Natural Products Research: Quo Vadis?

Rob Verpoorte
Natural Products Laboratory
IBL, Leiden University
PO Box 9502, 2300 RA Leiden
Verpoort@CHEM.LeidenUniv.NL
What have these in common?

Cannabis
(Hemp fibers)
Natural products are everywhere

- In your cars: natural fibers used in various materials
- In your printer ink: terpenoids
- In your clothing
- In dye of your jeans
- In your medicines
- In your shoes
  - So they are at the basis of our life
Useful plants

- Ca. 30 species for our staple food
- Ca. 100 species for fruits
- Ca. 100 species for vegetables
- 40,000-70,000 medicinal plants
- Many others (fibers, paper, wood, spices, ornamental, etc.)
Changing global situation

- Growing population
- Need for increased food production
- Need for novel medicines
  - Antibiotics, antiparasitic, anticancer, antiviral
- Pollution environment
- Sustainable, renewable production

New products and concepts from nature
Natural Products Research can contribute in many ways to make this a better world!
To survive in science you must be good in selfmotivation
5 most important natural products highlights of the past 5 decades

- Please give me your candidate Mine
- NIH plant screening: taxol, camptothecine
- TCM: artemisinin
- By chance: vinblastine and vincristine
- Metabolic engineering: Golden rice
- MEP terpenoid pathway
Nobel Prize for Medicine 2015

• Importance of the diseases of the poor
• Recognition of natural products as source of new drugs via bioprospecting
• Recognition of traditional medicine as source of new drugs
• First Nobel Prize for China
• Nobel Prize for a woman
• Nobel Prize for company for developing drugs for the poor

So in fact recognition of our whole field!
The Challenges

- Translate chemistry to genes
- Elucidate biosynthetic pathways
- Metabolic engineering
- Chemistry in plant-environment interactions
- Plants and health: medicines, food
- Novel fine chemicals from plants
Bioprospecting

The systematic search for:
- organisms
- genes
- biomolecules
- other compounds
- designs

that might have a potential use.
Sources of chemodiversity
Estimated numbers of species

- Higher plants $25 \times 10^4$
- Lower plants $12 \times 10^3$
- Vertebrates $12 \times 10^3$
- Insects $30 \times 10^6$
- Algae $10 \times 10^6$
- Fungi $15 \times 10^5$
- Prokaryotes $15 \times 10^5$

Total 10-100 million
If every organism contains one unique compound there are 10-100 million natural products.

Known: 150,000-200,000

Ca. 5000 new ones found per year

Still much to discover!
Some characteristics natural products

<table>
<thead>
<tr>
<th></th>
<th>Value end product</th>
<th>Activity range</th>
<th>amounts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medicines</strong></td>
<td>High</td>
<td>nM</td>
<td>Kg -tons</td>
</tr>
<tr>
<td><strong>Cosmetics</strong></td>
<td>High</td>
<td>μM-mM</td>
<td>Kg - tons</td>
</tr>
<tr>
<td><strong>Nutraceuticals, food additives</strong></td>
<td>Intermediate - low</td>
<td>mM</td>
<td>Tons - bulk</td>
</tr>
<tr>
<td><strong>Agrochemicals</strong></td>
<td>low</td>
<td>nM- μM</td>
<td>Tons - bulk</td>
</tr>
</tbody>
</table>
Small-molecule approved drugs 1981-2010

New medicines (NCEs) 1981-2010

Drug development 2017

The good news

• About half of all novel drugs are natural products or natural products derived!

The bad news

• The number of novel drugs is decreasing dramatically!
Lead discovery

Lead optimization

Identification clinical candidate

Investigational New Drug (IND) filing

Clinical studies (phase I-III)

New Drug Application (NDA)

registration

Screening 10,000-100,000 compounds

1-2 years

1-2 years

1-3 years

3-6 years

2-3 years

1-2 years

Ca. 7% of IND pass clinical trials

2003: 21 novel drugs

Screening 10,000-100,000 compounds

1-2 years

1-2 years

1-3 years

3-6 years

2-3 years

2003: 21 novel drugs
Reductionist approach in studying (medicinal) plants

• High throughput screening: for some targets upto 100,000 samples per 24 hrs, i.e. within three days all plants species can be screened for the activity

• Bioassay guided fractionation to isolate active compound
  – Chromatographic separation
  – Measure activity with simple bioassay
  – Repeat until pure active compound
Paradigm of modern drug development:

*Single compound single target*
Lock and key model for drug development, but the door does not change!
To find a novel drug is like:

Looking for the needle in the haystack
Allen Roses, vice-president of genetics at GlaxoSmithKline:

"The vast majority of drugs - more than 90 per cent - only work in 30 or 50 per cent of the people"

JPA Ioannidis: “Why most published Research Findings are False”

(www.plosmedicine.org)

- “For many current scientific research fields, claimed research findings may often be simple accurate measures of the prevailing bias”
- “Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true”
Models can be beautiful but are not always right.

Is this an as good or an as bad solution for the energy crisis as biofuel?
Wrong model for Adenosine A2A receptor drug development in past 20 years
Is there an other way to find the needle?

We have to rethink drug development!
Drug Development

Holistic Approach

Humans

Animals

Organs

Cells

Molecules

Reductionist Approach

past

present
2 out of 50,000 plant species were selected for the preparation of curare.
Learn from nature, learn from our ancestors!

develop Novel Models
Back to 2017: Is there no other way to find the needle?

- Go back to our ancestors approach of observation based discoