects described in the previous sections.

Vardar-Sukan reviewed the relationship between the operating variables and oxygen mass transfer [91, 92]. He noted that the mass transfer rate is affected by airflow rate, solubility of oxygen in medium, temperature, dissolved solids, surface active agents, oxygen partial pressure and rheological

properties of broth. The results were similar to those described above. The design variables identified were sparger design, impeller geometry, power input, vessel geometry and operating liquid volume.

Cho and his co-workers noted the effect aeration rate at a constant agitation rate of 150 rpm on the production of red pigment by *Paecilomyces sinclairii* in sucrose medium [93]. They found that the highest yields were obtained at the aeration rate of 3.5 volumes per volume per minute (vvm). However, best growth was seen at 1.5 vvm.

Makagiansar et al. studied the effect of agitation on the mean hyphal length of *Penicillium chrysogenum*. They found that the mean hyphal length did not correlate well with the tip speed but with the circulation frequency and energy dissipation around the impeller. A good correlation between the mean hyphal length and penicillin production with power dissipation function was obtained [94].

Interestingly, it was found that in the case of production of tocamycin (nucleoside antibiotic) by *Streptomyces chrestomyceticus*, the tip velocity was a critical variable [95]. This culture was shear sensitive and it was essential to reduce the tip velocity to 275 cm/sec after 20 h of cultivation irrespective of the size of the fermentor.

The effect of tip velocity is complex. Heydarian and co-workers noted that the increase in the tip velocity reduced the growth but had little effect on erythromycin production by *Saccharopolyspora erythrea* in the range of 350–1,000 rpm. However, it decreased by 10% at 1,250 rpm [96].

Dissolved oxygen concentration was mentioned above as an important variable. It was found that the yield of ascosteroid, a novel antifungal agent produced by *Ascotricha amphitrica*, was sensitive to the concentration of dissolved oxygen. The optimum level was 40% and any increase above that obtained by either increasing the agitation or the rate of aeration reduced the production [97].

Vardar and Lilly studied the effect of dissolved oxygen on product formation in penicillin fermentation. They noted that concentration below 30% sharply lowered the specific rate of production and no production was seen below 10% [98].

Cephamycin C synthesis by *Streptomyces clavuligerus* did not begin until the growth stopped and the dissolved oxygen concentration increased to 50%, which had dropped to 0 during the growth. If the concentration was maintained throughout the run at either 50% or 100% by feeding pure

oxygen, the production was initiated 10 h earlier and resulted in higher specific rate of production as well as higher yields. Under these conditions the rate of agitation did not have any effect on the specific rate of the volumetric productivity. These results indicated that mass transfer coefficient was not critical for production [99].

Zhou et al. noted that the production of Cephalosporin C by *Cephalosporium acremonium* was reduced at dissolved oxygen concentration below 20% and that of penicillin N was increased [100].

Similarly, in the case of bacitracin production by *Bacillus licheniformis* enrichment with oxygen increased the yield [101].

In general, increased dissolved oxygen concentration enhances production. However there are some exceptions. It was reported by Vandamme et al. that gramicidin S production by *Bacillus brevis* was reduced at higher dissolved oxygen concentrations [102].

An interesting effect of oxygen transfer rate on the ratio of red pigment and citrinin produced by *Monascus ruber* was seen. When oxygen transfer rate was increased either by increasing aeration or agitation, the production of citrinin was enhanced and that of the red pigment was reduced. Therefore, low transfer rate was optimum for higher ratio between the red pigment and citrinin [103].

As mentioned above, sterilization is an important variable at higher scale. This was demonstrated by Jain and Buckland during their studies on the scale-up of antibiotic efrotomycin. It was found that it was critical to sterilize glucose with the rest of the medium to obtain the optimum metabolism required for the synthesis of efrotomycin. If glucose was sterilized separately, the metabolism was shifted to rapid use of glucose in the medium and cessation of growth. Soybean oil, which is the main carbon source for the production, was not used and very low yield of the antibiotic was obtained [104].

The examples given above hopefully introduce the readers to the effects of various engineering aspects of microbial processes mentioned. No attempt has been made to be comprehensive as it is beyond the scope of this chapter.

3 Summary

The goal of this chapter was to introduce the readers to nutritional and engineering aspects of microbial process development. It was shown that carbon and nitrogen sources regulate the production of secondary metabolites and that significantly affects the choice of these two major constituents of the medium. Phosphate also plays a regulatory role. The yields and the spectrum of the compounds produced are sensitive to the trace nutrients. The awareness of these influences is critical for the success of medium design and optimization. The media can be developed using either one variable at a time method or with experimental design based on the statistics. The latter approach is preferred because it is economical, faster and takes into account interactions between medium components. Temperature and pH are important environmental variables and need to be studied for the optimization of the process. The scale-up of the process requires consideration of engineering variables such as, aeration, agitation, medium sterilization and heat transfer to mention a few. With the development of sensors to monitor the process parameters and that of the computers, it has become possible to control the fermentations to obtain the best results. The discussion in this chapter was organized to introduce the reader to the basic principles first followed by illustration with appropriate examples. It is hoped that the readers find this approach informative and useful.

References

1 Greasham RL, Herber WK (1997) Design and optimization of growth media. In: PM Rhodes, PF Stanbury (eds): *Applied microbial physiology*, IRL Press, New York, 53–74

2 Magasanik B (1961) Catabolite repression. *Cold Spring Harbor Symp Quant Biol* 26: 249–256

3 Cortes J, Liras P, Castro JM, Martin JF (1986) Glucose regulation of cephamycin biosynthesis in *Streptomyces lactamdurans* is exerted on the formation of a-aminoadipyl-cysteinyl-valine and desacetoxycephalosporin synthase. *J Gen Microbiol* 132: 1805–1814

4 Martin JF, Demain AL (1980) Control of antibiotic biosynthesis. *Microbiol Rev* 44: 232–251

5 Soltero FV, Johnson MJ (1954) Continuous addition of glucose for evaluation of penicillin producing cultures. *Appl Microbiol* 2: 41–44

6 Gallo M, Katz E (1972) Regulation of secondary metabolite biosynthesis. Catabolite

repression of phnoxazinone synthase and actinomycin formation by glucose. *J Bacteriol* 109: 659–667
7 Demain AL, Inamine E (1970) Biochemistry and regulation of streptomycin and mannosidostreptomycinase (α-D-mannosidase) formation. *Bacteriol Rev* 34: 1–19
8 Inamine E, Lago BD, Demain AL (1969) Regulation of mannosidase, an enzyme of streptomycin biosynthesis. In: D Perlman (ed): *Fermentation advances*. Academic Press New York, 199–221
9 Kimura A (1967) Biochemical studies on *Streptomyces sioyanensis* II. Mechanism of the inhibitory effect of glucose on siomycin formation. *Agric Biol Chem* 31: 845–852
10 Hurley LH, Bialek D (1974) Regulation of antibiotic production: catabolite inhibition and the dualistic effect of glucose on indolmycin production. *J Antibiot*ics 27: 49–56
11 Haavic HI (1974) Studies on the formation of bacitracin in *Bacillus licheniformis*: effect of glucose. *J Gen Microbiol* 81: 383–390
12 Haavic HI (1974) Studies on the formation of bacitracin in *Bacillus licheniformis*: role of catabolite repression and organic acids. *J Gen Microbiol* 84: 321–326
13 Demain AL (1963) Synthesis of cephalosporin C by resting cells of *Cephalosporium* sp. *Clin Med* 70: 2045–2051
14 Smith CG, Hinman JW (1963) Chloramphenicol. *Prog Ind Microbiol* 4: 137–163
15 DeMoss RD (1967) Violacein. In: D Gottlieb, PD Shaw (eds): *Antibiotics, vol 2*. Springer-Verlag, Berlin, 77–81
16 Ramsey HH, Hussain SM, Williams RP (1973) Inhibition by glucose of prodigiosin biosynthesis in *Serratia marcescens*. *Abstr Annu Meet Am Soc Microbiol* 239: 180
17 Kirsch EJ (1967) Mitomycins. In: D Gottlieb, PD Shaw (eds): *Antibiotics, vol 2: Biosynthesis. Springer-Verlag, Berlin, 66–76*
18 Majumdar MK, Majumdar SK (1971) Synthesis of neomycin by washed mycelium of *Streptomyces fradiae* and some physiological considerations. *Folia Microbiol* 16: 285–292
19 Basak K, Majumdar SK (1973) Utilization of carbon and nitrogen sources by *Streptomyces kanamyceticus* for kanamycin production. *Antimicrob Agents Chemother* 4: 6–10
20 Audhya TK, Russell DW (1975) Enniatin production by *Fusarium sambucinum*. Primary, secondary and unitary metabolism. *J Gen Microbiol* 86: 327–332
21 Sankaran L, Pogell BM (1975) Biosynthesis of puromycin in *Streptomyces alboniger*: regulation and properties of O-methylpuromycin O-methyltransferase. *Antimicrob Agents Chemother* 8: 721–732
22 Kominek LA (1975) Biosynthesis of novobiocin by *Streptomyces niveus*. *Antimicrob Agents Chemother* 1: 123–134
23 Martin JF, McDaniel LE (1974) The submerged culture production of the polyene antifungal antibiotics candidin and candihexin. *Dev Ind Microbiol* 15: 324–337
24 Howells JD, Anderson LE, Coffey GL, Senos GD, Underhill MA, Vogler DL, Erlich J (1972) Butirosin, a new aminoglycosidic antibiotic complex: bacterial origin and some microbiological studies. *Antimicrob Agents Chemother* 2: 79–83
25 Aharonowitz Y, Demain AL (1978) Carbon catabolite regulation of cephalosporin production in *Streptomyces clavuligerus*. *Antimicrob Agents Chemother* 14: 159–164
26 Buckland B, Gbewonyo K, Hallada T, Kaplan L, Masurekar P (1989) Nutritional and engineering aspects of microbial process development. In: AL Demain, GA Somkuti, JC Hunter-Cevera, HW Rossmoore (eds): *Novel microbial products for medicine and agriculture*. Elsevier, Amsterdam, 161–169
27 Petersen LA, Hughes DL, Hughes R, DiMichele L, Salmon P, Connors N (2001) Effects of

amino acids and trace element supplementation on pneumocandin production by *Glarea lozoyensis*: Impact on titer, analogue levels and the identification of new analogues of pneumocandin B_0. *J Ind Microbiol Biotechnol* 26: 216–221

28 Masurekar PS, Demain AL (1972) Lysine control of penicillin biosynthesis. *Can J Microbiol* 18: 1045–1048

29 Martin JF (1977) Control of antibiotic synthesis by phosphate. *Adv in Biochem Eng* 6: 105–127

30 Martin JF (2004) Phosphate control of biosynthesis of antibiotics and other secondary metabolites is mediated by PhoR-PhoP system: an unfinished story. *J Bacteriol* 186: 5197–5201

31 Venkateswarlu G, Krishna PM, Rao LV (1999) Production of rifamycin using *Amycolatopsis mediterranei* (MTCC14). *Bioprocess Eng* 20: 27–30

32 Lowe SE, Gustavson DR, Meyers DA, Veitch SA, Forenza S, Lam KS (1997) The effect of carbon source, temperature and aeration on the production of ascosteroide, a novel antifungal agent by *Ascotricha amphitrica*. *J Antibiot*ics 50: 412–417

33 Dekeva ML, Titus JA, Strohl WR (1985) Nutrient effects on anthracyclin production by *Streptomyces peucetius* in defined medium. *Can J Microbiol* 31: 287–294

34 Doull JL, Vining LC (1990) Nutritional control of actinorhodin production by *Streptomyces coelicolor* A3 (2): suppressive effects of nitrogen and phosphate. *Appl Microbiol* Biotechnol 32: 449–454

35 Kominek LA (1972) Biosynthesis of novobiocin by *Streptomyces niveus*. *Antimicrob Agents Chemother* 1: 123–134

36 Aharonowitz Y, Demain AL (1979) Nitrogen nutrition and regulation of cephalosporin production in *Streptomyces clavuligerus*. *Can J Microbiol* 25: 61–67

37 Chen M-H, Johns MR (1994) Effect of carbon source on ethanol and pigment production by *Monoascus purpureus*. *Enzyme and Microbial Technology* 16: 584–590

38 Wong HC, Koehler PE (1981) Production and isolation of an antibiotic from *Monoascus purpureus* and its relationship with pigment production. *J Food Sci* 46: 589–592

39 Lin TF, Demain AL (1991) Effect of nutrition of *Monoascus* sp on formation of red pigment. *Appl Microbiol* Biotechnol 36: 70–75

40 Guzman S, Ramos I, Moreno E, Ruiz B, Rodriguez-Sanoja R, Escalante L, Langley E, Sanchez S (2005) Sugar uptake and sensitivity to carbon catabolite regulation in *Streptomyces peucituss* var *caesius*. *Appl Microbiol Biotechnol* 60: 200–206

41 Slater H, Crow M, Everson L, Salmond GPC (2003) Phosphate availability regulates biosynthesis of two antibiotics, prodigiosin and carbapenem in *Serratia via* both quorum sensing-dependent and independent pathways. *Mole Microbiol* 47: 303–320

42 Basak K, Majumdar SK (1973) Utilization of carbon and nitrogen sources by *Streptomyces kanamyceticus* for kanamycin production. *Antimicrob Agents Chemother* 4: 6–10

43 James PDA, Edwards C (1988) The effect of culture conditions on growth and secondary metabolism in *Streptomyces thermoviolaceus*. *FEMS Microbiol Lett* 52: 1–6

44 Palmer DA, Bender CL (1993) Effects of environmental and nutritional factors on the production of polyketide phytotoxin coronatine by *Pseudomonas syringe* pv glycinea. *Appl Environ Microbiol* 59: 1619–1626

45 Myers RH, Montgomery DC (2002) *Response surface methodology*. John Wiley and Sons, New York

46 Wu CFJ, Hamada M (2000) *Experiments planning, analysis and parameter design optimization*. John Wiley and Sons, New York

47 Box GEP, Wilson KB (1951) On the experimental attainment of optimum conditions. *J Roy Stat Soc Series B* 13: 1–45
48 Spendley W, Hext GR, Himsworth FR (1962) Sequential application of Simplex design in optimization and evolutionary operation. *Technometrics* 4: 441–461
49 Plackett RL, Burman JP (1946) The design of optimum multifactorial experiments. *Biometrika* 33: 305–325
50 Box GEP, Behnken DW (1960) Some new three level designs for the study of quantitative variables. *Technometrics* 2: 455–475
51 Kennedy MJ, Prapulla SG, Thakur MS (1992) Designing fermentation media: a comparison of neural networks to factorial design. *Biotechnol Techniques* 6: 293–298
52 Kennedy MJ, Krouse D (1999) Strategies for improving fermentation medium performance: a review. *J Ind Microbiol Biotechnol* 23: 456–475
53 Nyttle NG, Chidambaram M (1993) Fuzzy logic control of a fed-batch fermentor. *Bioprocess Eng* 9: 115–118
54 Postlethwaite BE (1989) A fuzzy state estimator for fed-batch fermentation. *Chem Eng Res Des* 67: 267–272
55 Lee CC (1990) Fuzzy logic control systems: Fuzzy logic controller, Part II. *IEEE Trans Syst Man Cybernet* 20: 419–435
56 Pastrana L, Blanc PJ, Santerre AL, Loret MO, Goma G (1995) Production of red pigments by *Monoascus ruber* in synthetic medium with strictly controlled nitrogen source. *Proc Biochem* 30: 333–341
57 Roseiro JC, Esgalhado ME, Collaco A, Emery AN (1992) Medium development for xanthan production. *Proc Biochem* 27: 167–175
58 Florencio JA, Soccol CR, Furlanetto LF, Bonfim TMB, Kriegar N, Baron M, Fontana JD (1998) A factorial approach for sugarcane juice-based low cost culture medium: increasing the astaxanthin production by red yeast. *Phaffia rhodozyma*. *Bioproc Eng* 19: 161–164
59 Kesava SS, An G-H, Kim C-H, Rhee S-K, Choi E-S (1998) An industrial medium for improved production of carotenoids from a mutant strain of *Phaffia rhodozyma*. *Bioproc Eng* 19: 165–170
60 Dasu VV, Pande T (2000) Optimization of microbiological parameters for enhanced griseofulvin production using response surface methodology. *Bioproc Eng* 22: 45–49
61 Pujari V, Chandra TS (2000) Statistical optimization of medium components for improved synthesis of riboflavin by *Ermothecium ashbyii*. *Bioproc Eng* 23: 303–307
62 Wang JJ, Lee CL, Pan TM (2003) Improvement of monocolin K, g-aminobutyric acid and citrinin production ratio as a function of environmental conditions of *Monoascus purpureus* NTU 601. *J Ind Microbiol Biotechnol* 30: 669–676
63 Konya A, Jekkel A, Suto J, Salat J (1998) Optimization of compactin fermentation. *J Ind Microbiol Biotechnol* 20: 150–152
64 Cooper CM, Fernstrom A, Miller SA (1944) Performance of agitated gas-liquid contactors. *Ind Eng Chem* 36: 504–509
65 Richards JW (1961) Studies in aeration and agitation. *Progress Ind Microbiol* 3: 143–172
66 Blanch HW, Bhavaraju SM (1976) Non-Newtonian fermentation broths: rheology and mass transfer. *Biotech Bioeng* 18: 745–790
67 Buckland BC, Lilly MD (1993) Fermentation: an overview. In: G Stephanopoulos (ed): *Biotechnology*. VCH, Weinheim, 7–22

68 Oldshue JY (1966) Fermentation mixing scale-up techniques. *Biotechnol Bioeng* 8: 3–24

69 Roels JA, Van Den Berg J, Voncken RM (1974) The rheology of mycelial broths. *Biotechnol Bioeng* 14: 181–208

70 Deindorfer FH, Gaden EL (1955) Effect of liquid physical properties on oxygen transfer in penicillin fermentation. *Appl Microbiol* 3: 253–257

71 Deindorfer FH, West JM (1960) Rheological properties of fermentation broths. *Adv Appl Microbiol* 2: 265

72 Junker BH, Hesse M, Burgess B, Masurekar P, Connors N, Seely A (2004) Early phase process scale-up challenges for fungal and filamentous bacterial cultures. *Appl Biochem Biotechnol* 119: 242–277

73 Taguchi H (1971) The nature of the fermentation fluids. In: TK Ghose, A Fiechter (eds): *Advances in biochemical engineering*. Springer-Verlag, Berlin, 1–25

74 Buckland BC (1984) The translation of scale in fermentation process: impact of computer process control. *Bio/Technol* 2: 875–884

75 Buckland BC, Brix T, Fastert H, Gbewonyo K, Hunt G, Jain D (1985) Fermentation exhaust gas analysis using mass spectrometry. *Bio/Technol* 3: 982–988

76 Kuenzi MT (1978) Process design and control in antibiotic fermentations. In: *FEMS Symposium. Antibiotics and other secondary metabolites. Biosynthesis and production* 5: 39–56

77 Zabriskie DW (1979) Real-time application of computers in fermentation processes. *Ann NY Acad Sci* 326: 223–239

78 Constantinidis A (1979) Application of rigorous optimization methods to the control and operation of fermentation process. *Ann NY Acad Sci* 326: 193–221

79 Humphrey AE (1979) Fermentation process modeling: an overview. *Ann NY Acad Sci* 326: 17–33

80 Luedeking R, Piret EL (1959) A kinetic study of the lactic acid fermentation: batch process at controlled pH. *J Biochem Microbiol Tech Eng* 1: 393

81 Heijnen JJ, Roels JA, Stouthamer AH (1979) Application of balancing methods in modeling the penicillin fermentation. *Biotechnol Bioeng* 21: 2175–2201

82 Malmberg LH, Hu WS (1991) Kinetic analysis of cephalosporin biosynthesis in *Streptomyces clavuligerus*. *Biotechnol Bioeng* 38: 941–947

83 DiMassimo C, Willis MS, Montague GA, Tham MJ, Morris AJ (1991) Bioprocess model building using artificial neural network. *Bioproc Eng* 7: 77–82

84 Giudici R, Pamboukian CRD, Facciotti MCR (2004) Morphologically structured model for antitumoral retamycin production during batch and fed-batch cultivation of *Streptomyces olindersis*. *Biotechnol Bioeng* 86: 414–424

85 Kawase Y, Moo-Young M (1990) The effect of antifoam agents on mass transfer in bioreactors. *Bioproc Eng* 5: 169–173

86 Lee YH, Lerk S (1983) Aeration. *Ann Rep Ferm Proc* 6: 101–147

87 Yagi H, Yoshida F (1974) Oxygen absorption in fermentors – effects of surfactants, antifoams agents and sterilized cells. *J Ferment Technol* 52: 905–916

88 Benedek A, Heideger WJ (1971) Effects of additives on mass transfer in turbine aeration. *Biotechnol Bioeng* 13: 663–684

89 Kimura M, Nakao K (1986) Measurement of oxygen transfer coefficient using glucose oxidase reaction. *Chem Eng (Japan)* 31: 786–794

90 Whitaker A (1980) Fed-batch culture. *Proc Biochem* 15: 10–15, 32

91 Vardar-Sukon F (1985) Dynamics of oxygen mass transfer in bioreactors. Part I. Operating variables affecting mass transfer. *Proc Biochem* 20: 181–184

92 Vardar-Sukon F (1986) Dynamics of oxygen mass transfer in bioreactor. Part II. Design variables. *Proc Biochem* 21: 40–44
93 Cho YJ, Hwang HJ, Kim SW, Song CH, Yun JW (2002) Effect of carbon source and aeration rate on broth rheaology, and fungal morphology during the red pigment production by *Paecilomyces sinclairii* in batch bioreactor. *J Biotechnol* 95: 13–23
94 Makagiansar HV, Shamalou PA, Thomas CR, Lilly MD (1993) The influence of mechanical forces on the morphology and penicillin production of *Penicillium chrysogenum*. *Bioproc Eng* 9: 83–90
95 Flickinger MC, Greenstein M, Bremmon C, Conlin J (1990) Strain selection, medium development and scale-up of tocamycin production by *Streptomyces chrestomyceticus*. *Bioproc Eng* 5: 143–153
96 Heydarian SM, Mirjalili N, Ison AP (1999) Effects of shear on the morphology and erythromycin production in *Saccharopolyspora erythrea* fermentations. *Bioprocess Eng* 21: 31–39
97 Lowe SE, Gustavson DR, Veitch JA, Forenza S, Lam KS (1997) The effect of carbon source, temperature and aeration on the production of ascosteroid, a novel antifungal agent, by *Ascotricha amphitricha*. *J Antibiot* 50: 412–417
98 Vardar F, Lilly MD (1982) Effect of cycling dissolved oxygen concentration on product formation in penicillin fermentation. *Eur J Appl Microbiol Biotechnol* 14: 203–211
99 Rollins MJ, Jensen SE, Westlake DWS (1988) Effect of aeration on antibiotic production by *Streptomyces clavuligerus*. *J Ind Microbiol* 3: 357–364
100 Zhou W, Holzhauer-Rieger K, Dors M, Schugeri K (1992) Influence of dissolved oxygen concentration on the biosynthesis of cephalosporin C. *Enzyme Microbiol Technol* 14: 848–854
101 Flickinger MC, Perlman D (1979) Application of oxygen-enriched aeration in the production of bacitracin by *Bacillus licheniformis*. *Antimicrob Agents Chemother* 15: 282–293
102 Vandamme EJ, Ley man D, Visscher P, DeBuyser D, Vanstinkiste G (1981) Effect of aeration and pH on the gramicidin S production by *Bacillus brevis*. *J Chem Tech Biotechnol* 31: 247–257
103 Hajjaj H, Blanc PJ, Groussac E, Goma G, Uribelarrea JL, Loubiere P (1999) Improvement of red pigment/citrinin production ratio as a function of environmental conditions by *Monascus rubber*. *Biotechnol Bioeng* 64: 497–501
104 Jain D, Buckland BC (1988) Scale-up of efrotomycin fermentation using a computer controlled pilot plant. *Bioproc Eng* 3: 31–36

Progress in Drug Research, Vol. 65
(Frank Petersen and René Amstutz, Eds.)

Natural products from plant cell cultures

By Elizabeth McCoy and Sarah E. O'Connor

Department of Chemistry,
Massachusetts Institute of Technology,
Cambridge, MA 02139, USA
<soc@mit.edu>

Abstract

Plants produce complex small molecules – natural products – that exhibit anticancer, antimalarial and antimicrobial activity. These molecules play a key role in human medicine. However, plants typically produce these compounds in low quantities, and harvesting plant natural products is frequently expensive, time-consuming and environmentally damaging. Plant cell culture provides a renewable, easily scalable source of plant material. In this chapter we discuss the successes and pitfalls associated with natural product production in plant cell cultures.

1 Introduction

Plants are an important source of natural products; these complex, higher eukaryotes make many compounds that are not found in bacteria and fungi. Notably, plants are immobile and interact with their environment largely *via* the release of complex small molecules. This evolutionary pressure has resulted in the production of an extraordinarily diverse array of natural products by plants. Approximately 100,000 compounds from plants have been reported, with thousands more being discovered each year [1]. Terpene-derived compounds form the largest class of plant-derived natural products, with compounds in the alkaloid natural product class following at a close second [2]. Approximately 10% of all drugs considered to be 'basic and essential' by the WHO are isolated from plants [3]. A small selection of plant-derived drugs is shown in Figure 1.

Many natural products are isolated from wild plants, and harvesting of these molecules often has a significant, negative impact on the environment. Although some medicinal plants can be farmed – for example, *Hypericum perforatum*, which produces the herbal antidepressant St. Johns Wort – many medicinally important plants cannot be grown under controlled farming conditions [4]. A particularly compelling example is illustrated in the discovery of the anticancer agent Taxol that is isolated from the yew tree, as discussed in Section 3.

Plant cell culture has the potential to greatly expedite the production of natural products in a controlled environment. Controlled fermentation conditions avoid the variations in growing conditions that invariably occur with plants grown in the wild, such as differences in soil composition and attack by pathogens. Importantly, cell culture provides a renewable source of natural products, as plant cell culture can be produced and harvested

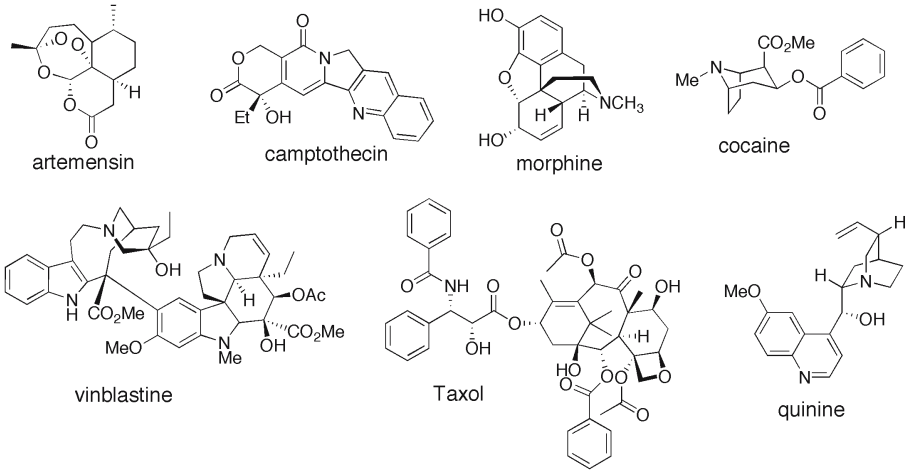


Figure 1
Representative pharmaceuticals found in plants

year round without damage to the environment. Plant cell culture was initiated in the 1930s [5, 6], and serious efforts to use plant cell culture as to produce natural products began in the 1970s [7]. These early studies unequivocally established that plant cell cultures produce many natural products; in modern research efforts, the practicality of natural product production in plant cells continues to be assessed [8].

Unfortunately, many pharmaceutically important plant natural products are not produced in cell culture. Biosynthesis of certain compounds requires the cellular differentiation that is not found in a homogenous cell culture. For example, the bis-indole alkaloid anticancer agent vinblastine from periwinkle has not been cultured in dedifferentiated cell cultures, because several of the biosynthetic genes are only expressed in epidermal cells of the whole plant [9]. Additionally, expression of natural products in cell cultures is often not stable over years of subculture. Certain plant extracts that contain complex mixtures of many products, such as *Ginkgo biloba*, are used medicinally. However, the relative ratios of the various compounds are often different in cell culture than they are in the plant [10]. From an engineering perspective, culturing of plant cells on a large scale presents technical challenges; the slow growing cells are susceptible to microbial contamination and are sensitive to mechanical stresses as

Table 1
Plant cell cultures used in industry (Adapted from reference [16])

Cell culture	Product	Function	Manufacturer
Duboisia	Scopolamine	Anticholinergicum	Sumitomo Chemical Industries (Japan)
Podophyllum	Podophyllotoxin	Antitumor	Nippon Oil (Japan)
Coptis japonica *Thalictrum minus*	Berberines	Antitumor	Mitsui Petrochemical Industries (Japan)
Taxus	Paclitaxel	Antitumor	ESCAgenetics (USA), Phyton Catalytic (USA/ Germany) Nippon Oil (Japan)
Coleus blumei	Rosmarinic acid	Anti-inflammatory	Nattermann (Germany)
Panax ginseng	Ginseng biomass	Dietary supplement	Nitto Denko (Japan)
Echinacea purpurea *Echinacea augustifolia*	Echinaceae	Immunostimulant	Diversa (Germany)
Lithospermum erythrorhizon	Shikonin	Pigment Antibiotic	Mitsui Petrochemical Industries (Japan)
Geramineae	Geraniol	Essential oil	Mitsui Petrochemical Industries (Japan)
Catharanthus roseus	Arbtin	Pigment	Mitsui Petrochemical Industries (Japan)
Carthamus tinctorius	Carthamin	Pigment	Kibun (Japan)
Vanilla planifolia	Vanillin	Flavor	ESCAgenetics (USA)
Beta vulgaris	Betacyanins	Color	Nippon Shinyaku (Japan)
Euphorbia milli *Aralia cordata*	Anthocyanins	Dye, color	Nippon Paint (Japan)

described in Section 2. Nevertheless, despite these challenges, a number of compounds have been successfully produced on an industrial scale using plant cell culture, as summarized in Table 1.

This chapter aims to give a brief background of the types of plant cell culture, the technology used to culture these cells, and a discussion of some of the genetic tools used to facilitate metabolic engineering in these cells. We then present a series of short sections that describe the biosynthesis of specific natural products in more detail. A comprehensive description of all natural products produced in plant cell culture is beyond the scope of this chapter. Instead, we selected compounds that are clinically relevant in

modern medicine, have been subjected to extensive optimization in plant cell culture and which also represent the vast structural variety of compounds that are found in plants. Plant natural products associated with herbal medicine, or that are used as pigments or dyes are not discussed in detail. Furthermore, we do not discuss in detail the production of plant extracts (for example, from *Ginkgo biloba*) that do not contain a defined mixture of specific natural products. Plants have also been used for the production of medically useful proteins and antibodies ('plantibodies') [11], but in this chapter small molecule production is the focus.

2 Techniques of plant cell culture

2.1 Classes of plant cell culture

In 1934 it was first reported that plant cells could be cultivated on synthetic media in an undifferentiated state and appeared to be capable of unlimited growth [5, 6]. In other words, for the first time, plant material could be used to generate tissue culture. Plant tissue culture has had a tremendous impact in the genetic and biochemical study of natural product biosynthesis and in some cases tissue culture can be used for the large-scale production of plant natural products. The major types of plant tissue culture are described below (Fig. 2). A recent survey has summarized over 90 reports of successful cell culture of medicinal plants [12].

To initiate plant tissue culture, seeds are typically sterilized and germinated under aseptic conditions on solid, agar containing media. The sterile plantlet is then used to generate cell culture. Since sterilization conditions are harsh, it is often easier to sterilize the surface of the seed instead of the soft tissue ultimately used to yield the tissue culture.

Common plant media are Murashige and Skoog, Gamborg, Schenk and Hildebrandt and White [13]. Each of these media recipes contains minerals, vitamins and a carbon source, usually sucrose. Although plant cell cultures are typically initiated on solid media, liquid media is required for large-scale production. It has been commonly observed that changing the amounts of minerals or carbon source in the media often has a profound impact on the levels of metabolite production [14]. The optimal media for natural product production, as with microbial organisms, must be deter-

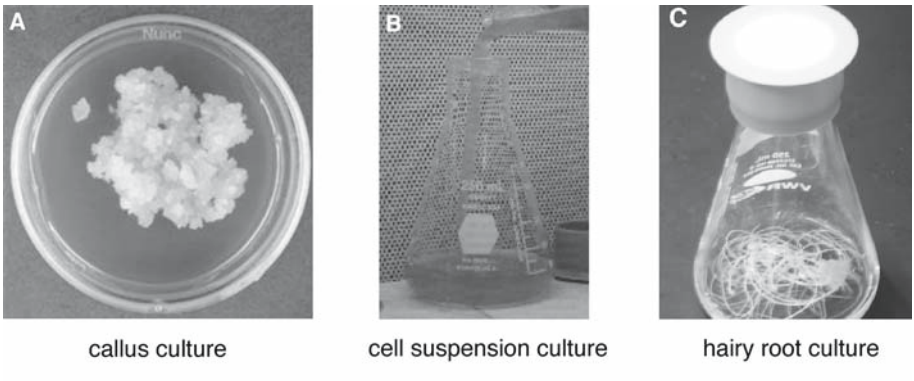


Figure 2
A Callus (copyright DSMZ GmbH, 2007), **B** cell suspension (from http://www1.coe.neu.edu/~clee/research/subculturing/index.html), and **C** hairy root cultures

mined empirically, though lowering the amount of phosphate and nitrate usually increases secondary metabolism [15].

An explant – a piece of the sterile plant tissue – from the seedling is then placed on solid culture, and dedifferentiated cells begin to grow. Notably, plants are characterized by totipotency, meaning that each cell can express its genetic potential and can form a complete, fertile plant. Therefore, to maintain cell cultures in a dedifferentiated state, exogenous phytohormones – a mixture of auxins and cytokines – are added to the media to prevent the cells from differentiating. The specific mixture used varies from plant to plant. Changes in the phytohormone ratio and concentration also affect what organ – shoot or root – is regenerated. Over extensive periods of subculturing, cultures may become hormone independent and retain their dedifferentiated state even in the absence of exogenous phytohormones. The mass of cells that result from the explant are known as callus culture. Callus is typically subcultured every 3–6 weeks and can be maintained for decades. These are heterogeneous cultures and the heterogeneity can be reduced by extensive subculturing, typically with the fastest growing cells being selected for. Callus cultures are typically slow growing and not suitable for large-scale bioreactors.

When callus is suspended in liquid media, the callus breaks into small cell aggregates and cell suspension cultures result. Single cell suspensions are rare, and plant suspension cultures are usually composed of aggregates

of cells approximately 40–200 μM in diameter [16]. Cell suspension cultures usually proliferate more rapidly than callus and are one of the most commonly used types of cultures in industry. These cultures are generally considered to be the most suitable cell types for large-scale cultures maintained in a bioreactor.

Hairy root cultures, which are obtained by infecting plants with *Agrobacterium rhizogenes*, are also widely used in basic research and in industrial production studies. When a plant is wounded or cut, signal molecules are produced by the plant at the site of the wound. This allows *Agrobacterium* to infect the plant through the wound, and the bacterial genes Rol A B and C are transferred to the plant and integrated into the plant genome [17]. These genes stimulate root production, resulting in the formation of hairy roots at the site of infection within several weeks. The roots can be removed, adapted to liquid media and grown on large scales. Auxin metabolism is altered in the transformed hairy roots, so exogenous auxins are not required to maintain the root propagation and keep the tissue from regenerating to the whole plant. Hundreds of plants can be transformed into hairy roots as reviewed [18, 19] though development of hairy root culture from some species, particularly the *Papaveraceae* and *Ranunculaceae*, has been difficult [20]. In general, hairy roots are genetically stable [21] and most have a doubling time of 1–2 days, which is similar to that of cell suspension cultures [22]. Not surprisingly, most compounds formed in hairy root cultures are also found in the roots of the differentiated plant. Interestingly however, hairy root cultures of *Artemisia annua* accumulate artemisinin, a natural product that is normally accumulated in the aerial parts of the plant, and might not be expected to be found in root cultures [23, 24] (see Section 3).

Root cultures can also be formed by adjusting the hormone balance of the callus culture media. When the appropriate ratio of phytohormones is added to the callus, roots are formed in the callus culture, which can then be excised and cultured. However, it is widely reported that hairy roots derived from *Agrobacterium* transformation are faster growing and generally accumulate higher titers of natural products than non-transformed root cultures [25].

Shoot cultures, which lack roots and undifferentiated callus cells, have also been developed. These cultures are either initiated from the seedling or can be generated from callus by adjusting the phytohormone content

of the medium. Shoot cultures can be cultivated on liquid or solid media in the presence of phytohormones. These cultures have been reported to produce certain natural products that cannot be produced in cell suspension culture – for example, artemisinin and vindoline [26]. Typically, shoot cultures accumulate compounds normally observed in the aerial parts of the plant. However, shoot cultures are not well suited for bioreactor production and are unlikely to replace plants for large-scale natural product production.

2.2 Scale up of plant cell culture

For laboratory studies, shake flask cultures are commonly used. However, industrial fermentation of natural products requires highly reproducible conditions in which the mixing, shear forces and gas content do not vary when the culture is grown at different volumes. A bioreactor has been employed in industrial fermentations to provide a highly controlled culture environment. Plant cells were first cultured in a 10 L bioreactor made of glass or steel in 1959 [27]. The first large-scale production of *Nicotiana tabacum* cells in a 20,000 L stirred reactor was reported in 1977 [28].

Plant cell cultures require low shear mixing, good aeration and absolute sterility. The stirred tank bioreactor predominates in the literature reports of plant cell culture (Fig. 3). A mechanical stirrer fitted with a variety of blades is used for mixing, where the best blade type depends upon the individual culture. Plant cells are highly shear sensitive, so slow mixing speeds of 100–300 rpm are used. The air lift fermentor, also commonly used with plant cell cultures, couples aeration to agitation and does not contain a mechanical stirrer (Fig. 3).

Compared to microbial cell cultures, plant cell cultures grow slowly with a doubling time of 1–2 days at best. Production levels are in the range of 10–500 mg per L per day, and products are usually stored intracellularly, not exported into the media as they are in most microbial fermentations. These factors lead to an overall less efficient production system so maximizing the cell density of the cultures is essential for making plant cell cultures practical.

Shoot cultures, due to mechanical frailty and light requirements are generally not suited for bioreactor production [16]. Root cultures, however

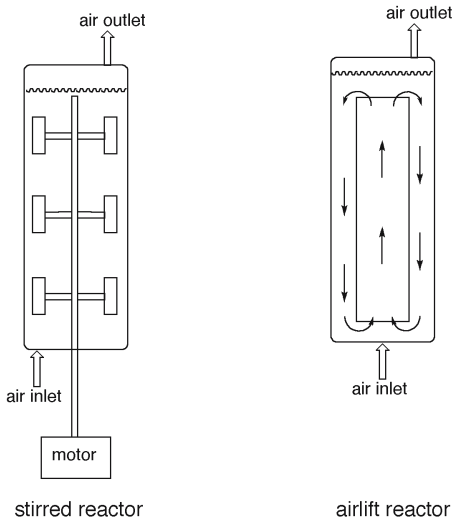
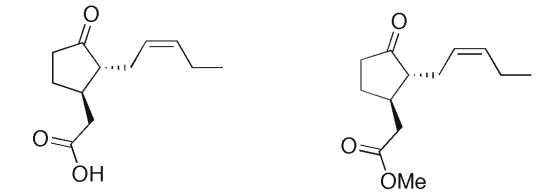


Figure 3
Schematic of stirred and airlift bioreactors

have been successfully cultivated in a variety of bioreactor set ups. Often, roots are treated as immobilized cells; the roots are allowed to adhere to mesh or stainless steel supports in the tank and no mechanical stirring is used. Immobilized cells have also been explored by adsorption of cell suspension cultures on glass, by covalent cross-linking or by entrapment of the cells in agarose [13]. However, since products are not exported into the media, the release of product from the immobilized cells is often problematic.

2.3 Elicitation of natural products in cell culture

A key advance in the development of plant cell cultures was the discovery of elicitors, or small molecules that stimulate natural product production. Elicitors fall into two main categories, biotic and abiotic. Biotic elicitors include fungal derived saccharides and proteins that stimulate a defensive response in plants. Abiotic elicitors include chemicals such as jasmonates, salicylic acid, and hormones (Fig. 4). Additionally, various metals such as



jasmonic acid methyl jasmonate

Figure 4
Chemical structures of jasmonate elicitors

vanadium, silver and lanthanum have also stimulated natural product production [29]. Elicitation has been used to improve the levels of paclitaxel and shikonin in industrial scale cultures (see Section 3).

2.4 Genetic manipulation in plant cell culture

Genetic manipulation of plant cells has the potential to profoundly impact plant natural product production. For example, genes expressed at low levels can be overexpressed at higher levels to make a slow step more efficient. Additionally, genes from other metabolic pathways can be transformed into the wild type pathway to expand the scope of natural products produced. Specific examples of metabolic engineering efforts designed to improved natural product production in plant cell culture are described in Sections 3.3 and 3.4.

Transformation of plants with foreign genes was largely impractical until the mechanism of crown gall formation by *Agrobacterium tumefaciens* was elucidated in the 1970s [30, 31]. It was discovered that the *Agrobacterium* transferred DNA material to the plant (T-DNA) and that this DNA was incorporated into the plant genome. Thus, a natural means of plant transformation was identified. The first transgenic tobacco plants expressing recombinant DNA in an integrated T-DNA sequence was reported in the early 1980s [32].

Unfortunately, *Agrobacterium* transformation is limited to certain species of plants, though recent advances have expanded this range [33]. Therefore, alternative methods of transformation are required. Particle

bombardment, reported in the late 1980s [34], uses a device to accelerate small tungsten particles coated with DNA at plant cells at a velocity of approximately 400 m/s. This method is generally applicable to plant cells, though subsequent rearrangement of the DNA sequence can be caused by the force used for transformation. Finally, plant cell protoplasts can be formed by partially removing the plant cell wall either enzymatically or mechanically. Foreign DNA can then be introduced into the cell by standard methods such as electroporation. However, protoplasts usually take long culturing periods for cells to recover, and some cell strains do not appear to survive the process [35].

A marker gene is also transferred to allow transformed cells to be easily selected [36, 37]. Common marker genes encode enzymes that confer antibiotic resistance. Aminoglycoside adenyltransferase, which confers resistance to spectinomycin, glycopeptide binding protein, which confers resistance to zeocin and hygromycin phosphotransferase, which confers resistance to hygromycin, are all widely used. Additionally, the marker gene mannose-6-phosphate isomerase confers upon transformed cells the ability to use mannose as the sole carbon source.

3 Natural products produced from plant cell culture

The production of many natural products in plant culture has been explored. This section describes the advances made in the production of several pharmaceutically important natural products.

3.1 Taxanes

Paclitaxel, also known by its trade name Taxol from Bristol Myers Squibb, is one of the success stories of plant cell culture (Fig. 5). This densely functionalized diterpene natural product exhibits potent anticancer activity. In 1963, the bark of *Taxus brevifolia* was found to harbor antitumor activity in a natural product extract screen at the National Cancer Institute [38]. Paclitaxel, or Taxol, was discovered to be the active agent of this extract in 1971 [39]. Paclitaxel acts by binding to tubulin, promoting the assembly of microtubules and stimulating apoptosis [40]. At the time of

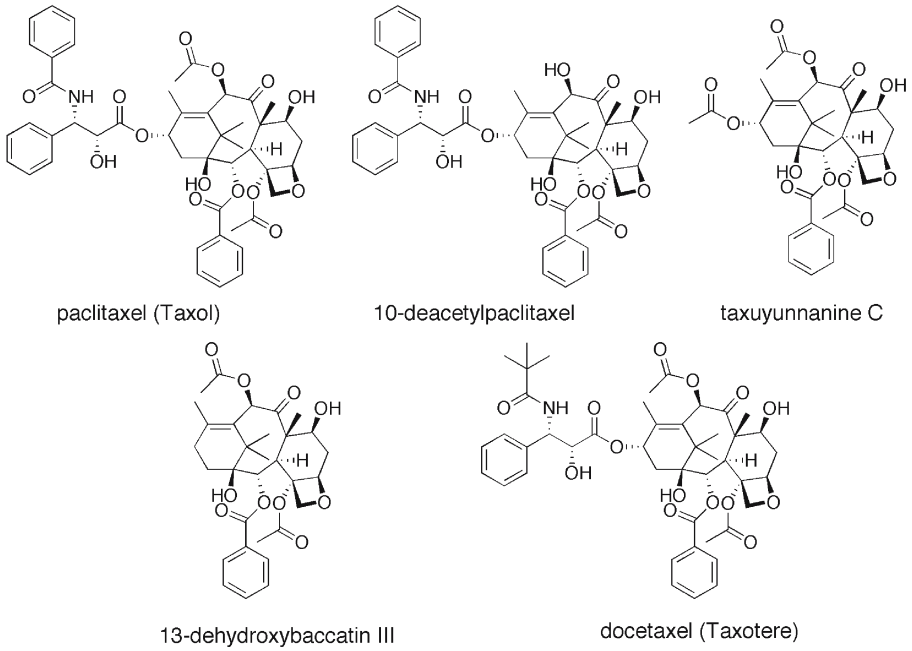


Figure 5
Structures of taxanes

its discovery, this represented a completely novel mechanism of cytotoxic activity. Recently an additional target for Taxol has been identified as Bcl-2, a human protein involved in apoptosis [41]. Paclitaxel is used against a variety of cancers including breast cancer, ovarian cancer, non-small cell lung cancer and Kaposi's sarcoma [42]. Current sales of paclitaxel and its derivatives are expected to exceed 2 billion US dollars [43].

Paclitaxel is found in the bark of *Taxus brevifolia*, the Pacific yew tree, in small amounts: approximately 0.01% of the dry weight of the bark [44]. The bark is non-renewable – harvesting of the bark results in the death of the tree – and for many years it appeared that no sustainable supply of the drug existed. The structural complexity of the molecule prevented an economic total synthesis for industrial scale production. A number of related diterpene compounds referred to as taxanes are found in a variety of other *Taxus* species. One of these related taxanes baccatin III, a paclitaxel intermediate, is found in the needles of *Taxus baccata*, the European Yew tree,

Table 2
Plant cell suspension cultures producing paclitaxel (Adapted from Baldi A, Bisaria VS, Srivasrava AK (2007) Biotechnological approaches for the production of some promising plant based chemotherapies. In: O Kayser, W Quax (eds.): *Medicinal plant biotechnology*. Wiley VCH, Weinheim, 128)

Cell type	Paclitaxel (mg/L)	Culture type	Reference
Taxus canadensis	117	Shake flask, MeJa	[74]
Taxus media	110	Shake flask, MeJa	[26]
Taxus media	21	Stirred reactor, MeJa, precursor	[82]
Taxus cuspidata	22	Wilson type	[22]
Taxus chinensis	138	Shake flask, silver nitrate	[49]

and is currently used to yield commercial supplies of paclitaxel through semi-synthetic strategies [45]. However, although harvesting of the needles it not lethal, it can still have a negative impact on *Taxus* populations. A paclitaxel source from plant cell culture, which is completely renewable, would provide a valuable strategy for paclitaxel production.

Callus cultures of various *Taxus* species were initiated, and the highest producing cultures were converted to suspension cultures. Disappointingly, *Taxus* cell suspension cultures proved to be exceptionally slow growing [43]. A critical advance was reported in 1996 when it was discovered that addition of methyl jasmonate substantially improved the production of paclitaxel to 0.5% of the dry weight of the plant cell culture [46]. These results were the first suggestion that plant cell cultures might be able to produce taxanes at levels appropriate for commercial production.

A variety of factors have been shown to improve the production of paclitaxel and other related taxanes. Methyl jasmonate, as mentioned above, has been shown to consistently improve levels of taxane production. Heat shock, mechanical stresses such as ultrasound, and peptide regulators (Phytosulfokine A) have also resulted in the improvements of taxane production [47]. Immobilization of cell suspension cultures of *T. cuspidata* on glass fiber mats resulted in production levels of paclitaxel (~0.012%) [48]. Levels of 10–22 mg/L are commonly reported but higher levels of about 100 mg/L have also been reported (Tab. 2).

Taxol and taxol precursor production in cell cultures has been scaled up, though use of these cultures for industrial production is still somewhat limited. Yields can vary from cell culture to culture, and production of

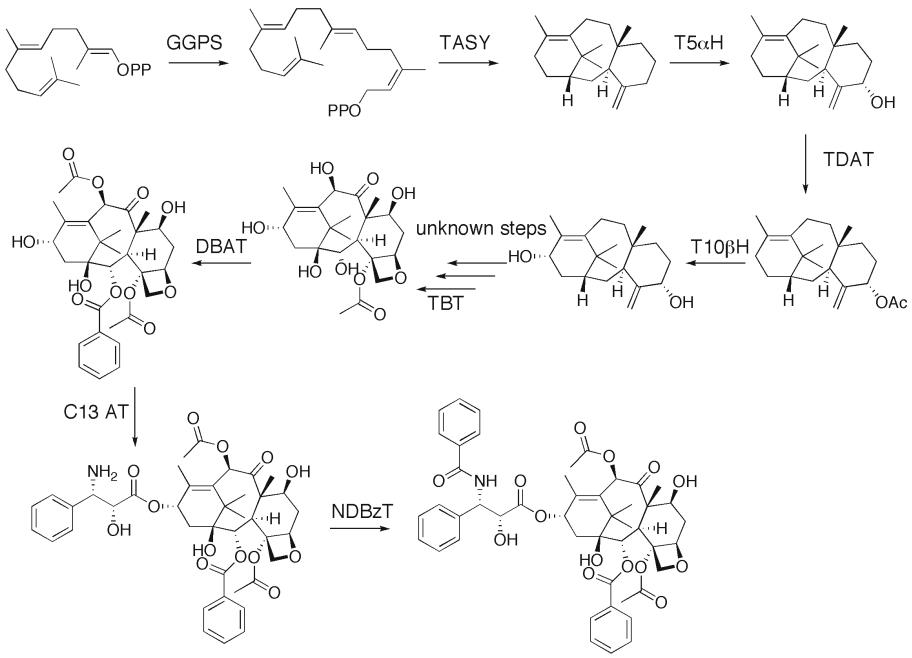


Figure 6
Paclitaxel biosynthetic pathway. GGPS (geranyl geranyl pyrophosphate synthase); TASY (taxadiene synthase); T5αH (taxadiene 5α hydroxylase); TDAT (taxadiene5a-ol O-acetyltransferase); T10βH (taxane 10β-hydroxylase); TBT (taxane 2a-O-benzoyltransferase); DBAT (10-deacetylbaccatin III-10-O-acetyltransferase); PPT (phenylpropanoyl transferase); NDBzT (3′-N-debenzoyl-2-deoxytaxol-N-benzoyltransferase)

paclitaxel is not favored over other taxanes. Nevertheless, bioreactors up to 75,000 L are being employed by ESC genetic (USA), Phyton (USA), Phyton Biotech (Germany), and Samyang Genex (Korea) for the commercial production of paclitaxel [16].

Use of *Taxus* cell cultures has also made a significant impact in the elucidation of the paclitaxel biosynthetic pathway. In a *tour de force*, sequencing of a highly productive, elicited cell culture line resulted in the identification of most of the genes in the pathway [49]. An abbreviated biosynthetic scheme is shown in Figure 6, though the exact sequence of biosynthetic transformations is not clear due to the difficulties in assigning the substrate specificity of many of the enzymes. A recent study that investigated the time resolved expression of these genes in *Taxus* cell culture may

shed further light on the timing of the biosynthetic transformations [50]. Additionally, expression of these genes in *E. coli* and *S. cerevisiae* has led to the microbial production of several advanced intermediates of paclitaxel [51].

An endophytic fungal strain isolated from the yew tree that produces small amounts of paclitaxel was reported in 1993 [52]. Since that time, several other taxane producing endophytic fungal strains have been isolated [53]. However, the amounts of taxanes produced by these fungal cultures are quite low (ng-µg per liter) and production levels are often unstable; therefore, these fungal strains are not likely to replace plant-derived cultures in the immediate future.

3.2 Camptothecin

Camptothecin (Fig. 7) is a potent anticancer agent, whose mode of action is topoisomerase I inhibition. However due to its severe side effects, camptothecin itself is not used for treatment of cancer though semi-synthetic derivatives of camptothecin have shown more useful clinical properties. First generation camptothecin derivatives irinotecan and topotecan (Fig. 7) are used for clinical treatment of lung cancer, and have also shown therapeutic value in the treatment of colon cancer, uterine cervical cancer and ovarian cancer [54]. Progress in the development of new camptothecin analogs to improve pharmacology is reviewed elsewhere [55, 56]. The demand for camptothecin and its derivatives is one ton per year and current worldwide sales of clinically used camptothecin derivatives are approximately one billion dollars per year [57]. Since relatively low levels of camptothecin are isolated from natural sources, plant cell culture has been extensively explored as a production model for camptothecin.

Camptothecin, a monoterpene indole alkaloid, was originally isolated from the sap of the stem wood of the Asian tree *Camptotheca acuminata* [58]. All parts of *C. acuminata* contain some camptothecin, with the highest levels found in young leaves at 0.4–0.5% dry weight (DW) [59]. Many derivatives of camptothecin have been isolated from plant extracts and cultures. The most abundant natural derivatives are the water soluble 10-hydroxycamptothecin [60] and 9-methoxycamptothecin [61], both of which display *in vitro* antitumor activity comparable to camptothecin [62].

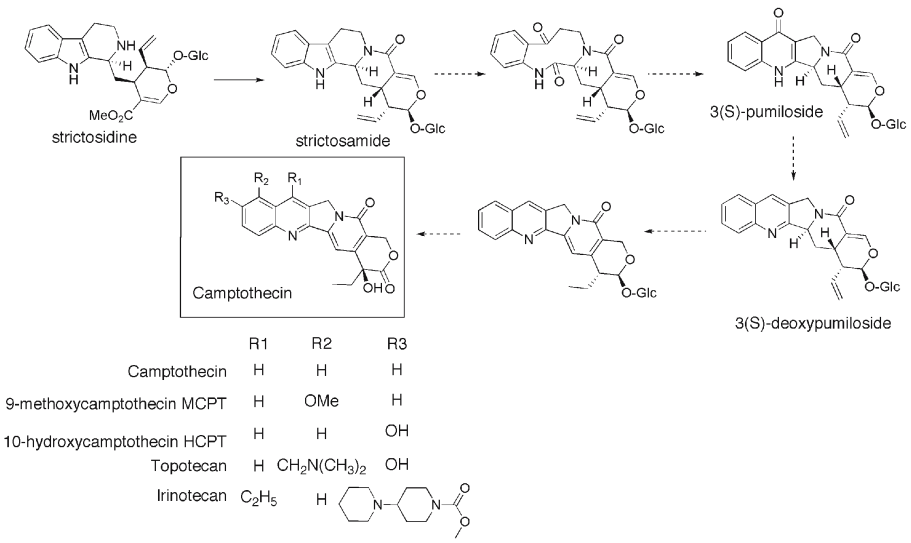


Figure 7
Biosynthetic scheme for camptothecin. Structure of camptothecin and derivatives

The initial step in biosynthesis of camptothecin – along with all other monoterpene indole alkaloid natural products – is the condensation of tryptamine and secologanin to form the biosynthetic intermediate strictosidine (Fig. 7). At this stage, the biosynthesis of camptothecin diverges from other terpene indole alkaloids; typically, strictosidine is immediately deglycosylated, but in the camptothecin pathway strictosidine instead forms the strictosamide intermediate (Fig. 7). The biosynthetic conversion of strictosamide to camptothecin is not known on the enzymatic level. However, intermediates observed in camptothecin producing cultures [63, 64] has led to the proposal of a reasonable series of biosynthetic transformations (Fig. 7). The lack of genetic and enzymatic knowledge of the biosynthesis of this important class of alkaloids greatly hinders metabolic engineering efforts to improve the production of camptothecin and its derivatives.

Although camptothecin production in plants is sustainable, a readily fermentable source could lower the cost and provide improved access to the compound and its derivatives. Investigation into the production of

camptothecin and its derivatives in plant tissue culture was first reported for *Camptotheca acuminata*. Yields of 0.0002% DW were reported for *C. acuminata* cell suspension cultures [65], a significant drop from the levels found in the whole plant, which range from 0.02–0.5% DW depending on the tissue analyzed [59] (Tab. 3). Media optimization studies revealed that a nitrogen concentration of 70 mM in the media gave the highest biomass of *C. acuminata* cell suspension cultures, while a NH^{4+}/NO^{3-} molar ratio of 5:1 (total of 40 mM nitrogen) gave the maximum camptothecin yield [66]. This is similar to results found in suspension cultures of *Atropa belladonna* for the production of tropane alkaloids [67]. Implementation of a two stage bioreactor, where the culture is grown in growth medium containing 70 mM nitrate for the first 18 days, followed by growth in production medium containing 40 mM NH^{4+}/NO^{3-} (5:1), gave the highest reported yield of camptothecin at 12.8 mg/L (Tab. 3). Callus cultures of *C. acuminata* with production of 0.2% DW have also been reported [68].

Mature trees of *Nothapodytes foetida* produce 0.075% DW camptothecin and 0.013% 9-methoxycamptothecin in the shoots, and 0.5 and 0.06% DW camptothecin and 9-methoxycamptothecin, respectively, in the seeds [69]. The first cell cultures of *N. foetida* produced 100 times less camptothecin than intact plants [69, 70]. However, more recently established cell suspension cultures of *N. foetida* have been found to produce camptothecin in levels comparable to *C. acuminata* cultures (approximately 0.1% DW) (Tab. 3) [71]. This cell line also produced high amounts of 9-methoxycamptothecin (0.08% DW), which is not a major alkaloid of *C. acuminata*.

Differentiated hairy root cultures of *C. acuminata* have been established by transformation with *Agrobacterium rhizogenes* [72]. These cultures displayed respectable production of camptothecin at 0.1% DW, comparable to levels found in the intact plant [59]. Additionally *C. acuminata* hairy roots yielded 10-hydroxycamptothecin in 0.015% DW compared to 0.002% DW in the intact plant [59]. Hairy root cultures of *Ophiorrhiza pumila*, *O. liukiuensis* and *O. kuroiwai* have also yielded camptothecin and its derivatives. *O. pumila* hairy root culture medium was found to contain substantial amounts of camptothecin, which could be increased by the addition of a polystyrene resin Dianion HP-20 to the medium [73]. The production of camptothecin after resin addition was only slightly enhanced; however, the overall secretion into the medium was increased five-fold, dramatically simplifying the isolation and purification process.

Table 3 Select examples of production of camptothecin and its analogs in various plants and plant cell culture
CSC is cell suspension culture, CPT is camptothecin, HCPT is 10-hydroxycamptothecin, MCPT is 9-methoxycamptothecin

Species	Culture	Conditions	CPT	HCPT	MCPT	Ref.
			% DW (mg/L)			
C acuminate	Leaves		0.4–0.5	0.003		[59]
	Roots		0.04	0.002		[59]
	Bark		0.18–0.2	0.009		[59]
	CSC	Control	0.009			[66]
		2 stage bioreactor Optimal nitrogen	0.036 (12.8)			
	Callus		0.232	0.008		[68]
	Hairy roots		0.1	0.015		[72]
N foetida	Shoot		0.075		0.013	[69]
	Seed		0.5		0.06	[69]
	CSC		0.11(35)		0.08 (26)	[71]
O pumila	Leaves		0.03			[73]
	Young roots		0.1			[73]
	Hairy root		0.1 (8.9)			[73]
		3L bioreactor	(8.7)			[75]
O luikiunesis	Hairy root		0.083 (4.6)		(0.5)	[76]
O. Kuroiwai	Hairy root		0.022 (9.4)		(0.3)	[76]

A similar effect was seen after addition of resin to *Catharanthus roseus* hairy roots producing the terpene indole alkaloid ajmalicine [74]. Production of this *O. pumila* hairy root line in a 3 L bioreactor did not affect production of camptothecin compared to cells cultured in shake flasks [75]. *O. liukiuensis* and *O. kuroiwai* hairy root lines have been established with different levels of camptothecin and 9-methoxycamptothecin levels compared to *O. pumila* (Tab. 3) [76].

Despite these advances however, the production of camptothecin in cell cultures is far below the levels needed for commercial production of these compounds. The identification of enzymes involved in the biosynthesis of this important medicinal compound will undoubtedly impact its production in plant cell cultures.

3.3 Tropane alkaloids

The tropane alkaloids hyoscyamine and scopolamine (Fig. 8) function as acetylcholine receptor antagonists and are clinically used as parasympatholytics. Scopolamine has a higher commercial market than hyoscyamine, though its yield from plants is much lower. The tropane alkaloids are biosynthesized primarily in plants of the family *Solonaceae*, which includes *Hyoscyamus*, *Duboisia*, *Atropa*, and *Scopolia*; each of these species are capable of biosynthesis of both nicotine and tropane alkaloids. Early studies revealed that tropane alkaloids are formed in the roots and then transported to the aerial parts of the plant [77]. This localization information suggested that *Agrobacterium rhizogenes* induced hairy root cultures could be used for high yield production of tropane alkaloids. Several hairy root cell lines that produce tropane alkaloids have been established (Tab. 4) [78].

Metabolic engineering is a powerful strategy often used to improve production of secondary metabolites. However, engineering efforts require knowledge of the enzymes involved in the biosynthetic pathway and the mechanism by which these enzymes are regulated. Tropane alkaloid biosynthesis has been studied at the biochemical level and several enzymes from the biosynthetic pathway have been isolated and cloned [79]. A summary of the biosynthetic steps is shown in Figure 8. Knowledge of the enzymes responsible for catalyzing key biosynthetic steps, in conjunction with *Agrobacterium* mediated transformation of plants, provided an excellent platform to initiate metabolic engineering efforts with this plant alkaloid. The levels of tropane alkaloid production in a variety of hairy root cultures were altered by overexpression of methyltransferase putrescine-N-methyltransferase (PMT) and hyoscyamine 6B-hydroxylase (H6H) (Fig. 8).

Hyoscyamine 6B-hydroxylase (H6H) catalyzes the hydroxylation of hyoscyamine to 6B-hydroxyhyoscyamine, as well as the epoxidation to scopolamine (Fig. 8) [80, 81]. H6H was first cloned from *Hyoscyamus niger* [82]. Transformation and overexpression of the H6H clone from *H. niger* into *Atropa belladonna*, under control of the cauliflower mosaic virus 35S (CMV 35S) promoter, resulted in plants with increased scopolamine production [83]. This demonstration of improved alkaloid production by heterologous expression of a pathway enzyme in intact plants led researchers to seriously explore genetic strategies as a method for increasing produc-

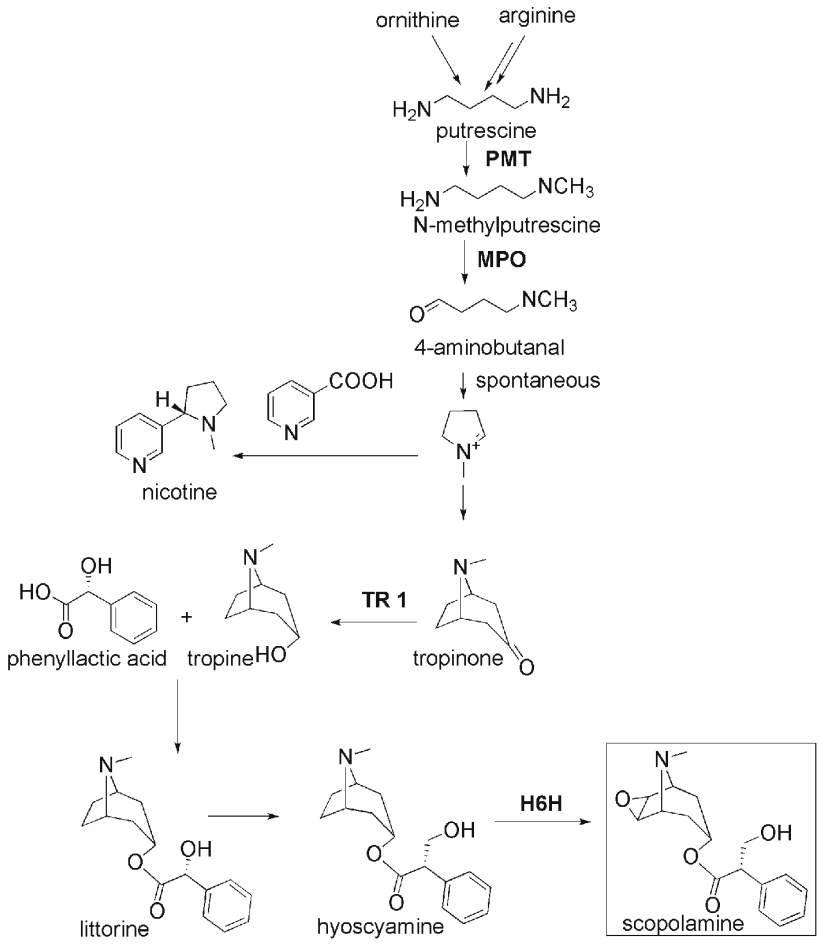


Figure 8
Tropane alkaloid biosynthesis. PMT (putescine N-methyltransferase); MPO (N-methylputrecine oxidase); TR 1(tropinone reductase 1); H6H (hyoscyamine 6-βhydroxylase)

tion and altering the alkaloid profiles of tropane alkaloid producing plant hairy root cultures.

Transformation of plants with *Agrobacterium rhizogenes* is a common technique used to heterologously express genes in plant hairy root culture as described in Section 2. Using this transformation system, the cDNA encoding *H. niger* H6H was expressed in hairy root cultures of various

Table 4
Select examples of scopolamine and hyoscyamine production in wild type and transformed hairy root cultures
*Hn*H6H is *Hyoscyamus niger* H6H, NtPMT is *Nicotiana tabacum* PMT

Species	Condition	Scopolamine	Hyoscyamine	Ref
		%DW (mg/L)		
Duboisia hybrid	Tranformed with *Hn*H6H	2.48 (74.5)	0.14 (4.5)	[88]
	Control	1.16(24.7)	0.39 (8.3)	
Atropa baetica	Transformed with *Hn*H6H	0.385	0.02	[84]
	Control	0.29	0.50	
Scopolia parviflora	Transformed with *Nt*PMT	0.52	0.35	[89]
	Transformed with *Hn*H6H	0.81		[87]
	Control	0.08	0.10	
Hyoscyamus niger	Transformed with *Hn*H6H and *Nt*PMT	(411.2)		[85]
	Control	(43.7)		
	Transformed with *Hn*H6H	(184.4)		
Hyoscyamus muticus	Transformed with *Hn*H6H	0.11 (14.4)	0.98 (125.4)	[86]
	Control	(0.12)	(134.9)	

Solonaceae plants (Tab. 4). The overexpression of H6H under control of the CMV 35S promoter in transgenic hairy root cultures of *Atropa baetica* [84]. *Hyoscyamus niger* [85], *Hyoscyamus muticus* [86], *Scopolia parviflora* [87], and a *Duboisia* hybrid [88] cultures systematically resulted in an increase in the production of scopolamine when compared to control cultures not expressing H6H (Tab. 4). Therefore, transgenic expression of H6H appears to be a general method for increasing scopolamine production in hairy root cultures of tropane alkaloid producing plants.

The conversion of putrescine to the N-methylputrescine by putrescine N-methyltransferase (PMT) is the first committed step in the biosynthesis of both tropane alkaloids and nicotine (Fig. 8). Heterologous expression of *Nicotiana tabacum* PMT in *Scopolia parviflora* resulted in an eight-fold

increase in scopolamine production and a 4.2-fold increase in hyoscyamine production [89]. A similar effect was observed with the expression of *Nicotiana tabacum* PMT in *H. muticus* and *D. metel* [90]. However the expression of *Nicotiana tabacum* PMT in other tropane alkaloid producing hairy root cultures including *Hyoscyamus niger*, *Atropa belladonna*, *Duboisia* hybrid had no effect on alkaloid production [85, 91, 92]. H6H and PMT were co-expressed in *Hyoscyamus niger* on a plasmid containing two separate expression cassettes both under the control of the CMV 35S promoter. Co-expression of H6H and PMT resulted in the highest production of scopolamine in hairy root culture reported to date of 411.2 mg/L. This was a 10-fold increase over control cultures and a 2–3-fold increase over cultures only expressing H6H (Tab. 4) [85].

The tropane alkaloid biosynthetic pathway has been used as a model system for the generation of high alkaloid producing cell cultures through genetic engineering. The extent to which these cultures can be modified is dependent on the identification of new genes encoding regulatory and biosynthetic enzymes involved in tropane alkaloid secondary metabolism.

3.4 Benzylisoquinoline alkaloids

The benzylisoquinoline alkaloids (BIA) include medicinally important compounds such as the analgesics morphine and codeine, and the antibiotics berberine, and sanguinarine (Fig. 9). Although used for over a century, morphine and codeine are still two of the most important analgesics in medicine worldwide. Additionally, the morphinan alkaloid thebaine is converted semi-synthetically to yield the analgesics oxycodone and buprenorphine (Fig. 9). Due to their complexity and multiple chiral centers, commercial synthesis of morphine, codeine and other BIA is not economically feasible, and opium poppy plants remain the main commercial source of the alkaloids [93].

Development of plant cell cultures for the BIA producing medicinal plants *Eschscholzia californica*, *Papaver somniferum* and *Coptis japonica* aided in the isolation and cloning of many of the enzymes involved in the biosynthesis of BIA (for reviews see [94, 95]). In brief, the biosynthesis of BIA begins with the condensation of dopamine and 4-hydroxyphenylacetaldehyde, both derived from tyrosine, by norcoclaurine synthase to form (S)-

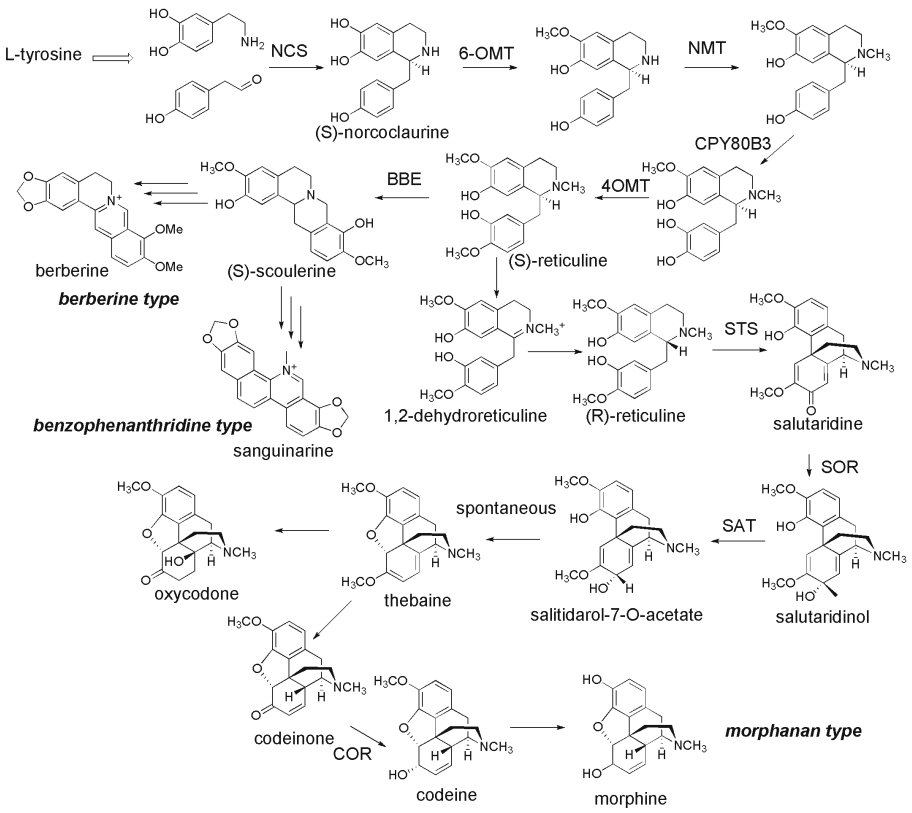


Figure 9
Tetrahydroisoquinoline biosynthesis. NCS (norcoclaurine synthase); 6-OMT (O-methyl transferase); NMT (N-methyltransferase); CPY80B3 (cytochrome P450); 4OMT (O-methyl transferase) BBE (berberine bridge enzyme); STS (salutaridinol synthase); SAT (salutidarinol-7-O-acetyltransferase), COR (codeine reductase)

norcoclaurine (Fig. 9). (S)-Norcoclaurine is converted to (S)-reticuline by a number of enzymes, all of which have been isolated and cloned. (S)-reticuline is the last common intermediate for the berberine, benzophenanthridine, and morphinan alkaloids. In morphine biosynthesis, (S)-reticuline is converted to (R)-reticuline through a 1, 2-dehydroreticuline intermediate. Many of the downstream enzymes of this pathway have been isolated and cloned; for example, codeinone reductase and salutidarinol-7-O-acetyltransferase have very recently been cloned [96, 97].

In the biosynthesis of berberine, an antibacterial benzylisoquinoline alkaloid, (S)-reticuline is converted to (S)-scoulerine (Fig. 9). All the enzymes involved in the biosynthesis have been isolated and cloned except the final enzyme. Overproduction of berberine in *C. japonica* cell suspension cultures was achieved by selection of a high producing cell line [98], with reported productivity of berberine reaching 7 g/L [99]. This was one of the first demonstrations of production of a benzylisoquinoline alkaloid in cell culture at levels necessary for economic production. Another BIA, sanguinarine, a benzophenanthridine type alkaloid, is used as an antibiotic and antiplaque agent in toothpastes and mouthwashes. It is commercially isolated from the root of *Sanguinaria canadensis* L. The production of sanguinarine in cell suspension cultures has been optimized by elicitation [100], medium optimization [101], resin addition [102], and metabolic engineering [103] to a maximum reported yield of 2% DW in cell suspension cultures of *Papaver somniferum* after elicitation.

The morphinan alkaloids, the most desirable benzylisoquinoline alkaloids from an economic perspective, are not produced in high amounts in cell suspension cultures. Several strategies for increasing the production of morphinan alkaloids in plants and plant cell cultures have been employed including media optimization, differentiated cell cultures, RNA mediated suppression of pathway enzymes, and overexpression of rate limiting pathway enzymes.

The low production of morphinan alkaloids in cell suspension cultures is thought to be due to the need for differentiation for production of these secondary metabolites. Removal of hormones from the media is often used to induce embryogenesis and promote cellular differentiation. In one study, a two stage process where cell suspension culture were grown in media containing hormones and then transferred to hormone free media for production of secondary metabolites, resulted in increased production of morphinan alkaloids [104]. Often the removal of hormones results in the formation of shoots roots and meristemoids. In this study the culture maintained its dispersed character after removal of hormones, and was maintained in hormone free medium for 9 months. Removal of exogenous hormones from the medium of cell suspension cultures of *Papaver somniferum* resulted in the accumulation of 0.3% DW and 0.25% DW codeine and morphine respectively, a three-fold increase over control cultures supplied with hormones [104]. However, these levels are still significantly lower than the

1.23–2.45% commonly found in field grown whole plant extracts and the 20% DW found in the latex of *Papaver somniferum*. Since differentiated cell cultures often lead to increased production of certain secondary metabolites, hairy root cultures of *Papaver somniferum* and *Eschscholzia californica* were established in the hopes that benzylisoquinoline alkaloid production would be increased. Production of morphine, codeine and sanguinarine was achieved in several hairy root lines of *P. somniferum* [105]. Transformed root cultures obtained by infection with *Agrobacterium* contained a 1.4-fold higher alkaloid content compared to non-transformed roots and yielded a three-fold increase in production of codeine (0.18% DW) when compared to non-transformed roots. Morphine was produced in similar concentration to non-transformed roots at 0.26% DW. The hairy roots also produced sanguinarine (0.02% DW), which is not produced in non-transformed roots. Additionally, the hairy root culture media contained 0.26 and 0.014% DW morphine and sanguinarine respectively.

RNA mediated suppression of pathway enzymes is a potential tool for the metabolic engineering of secondary metabolism. Blocking the action of enzymes at pathway branch points could lead to an increase in the levels of a central intermediate that could in turn be converted to a greater amount of the desired end product. Suppression of enzymes may also give insight into regulatory effects of certain pathway enzymes. Both berberine and morphinan alkaloids arise from a central intermediate (S)-reticuline (Fig. 9). One strategy to increase production of the morphinan alkaloids is to prevent (S)-reticuline from being converted to the berberine or benzophenanthridine type alkaloids (Fig. 9). Presumably, more (S)-reticuline will then be available for conversion to alternative end products though additional metabolic engineering.

Berberine bridge enzyme (BBE) catalyzes the conversion of S-reticuline to S-scoulerine, the intermediate that leads to the berberine and benzophenanthridine alkaloids, berberine and sanguinarine respectively. Antisense RNA-mediated suppression of BBE in cell suspension cultures of California poppy, which produces benzophenanthridine alkaloids but not morphinan or berberine alkaloids, resulted in a 10-fold decrease in total alkaloid production [106]. The suppression of BBE did not result in the accumulation of the (S)-reticuline biosynthetic intermediate. Antisense suppression of BBE in hairy root cultures of California poppy gave results similar to those observed in cell suspension cultures. Overexpression of

Papaver somniferum BBE, under the control of the cauliflower mosaic virus 35S promoter, in hairy root cultures resulted in a five-fold increase in total alkaloid content [107].

In a separate study, a RNAi knockdown of BBE in California poppy cell suspension cultures resulted in the accumulation of (S)-reticuline [108], which was not accumulated in any previous RNA antisense mediated suppression of cell suspension and hairy root cultures. The reason for the differences in these results is unknown. BBE RNAi knockdown cultures accumulated a maximum of 0.031% fresh weight (S)-reticuline *versus* 0.00045% fresh weight found in control cultures. Additionally, three healthy growing BBE knockdown cell lines secreted significant amounts of (S)-reticuline into the medium (6 mg/20 mL after 2 weeks). (S)-reticuline has been shown to be a substrate for the manufacturing of various compounds that possess antimalarial and anticancer activity, so this cell line may have some commercial interest [93]. Additional products were 7-O-methylreticuline ((S)-laudanine) and 1, 2-dehydroreticuline, both minor components in wild type cultures and formed in greater amounts in BBE suppressed cultures. This study showed the accumulation of the central branch point intermediate, (S)-reticuline, and reduction of the naturally produced benzophenanthridine class of benzylisoquinoline alkaloids in California poppy. Antisense BBE was transformed into opium poppy plant and showed a similar increase in (S)-reticuline along with other alkaloids; including laudanine, laudanosine, 1, 2-dehydroreticuline, salutaridine and (S)-scoulerine and no change in the relative concentration of the morphinan alkaloids [109].

In attempts to accumulate the precursor to the analgesic oxycodone, thebaine, and decrease production of morphine (a precursor to the recreational drug heroin), RNAi mediated silencing of codeinone reductase, in opium poppy plant, was performed [93]. Codeine reductase is the penultimate step in morphine biosynthesis. Opium poppy was transformed with a chimeric cDNA hairpin RNA construct designed to silence codeinone reductase. However, silencing of codeinone reductase resulted in the accumulation of (S)-reticuline, but not the substrate codeinone or other compounds on the pathway from (S)-reticuline to codeine [93]. Allen and coworkers postulated that this accumulation could be due to several factors. First, accumulation of codeinone and morphinone could result in negative feedback on one of the enzymes, such as the reductase responsible for the

reduction of (S)-reticuline to 1,2-dehydroreticuline. Also, the biosynthetic intermediates and final product may regulate the transcription of pathway enzymes, though analysis of the transcript levels of a number of the morphine biosynthetic enzymes showed no change in suppressed plants. Finally, codeinone reductase could be a part of a multienzyme complex, which can not function when one of the enzymes is removed. This study highlights that the complex metabolic networks found in plants are not easily or predictably redirected.

Secondary metabolite profiles can also be altered by the overexpression of enzymes that are rate limiting in the synthesis of the desired product. Since the silencing of codeinone reductase leads to accumulation of (S)-reticuline, this enzyme may play a role in regulating the morphinan alkaloid biosynthetic pathway and overexpression of the reductase may lead to increased production of morphinan alkaloids. The overexpression of codeinone reductase in opium poppy plants did in fact result in an increase in morphine and other morphinan alkaloids, such as morphine, codeine, and thebaine compared to the control plants [110]. This demonstrates that increase of morphinan alkaloid levels can be achieved by increased expression of a single pathway enzyme.

In another study, a cytochrome P450 (CYP80B3) responsible for the oxidation of (S)-N-methylcoclaurine to (S)-3'-hydroxy-N-methylcocluarine (Fig. 9), was overexpressed in opium poppy plants, and morphinan alkaloid production in the latex was increased to 4.5 times the level in wild type plants [111]. Additionally, expression of antisense CYP80B3 resulted in a decrease in morphinan alkaloids to 16% of the wild type level. The increase (or decrease) in alkaloid levels was found to correlate with overexpression (or suppression) of a number of other biosynthetic enzyme transcript levels in plants overexpressing (or suppressing) CYP80B3. These results support the regulatory role CYP80B3 plays in the production of benzylisoquinoline alkaloids.

Moreover, this work highlights the complexity of the regulation of biosynthesis of secondary metabolites, where enzymes that perform the catalytic roles also have a regulatory function using mechanisms that are not yet understood. Understanding of these mechanisms will impact the future of metabolic engineering of secondary metabolism in both plants and plant cell cultures. The effective expression and suppression of pathway enzymes in hairy root cultures, cell suspension cultures and whole

plants indicates a positive future for the optimization of production of benzylisoquinoline alkaloids compounds in plant cell culture for drug manufacturing.

3.5 Artemisinin

Approximately 100 million people are infected with malaria each year, with 1 million dying from the disease [112]. The third world is particularly affected and there is a great need for inexpensive, easily administered, effective drugs. Malaria is caused by the parasite *Plasmodium*, and a number of strains resistant to older antimalarials have recently emerged [113]. Notably, used in combination with older antimalarials, artemisinin and its derivatives appear to be effective against drug resistant *Plasmodium falciparum* strains found in Southeast Asia and Africa (Fig. 10) [114]. The biological mechanism of action of this drug is not entirely clear. Although a variety of biological mechanisms have been proposed, the sarco/endoplasmic reticulum Ca^{2+} ATPase orthologue PfATP6 appears to be the likely target [115].

The extracts of *Artemisia annua* (annual or sweet wormwood), a plant that is native to China but is now endemic in many countries, were found in the 1970s to have antimalarial activity. The active agent, artemisinin (also known as qinghaosu [116] was isolated in 1972 from the aerial portions of the plant at concentrations of 0.01–0.86% [117]. The compound was structurally elucidated in 1979 and shown to be a sesquiterpene natural product containing an endoperoxide moiety (Fig. 10).

The low production levels in the plant make commercialization of artemisinin difficult. Total synthesis does not provide a commercially viable solution. Therefore, plant cell culture is being examined as a viable strategy for artemisinin production. Although production levels to date still remain low, it is hoped that optimization and genetic efforts can improve the prospects of plant cell culture as a source for artemisinin.

Efforts at establishing artemisinin producing *A. annua* cultures began in the early 1980s [118]. Although a range of artemisinin levels have been found in callus, shoot and root cultures, no artemisinin has been found in cell suspension cultures, suggesting that some degree of differentiation is required for production [114]. Transformation of *A. annua* with *Agrobacterium* rhizogenes yielded hairy root cultures that produced artemisinin at

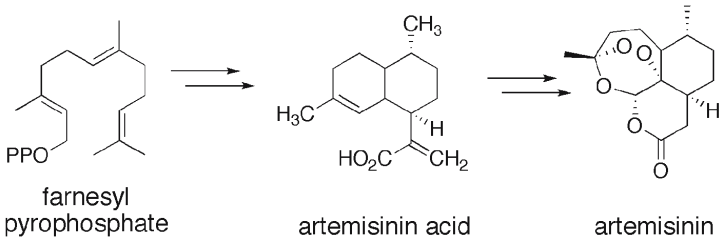


Figure 10
Artemensinin

modest levels. Currently, many research efforts are directed toward optimizing artemisinin production in hairy root cultures [116]. Bioreactors designed for the production of artemisinin from hairy root cultures have been developed, though much more optimization needs to be done before commercial levels are obtained.

The biosynthesis of artemisinin involves cyclization of farnesyl pyrophosphate to the germacrane skeleton (amorpha-4,11-diene) by a terpene cyclase. The exact order of the steps following cyclization has not been elucidated. The gene encoding the terpene cyclase has been cloned, as well as a number of genes involved in the biosynthesis of the farnesyl pyrophosphate precursor [119–122]. Since the complete biosynthetic pathway has not been elucidated, artemisinin cannot be expressed in a microbial host. However, attempts have been made to overexpress the genes in the mevalonic acid biosynthetic pathway as well as the terpene cyclase gene (amorpha-4,11-diene synthase) in *E. coli* to yield 24 µg/mL of the artemisinin precursor amorphadiene [123]. Notably, an efficient transformation system for *A. annua* is available, suggesting that as additional genes in the biosynthetic pathway are elucidated, genetic manipulation in plant cell cultures may improve production levels [124].

3.6 Shikonin

Shikonin represents one of the earliest of the plant cell culture success stories (Fig. 11). Shikonin is a naphthoquinone pigment accumulated in a variety of boraginaceous species [125]. Shikonin displays antimicrobial,

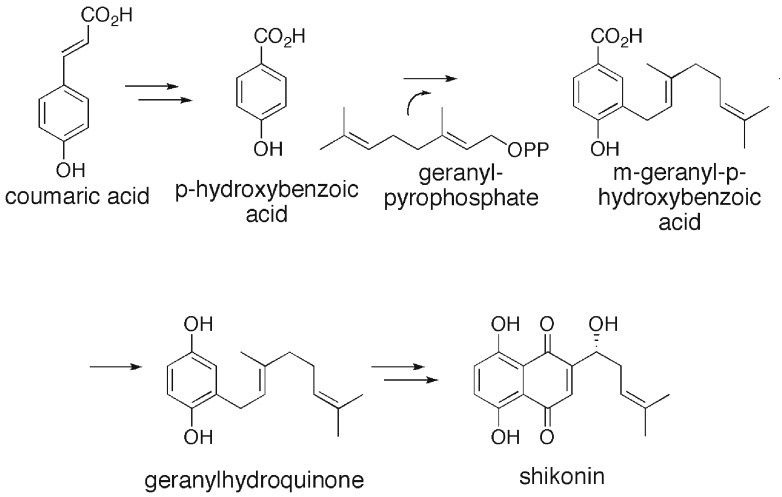


Figure 11
Shikonin

anti-inflammatory and antitumor activity. Mechanistic studies indicate that shikonin acts as a topoisomerase I inhibitor and that it also induces apoptosis in certain leukemia cell lines [126]. Shikonin appears to inhibit angiogenesis in *in vitro* studies. Shikonin is also used as a lipstick pigment in certain countries [127]. It was first demonstrated that *Lithospermum erythrorhizon* callus cultures yielded shikonin acyl esters in the 1970s [128]. High producing cell lines, producing up to 1.2 mg/g (fresh weight), could be selected for visually, since shikonin is red. Production in cell suspension cultures was subjected to extensive optimization both in terms of generation of cell biomass and product production. It was found that sucrose, oligogalacturonide and methyl jasmonate upregulated the production of shikonin [127]. A 750 L bioreactor culture yielded 1.4–2.3 g/L of shikonin in 23 days [129]. Commercial production of shikonin was reported by Mitsui Petrochemical Industry in 1983 [127].

A hairy root culture of *L. erythrorhizon* was established by infection with *Agrobacterium rhizogenes*. The hairy root culture produced shikonin in the root culture medium and also secreted it into the medium [130]. Addition of absorbents increased the concentration of shikonin produced to levels

of approximately 9 mg/day. The plant cell cultures were also used to investigate the mechanism of biosynthesis and regulation [131]. Feeding studies in the cell cultures were used to demonstrate that shikonin is derived from *p*-hydroxy-benzoic acid and geranyl diphosphate [132]. A proposed biosynthetic pathway is shown in Figure 11. However, most of the enzymes that catalyze the biosynthetic transformations remain to be cloned.

3.7 Podophyllotoxin

Podophyllotoxin, a liganin which demonstrates anti-proliferative activity [133], binds to the protein tubulin and prevents association of tubulin into microtubules (Fig. 12) [134, 135]. Unfortunately however, podophyllotoxin proved to be too toxic for clinical use. A research program at Sandoz isolated glycosylated podophyllotoxins, and discovered that these compounds although less toxic, were, unfortunately, also less cytostatic. A chemistry program was initiated to improve the anticancer activity of the lignan glycoside [136]. A variety of semi-synthetic podophyllotoxin derivatives were screened for antitumor activity and the clinical drugs etoposide, etopophos and teniposide were developed from these studies. Etoposide showed a much higher cytostatic *in vitro* activity than the precursor compound and is now licensed to and produced by Bristol-Myers Squibb (Fig. 12) [136-138]. Interestingly, etoposide was later found to inhibit topoisomerase II, and does not function by inhibiting microtubule formation, indicating that derivatization of the original pharmacophore can result in a change in biological mechanism of action.

Podophyllotoxin is produced in highest levels (4.3% DW) in *Podophyllum hexandrum* [139]. Many *Podophyllum* species are endangered, so establishment of a renewable plant cell culture source of this compound would be highly desirable. Although a cell culture that produces appropriate amounts of Podophyllum have not been established yet, efforts are being made to develop a cell culture line capable of sustaining podophyllotoxin production. However, cell cultures of *P. hexandrum* produced greatly reduced levels of the compound [140]. A variety of plant species have been screened for podophyllotoxin production and it appears that *Linum album*, *L. flavum* and *L. nodiflorum* all produce podophyllotoxin and grow well as cell suspension cultures. Cell suspension cultures of *L. album* produce

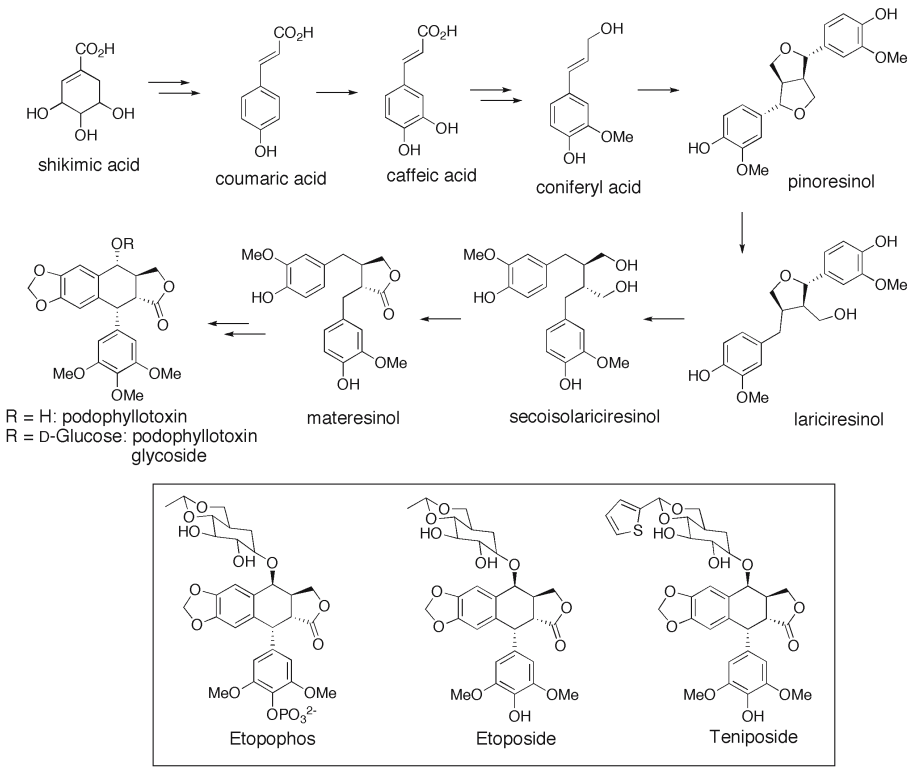


Figure 12
Podophyllotoxin biosynthesis and chemical structures of derivatives

0.2–0.5% DW podophyllotoxin, which, although an improvement over earlier cell lines, still does not approach the levels observed in the whole plant [141]. Root cultures of *L. flavum* have been shown to produce up to 5% DW podophyllotoxin, suggesting that root cultures may provide the most viable production strategy for this compound in culture [139].

4 Future directions

Plant derived natural products play a critical role in modern medicine. However, harvesting these products from the plant is expensive, requires

large amounts of plant material and is often damaging to the environment. Due to the structural complexity of natural products, total synthesis does not present an economical option for industrial production for many of these metabolites. Plant cell cultures have been established for a wide variety of species using nominal amounts of plant material, a particularly useful strategy for the production of metabolites from endangered or rare plants. Although plant cell culture represents an important alternative for natural product production, only a small fraction of plant natural products can be expressed at industrially useful levels in culture. However, as biosynthetic genes are cloned, the opportunities to metabolically engineer plant cells become more feasible, and the commercial prospects of plant cell cultures may improve.

References

1 Verpoorte R, van der Heijden R, Ten Hoopen HJG, Memelink J (1999) Metabolic engineering of plant secondary metabolite pathways for the production of fine chemicals. *Biotech Lett* 21: 467–479

2 Verpoorte R, van der Heijden R, Schripsema J, Hoge JHC, Ten Hoopen HJG (1993) Plant cell biotechnology for the production of alkaloids: present status and prospects. *J Nat Prod* 56: 186–207

3 Rates S (2001) Plants as sources of drugs. *Toxicon* 39: 603–613

4 Julsing MK, Quax WJ, Kayser O (2007) The engineering of medicinal plants: prospects and limitations of medicinal plant biotechnology. In: O Kayser, W Quax (eds): *Medicinal plant biotechnology*. Wiley-VCH, Weinheim, 3–8

5 White PR (1934) Potentially unlimited growth of excised tomato root tips in a liquid medium. *Plant Physiol* 9: 585–600

6 Gautheret RJ (1934) Culture du tissues cambial. *C R Acad Sci* 198: 151–158

7 Heinstein PF (1985) Future approaches to the formation of secondary natural products in plant cell suspension cultures. *J Nat Prod* 48: 1–9

8 Hughes EH, Shanks JV (2002) Metabolic engineering of plants for alkaloid production. *Met Eng* 4: 41–48

9 Vazquez-Flota F, de Luca V, Carrillo-Pech M, Canto-Flick A, Miranda-Ham M (2002) Vindoline biosynthesis is transcriptionally blocked in *Catharanthus roseus* cell suspension cultures. *Mol Biotechnol* 22: 1–8

10 Briskin DP, Kobayashi H, Lila MA, Gawienowski M (2004) Growth in tissue culture and *in vitro* production of kavapyrones. In: LA Dyer, ADN Palmer (eds): *Piper. A model genus for studies of phytochemistry, ecology and evolution*. Kluwer Academic Press, New York, 140–155

11 Fischer R, Stoger E, Schillberg S, Christou P, Twyman, RM (2004) Plant-based production of biopharmaceuticals. *Curr Opin Plant Biol* 7: 152–158

12 Vanisree M, Lee CY, Lo SF, Nalawade MS, Lin CY, Tsay TS (2004) Studies on the produc-

tion of some important secondary metabolites from plants by plant tissue cultures. *Bot Bull Acad Sing* 45: 1–22

13 Kreis W (2007) *In vitro* culturing techniques of medicinal plants. In: O Kayser, W Quax (eds): *Medicinal plant biotechnology*. Wiley-VCH, Weinheim, 159–185

14 Schlatman JE, Moreno PRH, Selles M, Vinke JL, Ten Hoopen HUG, Verpoorte R, Heijnen JJ (1995) Two stage batch production of ajmalicine by *Catharanthus roseus*: the link between growth and production stage. *Biotechnol Bioeng* 47: 53–59

15 Nettleship L, Slaytor M (1974) Adaptation of *Peganum harmala* callus to alkaloid production. *J Exp Bot* 25: 1114–1123

16 Eible R, Eibl D (2002) Bioreactors for plant cell and tissue cultures. In: KM Oksman-Caldentey, WH Barx (eds): *Plant biotechnology and transgenic plants*. Marcel Dekker, New York, 166

17 Hamill JD, Lidgett AJ (1997) Hairy root cultures: opportunities and key protocols for studies in metabolic engineering. In: PM Doran (ed): *Hairy roots: culture and applications*. Harwood Academic, Amsterdam, 1–31

18 Tepfer D (1990) Genetic transformations using *Agrobacterium rhizogenes*. *Physiol Plan*t 79: 140–146

19 Giri A, Narasu, ML (2000) Transgenic hairy roots: recent trends and applications. *Biotechnol Adv* 18: 1–22

20 Mugnier J (1988) Establishment of new axenic hairy root lines by inoculation with *Agrobacterium rhizogenes*. *Plant Cell Rep* 7: 9–12

21 Aird ELH, Hamill JD, Rhodes MJC (1988) Cytogenetic analysis of hairy root cultures from a number of plant species transformed by *Agrobacterium rhizogenes*. *Plant Cell Tissue Organ Cult* 15: 47–57

22 Doran PM (2002) Properties and applications of hairy root cultures. In: KM Oksman-Caldentey, WH Barx (eds): *Plant biotechnology and transgenic plants*. Marcel Dekker, New York, 143–161

23 Weathers PJ, Cheetham RD, Follansbee E, Teoh T (1994) Artemisinin production by transformed roots of *Artemisia annua*. *Biotechnol Lett* 16: 1281–1286

24 Wallaart TE, Pras N, Quax WJ, (1999) Isolation and identification of dihydroartemisinic acid hydroperoxide from *Artemisia annua*: a novel biosynthetic precursor of artemisinin. *J Nat Prod* 62: 1160–1162

25 Saito K, Mizukami H (2002) Plant cell cultures as producers of secondary compounds In: KM Oksman-Caldentey, WH Barx (eds): *Plant biotechnology and transgenic plants*. Marcel Dekker, New York, 77–109

26 Kreis, W (2007) *In vitro* culturing techniques of medicinal plants In: O Kayser, W Quax (eds): *Medicinal plant biotechnology*. Wiley-VCH, Weinheim, 173

27 Tulecke W, Nickell LG (1959) Production of large amounts of plant tissue by submerged culture. *Science* 130: 863–864

28 Noguchi M, Matsumoto T, Hirata Y, Yamamoto K, Katsuyama A, Kato A, Azechi S, Kato K (1977) Improvement pf growth rates of plant cell cultures. In: W Barz, E Reinhard, MH Zenk (eds): *Plant tissue culture and its biotechnological application*. Springer Verlag, Berlin, 85–94

29 Briskin DP (2007) Biotechnological methods for the selection of high-yielding cell lines and production of secondary metabolites in medicinal plants In: O Kayser, W Quax (eds): *Medicinal plant biotechnology*. Wiley-VCH, Weinheim, 187–199

30 van Larebeke N, Engler G, Holsters M, van den Elsacker S, Zaenen I, Schilperoort RA,

Schell J (1974) Large plasmid in *Agrobacterium tumefaciens* essential for crown gall inducing ability. *Nature* 252: 975–976
31 Chilton MD, Drummond MH, Merlo DJ, Sciaky D, Montoya AL, Gordon MP, Nester EW (1977) Stable incorporation of DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. *Cell* 11: 263–271
32 Herrera-Estrella L, Depicker A, van Montagu M, Schell J (1983) Expression of chimaeric genes transferred into plants using a Ti-plasmid derived vector. *Nature* 303: 209–213
33 Niu X, Li X, Veronesse P, Bressan RA, Weller SC, Hasegawa PM (2000) Factors affecting *Agrobacterium tumefacians* transformation of peppermint. *Plant Cell Rep* 19: 304–310
34 Klein TM, Wolf ED, Wu R, Stanford JC (1987) High velocity microprojectiles for delivering nucleic acids into living cells. *Nature* 327: 70–73
35 Twyman RM, Christou P, Stoger E (2002) Genetic transformation of plants and their cells. In: KM Oksman-Caldentey, WH Barx (eds): *Plant biotechnology and transgenic plants*. Marcel Dekker, New York, 111–141
36 Wilmink A, Dons JJM (1993) Selective agents and marker genes for use in transformation of monocotyledonous plants. *Plant Mol Biol Rep* 11: 165–185
37 Joersbo M, Donaldson I, Kreiberg J, Petersen SG, Brunstedt J, Okkels FT (1998) Analysis of mannose selection used for transformation of sugar beet. *Mol Breeding* 4: 111–117
38 Holm Y, Hiltunen R (2002) Plant derived drugs and extracts. In: KM Oksman-Caldentey, WH Barx (eds): *Plant biotechnology and transgenic plants*. Marcel Dekker, New York, 23–44
39 Wani MC, Taylor HL, Wall ME, Coggon P, McPhail AT (1971) Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifoli*. *J Am Chem Soc* 93: 2325–2327
40 Schiff PB, Fant J, Horwitz SB (1979) Promotion of microtubule assembly *in vitro* by Taxol. *Nature* 277: 665
41 Rodi DJ, Janes RW, Sanganee HJ, Holton RA, Makowski L (1999) Screening of a library of phage displayed peptides identifies human bcl-2 as a taxol binding protein. *J Mol Biol* 285: 197–204
42 Rowinsky EK, Cazenave LA, Donehower RC (1990) Taxol: a novel investigational antimicrotuble agent. *J Natl Cancer Inst* 82: 1247–1259
43 Tabata H (2004) Paclitaxel production by plant cell culture technology. *Adv Biochem Engin Biotechnol* 7: 1–23
44 Whiterup KM, Look SA, Stasko MW, Ghiorzi TJ, Muschik GM, Cragg GM (1990) *Taxus* needles contain amounts of taxol comparable to the bark of *Taxus brevifolia*: analysis and isolation. *J Nat Prod* 53: 1249
45 Holton RA, Biediger RJ, Boatman PD (1995) Semisynthesis of taxol and taxotere. In: M Suffness (ed): *Taxol: science and applications*. CRC, Boca Raton, FL, 3–25
46 Yukimune Y, Tabata H, Higashi Y, Hara Y (1996) Methyl jasmonate induced overproduction of paclitaxel and baccatin III in *Taxus* cell suspension cultures. *Nat Biotechnol* 14: 1129–1132
47 Frense D (2007) Taxanes: perspectives for biotechnological production. *Appl Microbiol Biotechnol* 73: 1233–1240
48 Fett-Neto AG, Dicosmo F, ReynoldsWF, Sakata K (1992) Cell culture of *Taxus* as a source of the antineoplastic drug taxol and related taxanes. *Biotechnology* 10: 1572–1575
49 Jennewein S, Wildung MR, Chau M, Walker K, Croteau R (2004) Random sequencing

of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Natl Acad Sci* 101: 9149–9154

50 Nims E, Dubois CP, Robers SC, Walker EL (2006) Expression profiling of genes involved in paclitaxel biosynthesis for targeted metabolic engineering. *Met Eng* 8: 385–394

51 Withers ST, Keasling JD (2007) Biosynthesis and engineering of isoprenoid small molecules. *Appl Microbiol* Biotechnol 73: 980–990

52 Stierle A, Strobel G, Stierle D (1993) Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of pacific yew. *Science* 260: 214–216

53 Shrestha K, Strobel GA, Shrivastava SP, Gewali MB (2001) Evidence for paclitaxel from three new endophytic fungi of Himalayan yew of Nepal. *Planta Med* 67: 374–376

54 Lorence A, Nessler CL (2004) Camptothecin, over four decades of surprising findings. *Phytochemistry* 65: 2735–2749

55 Thomas CJ, Rahier NJ, Hecht SM (2004) Camptothecin: current perspectives. *Bioorg Med Chem* 12: 1585–1604

56 Li QY, Zu YG, Shi RZ, Yao LP (2006) Review camptothecin: current perspectives. *Curr Med Chem* 13(17): 2021–2039

57 Oberlies NH, Kroll DJ (2004) Camptothecin and taxol: historic achievements in natural products research. *J Nat Prod* 67: 129–135

58 Wall ME, Wani MC, Cook CE, Palmer KH, McPhail AT, Sim GA (1966) Plant antitumor agents. I. Isolation and structure of camtothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminata*. *J Am Chem Soc* 88: 3888–3890

59 Lopez-Meyer M, Nessler CL, McKnight TD (1994) Sites of accumulation of the antitumor alkaloid camptothecin in *Camptotheca acuminata*. *Planta Med* 60: 558–560

60 Wani MC, Wall ME (1969) Plant antitumor agents. II. Structure of two new alkaloids from *Camptotheca acuminata*. *J Org Chem* 34: 1364–1367

61 Govindachari TR, Viswanathan N (1972) 9-Methoxycamptothecin. New alkaloid from *Mappia foetida*. *Ind J Chem* 10: 453–454

62 Wani MC, Ronman PE, Lindley JT, Wall ME (1980) Plant antitumor agents. 18. Synthesis and biological activity of camptothecin analogs. *J Med Chem* 23: 554–560

63 Hutchinson CR (1981) Camptothecin – chemistry, biogenesis and medicinal chemistry. *Tetrahedron* 37: 1047–1065

64 Kitajima M, Yoshida S, Yamagata K, Nakamura M, Takayama H, Saito K, Seki H, Aimi N (2002) Camptothecin-related alkaloids from hairy roots of *Ophiorrhiza pumila*. *Tetrahedron* 58: 9169–9178

65 Sakato K, Tanaka H, Mukai N, Misawa M (1974) Isolation and identification of camptothecin from cells of *Camptotheca acuminata* suspension cultures. *Agric Biol Chem* 38: 217–218

66 Pan XW, Xu HH, Xin L, Gao X, Lu YT (2004) Improvement of growth and camptothecin yield by altering nitrogen source supply in cell suspension cultures of *Camptotheca acuminata*. *Biotechnol Lett* 26: 1745–1748

67 Bensaddek L, Gillet F, Saucedo JEN and Fliniaux MA (2001) The effect of nitrate and ammonium concentrations on growth and alkaloid accumulation of *Atropa belladonna* hairy roots. *J Biotechnology* 85: 35–40

68 Wiedenfeld H, Furmanowa M, Roeder E, Guzewska J, Gustowski W (1997) Camptothecin and 10-hydroxycamptothecin in callus and plantlets of *Camptotheca acuminata*. *Plant Cell Tiss Organ Cult* 49: 213–218

69 Roja G, Heble MR (1994) The quinoline alkaloids camptothecin and 9-methoxycamptothecin from tissue cultures and mature trees of *Nothapodytes foetida*. *Phytochemistry* 36: 65–66
70 Ciddi V, Shuler ML (2000) Camptothecin from callus cultures of *Nothapodytes foetida*. *Biotechnol Lett* 22: 129–132
71 Fulzele DP, Satdive RK, Pol BB (2001) Growth and production of camptothecin by cell suspension cultures of *Nothapodytes foetida*. *Planta Med* 67: 150–152
72 Lorence A, Medina-Bolivar F, Nessler CL (2004) Camptothecin and 10-hydroxycamptothecin from *Camptotheca acuminata* hairy roots. *Plant Cell Reports* 22: 437–441
73 Saito K, Sudo H, Yamazaki M, Koseki-Nakamura M, Kitajima M, Takayama H, Aimi N (2001) Feasible production of camptothecin by hairy root culture of *Ophiorrhiza pumila*. *Plant Cell Rep* 20: 267–271
74 Lee-Parsons CWT, Shuler ML (2002) The effect of ajmalicine spiking and resin addition timing on the production of indole alkaloids from *Catharanthus roseus* cell cultures. *Biotechnol Bioeng* 79: 408–415
75 Sudo H, Yamakawa T, Yamazaki M, Aimi N, Saito K (2002) Bioreactor production of camptothecin by hairy root cultures of *Ophiorrhiza pumila*. *Biotechnol Lett* 24: 359–363
76 Asano T, Watase I, Sudo H, Kitajima M, Takayama H, Aimi N, Yamazaki M, Saito K (2004) Camptothecin production by *in vitro* cultures of *Ophiorrhiza liukiuensis* and *O. kuroiwai*. *Plant Biotechnol* 21: 275–281
77 Hashimoto T, Yamada Y (1983) Scopolamine production in suspension – cultures and redifferentiated roots of *Hyoscyamus niger*. *Planta Medica* 47: 195–199
78 Knopp E, Strauss A, Wehrli W (1988) Root induction on several *Solanaceae* species by *Agrobacterium rhizogenes* and the determination of root tropane alkaloid content. *Plant Cell Rep* 7: 590–593
79 Oksman-Caldenty KM, Arroo R (2000) Regulation of tropane alkaloid metabolism in plants and plant cell cultures. In: R Verpoorte, AW Alfermann (eds): *Metabolic engineering of plant secondary metabolism*. Kluwer Academic Publishers, Netherlands, 253–281
80 Hashimoto T, Kohno J, Yamada Y (1989) 6b-Hydroxyhyoscyamine epoxidase from cultured roots of *Hyoscyamus niger. Phytochemistry* 28: 1077–1082
81 Yamada Y, Hashimoto T (1989) Substrate specificity of the hyoscyamine 6β-hydroxylase from cultured roots of *Hyoscyamus niger. Proc Japan Acad B* 65: 156–159
82 Matsuda J, Okabe S, Hashimoto T, Yamada Y (1991) Molecular cloning of hyoscyamine 6β-hydroxylase, a 2-oxoglutarate-dependent dioxygenase, from cultured roots of *Hyoscyamus niger. J Biol Chem* 266: 9460–9464
83 Yun DJ, Hashimoto T, Yamada Y (1992) Metabolic engineering of medicinal-plants – transgenic *Atropa belladonna* with an improved alkaloid composition. *Proc Natl Acad Sci* 89: 11799–11803
84 Zarate R, el Jaber-Vazdekis N, Medina B, Ravelo AG (2006) Tailoring tropane alkaloid accumulation in transgenic hairy roots of *Atropa baetica* by over-expressing the gene encoding hyoscyamine 6β-hydroxylase. *Biotechnology Lett* 28: 1271–1277
85 Zhang L, Ding RX, Chai YR, Bonfill M, Moyano E, Oksman-Caldentey KM, Xu TF, Pi Y, Wang ZN, Zhang HM et al (2004) Engineering tropane biosynthetic pathway in *Hyoscyamus niger* hairy root cultures. *Proc Natl Acad Sci* 101: 6786–6791
86 Jouhikainen K, Lindgren L, Jokelainen T, Hiltunen R, Teeri TH, Oksman-Caldentey KM (1999) Enhancement of scopolamine production in *Hyoscyamus muticus* L. hairy root cultures by genetic engineering. *Planta* 208: 545–551

87 Kang YM, Lee OS, Jung HY, Kang SM, Lee BH, Karigar C, Prasad T, Bahk JD, Choi MS (2005) Overexpression of hyoscyamine 6β-hydroxylase (h6h) gene and enhanced production of tropane alkaloids in *Scopolia parviflora* hairy root lines. *J Microbiol Biotechnol* 15: 91–98

88 Palazon J, Moyano E, Cusido RM, Bonfill M, Oksman-Caldentey KM, Pinol MT (2003) Alkaloid production in *Duboisia* hybrid hairy roots and plants overexpressing the h6h gene. *Plant Science* 165: 1289–1295

89 Lee OS, Kang YM, Jung HY, Min JY, Kang SM, Karigar CS, Prasad DT, Bahk JD, Choi MS (2005) Enhanced production of tropane alkaloids in *Scopolia parviflora* by introducing the PMT (putrescine N-methyltransferase) gene. *In Vitro Cellular Dev Biol Plant* 41: 167–172

90 Moyano E, Jouhikainen K, Tammela P, Palazon J, Cusido RM, Pinol MT, Teeri TH, Oksman-Caldentey KM (2003) Effect of *pmt* gene overexpression on tropane alkaloid production in transformed root cultures of *Datura metel* and *Hyoscyamus muticus*. *J Exp Bot* 54: 203–211

91 Rothe G, Hachiya A, Yamada Y, Hashimoto T, Drager B (2003) Alkaloids in plants and root cultures of *Atropa belladonna* overexpressing putrescine N-methyltransferase. *J Exp Bot* 54: 2065–2070

92 Moyano E, Fornale S, Palazon J, Cusido RM, Bagni N, Pinol MT (2002) Alkaloid production in *Duboisia* hybrid hairy root cultures overexpressing the *pmt* gene. *Phytochemistry* 59: 697–702

93 Allen RS, Millgate AG, Chitty JA, Thisleton J, Miller JAC, Fist AJ, Gerlach WL, Larkin PJ (2004) RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nat Biotech* 22: 1559–1566

94 Facchini PJ (2001) Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Ann Rev Plant Physiol Plant Mol Biol* 52: 29–66

95 Facchini PJ, Bird DA, Bourgault R, Hagel JM, Liscombe DK, MacLeod BP, Zulak KG (2005) Opium poppy: a model system to investigate alkaloid biosynthesis in plants. *Can J Bot* 83: 1189–1206

96 Unterlinner B, Lenz R, Kutchan TM (1999) Molecular cloning and functional expression of codeinone reductase: the penultimate enzyme in morphine biosynthesis in the opium poppy *Papaver somniferum*. *Plant J* 18: 465–475

97 Grothe T, Lenz R, Kutchan TM (2001) Molecular characterization of the salutaridinol 7-O-acetyltransferase involved in morphine biosynthesis in opium poppy *Papaver somniferum*. *J Biol Chem* 276: 30717–30723

98 Sato F, Yamada Y (1984) High berberine-producing cultures of *Coptis japonica* cells. *Phytochemistry* 23: 281–285

99 Matsubara K, Kitani S, Yoshioka T, Morimoto T, Fujita Y (1989) High-density culture of *Coptis japonica* cells increases berberine production. *J Chem Technol Biotechnol* 46: 61–69

100 Eilert U, Kurz WGW, Constabel F (1985) Stimulation of sanguinarine accumulation in *Papaver somniferum* cell-cultures by fungal elicitors. *J Plant Physiol* 119: 65–76

101 Lamboursain L, Jolicoeur M (2005) Critical influence of *Eschscholzia californica* cells nutritional state on secondary metabolite production. *Biotechnol Bioeng* 91: 827–837

102 Klvana M, Legros R, Jolicoeur M (2005) *In situ* extraction strategy affects benzophen-

anthridine alkaloid production fluxes in suspension cultures of *Eschscholtzia californica*. *Biotechnol Bioeng*ineer 89: 280–289

103 Inui T, Tamura K, Fujii N, Morishige T, Sato F (2006) Overexpression of *Coptis japonica* norcoclaurine 6-O-methyltransferase overcomes the rate-limiting step in benzylisoquinoline alkaloid biosynthesis in *E. californica*. *Plant Cell Physiol* 48: 252–262

104 Siah CL, Doran PM (1991) Enhanced codeine and morphine production in suspended *Papaver somniferum* cultures after removal of exogenous hormones. *Plant Cell Rep* 10: 349–353

105 Le Flem-Bonhomme V, Laurain-Mattar D, Fliniaux MA (2004) Hairy root induction of *Papaver somniferum* var. *album*, a difficult-to-transform plant, by *A. rhizogenes* LBA 9402. *Planta* 218: 890–893

106 Park SU, Yu M, Facchini PJ (2002) Antisense RNA-mediated suppression of benzophenanthridine alkaloid biosynthesis in transgenic cell cultures of California poppy. *Plant Physiol* 128: 696–706

107 Park SU, Yu M, Facchini PJ (2003) Modulation of berberine bridge enzyme levels in transgenic root cultures of California poppy alters the accumulation of benzophenanthridine alkaloids. *Plant Mol Biol* 51: 153–164

108 Fujii N, Inui T, Iwasa K, Morishige T, Sato F (2007) Knockdown of berberine bridge enzyme by RNAi accumulates (S)-reticuline and activates a silent pathway in California poppy cells. *Transgenic Res* 16: 363–375

109 Frick S, Chitty J, Kramell R, Schmidt J, Allen R, Larkin P, Kutchan T (2004) Transformation of opium poppy with antisense berberine bridge enzyme gene *via* somatic embryogenesis results in an altered ratio of alkaloids in latex but not in roots. *Transgenic Res* 13: 607–613

110 Larkin PJ, Miller JAC, Allen RS, Chitty JA, Gerlach WL, Frick S, Kutchan TM, Fist AJ (2007) Increasing morphinan alkaloid production by overexpressing codeinone reductase in transgenic *Papaver somniferum*. *Plant Biotechnol* J 5: 26–37

111 Frick S, Kramell R, Kutchan TM (2007) Metabolic engineering with a morphine biosynthetic P450 in opium poppy susrpasses breeding. *Met Eng* 9: 169–176

112 Zhao Y, Liu C (2007) Biological production of artemisinin for malarial therapy. In: O Kayser, W Quax (eds): *Medicinal plant biotechnology*. Wiley-VCH, Weinheim, 529–542

113 Wernsdorfer WH (1994) Epidemiology of drug resistance in malaria. *Acta Trop* 56: 143–156

114 Krishna S, Uhlemann AC, Haynes RK (2004) Artemisinins: mechanisms of action and potential for resistance. *Drug Resistance Update* 7: 233–244

115 Uhlemann AC, Cameron A, Eckstein-Ludwig U, Fischbarg J, Iserovich P, Zuniga FA, East M, Lee A, Brady L, Haynes RK et al (2005) A single amino acid can determine the sensitivity of the SERCAs to artemisinins. *Nat Struct Mol Biol* 12: 628–629

116 Weathers PJ, DeJesus-Gonzalez L, Kim YJ, Souret FF, Towler MJ (2004) Alteration of biomass and artemisinin production in *Artemisia annua* hairy roots by media sterilization method and sugars. *Plant Cell Rep* 23: 414–418

117 Abdin MZ, Israr M, Rehman RU, Jain SK (2003) Artemisinin, a novel antimalarial drug: biochemical and molecular approaches for enhanced production. *Planta Med* 69: 289–299

118 He XC, Zeng MY, Li GF, Liang Z (1983) Callus induction and regeneration of plantlets from *Artemisia annua* and changes in qinghaosu contents. *Acta Bot Sin* 25: 87–90

119 Akhila A, Thakur RS Popli SP (1987) Biosynthesis of artemisinin in *Artemisia annua*. *Phytochemistry* 16: 1927–1930

120 Mercke P, Bengtsson M, Bouwmeester HJ, Posthumus MA, Brodelius PE (2000) Molecular cloning expression and characterization of amorpha-4,11-diene synthase a key enzyme of artemisinin biosynthesis in *Artemisia annua*. *Arch Biochem Biohys* 381: 173-180

121 Liu Y, Ye HC, Wang H, Li GF (2003) Molecular cloning, *E coli* expression and genomic organization of squalene synthase gene from *Artemisia annua*. *Acta Bot Sin* 45: 608-613

122 Matsushita Y, Kang WY, Charlwood BV (1996) Cloning and analysis of a cDNA encoding farnesyl diphosphate synthase from *Artemisia annua*. *Gene* 172: 207–209

123 Martin VJJ, Pitera DJ, Withers ST, Newan JD, Kiesling JD (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotech* 21: 796–802

124 Bouwmeester HJ, Wallaart TE, Janssen MHA, Loo B, Jansen BJM, Schmidt MAPCO, Kracker JWD, Konig WA, Franssen MCR (1999) Amorpha-4,11-diene synthase catalyzes the first probable step in artemisinin biosynthesis. *Phytochemistry* 52: 843–854

125 Yazaki K, Matsuoka H, Ujihara T, Sato, F (1999) Shikonin biosynthesis in *Lithospermum erythrorhizon*. Light-induced negative regulation of secondary metabolism. *Plant Biotechnol*ogy 16: 335–342

126 Yoon Y, Kim YO, Lim NY, Jeon WK, Sung HJ (1999) Shikonin an ingredient of *Lithospermum erythrorhizon* induced apoptosis in HL60 human premyelotic leukemia cell line. *Planta Med* 65: 532–535

127 Oinuma M, Shimoyama S, Noda, Y (1997) Red colorants their production and cosmetics containing the colorants. *Jpn Kokai Tokkyo Koho* 8 pp. JP 09296125

128 Tabata M, Mizukami H, Hiraoka N, Konoshima M (1974) Pigment formation in callus cultures of *Lithospermum erythrorhizon*. *Phytochemistry* 13: 927–932

129 Fujita Y, Tabata M, Nishi A, Yamada Y (1982) New medium and production of secondary compounds with two staged culture medium. In: A Fujiwara (ed): *Plant tissue culture*. Maruzen, Tokyo, 399–400

130 Sim SJ, Chang HN (1999) Fed-batch hairy root cultures with *in situ* separation. *Biotechnol Bioprocess Eng* 4: 106–111

131 Yazaki K (2001) Root-specific production of secondary metabolites: regulation of shikonin biosynthesis by light in *Lithospermum erythrorhizon*. *Natural Medicines* 55: 49–54

132 Okamoto T, Yazaki K, Tabata M (1995) Biosynthesis of shikonin derivatives from L-phenylalanine *via* deoxyshikonin in *Lithospermum* cell cultures and cell-free extracts. *Phytochemistry* 38: 83–88

133 Hartwell JL, Schrecker AW (1951) Components of podophyllin. V. The constitution of podophyllotoxin. *J Amer Chem Soc* 73: 2909–2916

134 Jardine I (1980) Podophyllotoxins. In: JM Cassady, JD Douros (eds): *Anticancer agents based on natural product models*. Academic, New York, 319–351

135 Cortese F, Bhattacharyya B, Wolff J (1977) Podophyllotoxin as a probe for the colchicine binding site of tubulin. *J Biol Chem* 252: 1134–1140

136 Stahelin HF, von Wartburg A (1991) The chemical and biological route from podophyllotoxin glucoside to etoposide. *Cancer Res* 51: 5–15

137 Canel C, Moraes RM, Dayan FE, Ferreira D (2000) Podophyllotoxin. *Phytochemistry* 54: 115-120

138 Staehelin H, Von Wartburg A (1989) From podophyllotoxin glucoside to etoposide. *Prog Drug Res* 33: 169–266

139 Arroo RRJ, Alfermann AW, Medarde M, Peterse M, Pras N, Woolley JG (2002) Plant cell factories as a source for anti-cancer lignans. *Phytochem Rev* 1: 27–35
140 Sharma TR, Singh BM, Sharma NR, Chauhan RS (2000) Identification of high podophyllotoxin producing biotypes of *Podophyllum hexandrum* Royle from North-Western Himalaya. *J Plant Biochem Biotechnol* 9: 49–51
141 Fuss E (2003) Lignans in plant cell and organ cultures: an overview. *Phytochem Rev* 2: 307–320

Index

The PDR-Series
Progress in Drug Research

Progress in Drug Research is a prestigious book series which provides extensive expert-written reviews on highly topical areas in current pharmaceutical and pharmacological research.

Founded in 1959 by Ernst Jucker, the series moved from its initial focus on medicinal chemistry to a much wider scope. Today it encompasses all fields concerned with the development of new therapeutic drugs and the elucidation of their mechanisms of action, reflecting the increasingly complex nature of modern drug research. Invited authors present their biological, chemical, biochemical, physiological, immunological, pharmaceutical, toxicological, pharmacological and clinical expertise in carefully written reviews and provide the newcomer and the specialist alike with an up-to-date list of prime references.

Starting with volume 61, *Progress in Drug Research* is continued as a series of monographs.

Forthcoming titles:

Natural Compounds as Drugs, Volume II (PDR 66), F. Petersen, R. Amstutz (Editors), 2008

Available volumes:

Systems Biological Approaches in Infectious Diseases (PDR 64), H. I. Boshoff, C. E. Barry III (Editors), 2007
Advances in Targeted Cancer Therapy (PDR 63), R. M. Schultz (Editor), 2005
Imaging in Drug Discovery and Early Clinical Trials (PDR 62), M. Rudin (Editor), 2005
Peptide Transport and Delivery into the Central Nervous System (PDR 61), L. Prokai, K. Prokai-Tatrai (Editors), 2003
Progress in Drug Research, volumes 1–60, E. Jucker (Editor), 1959–2003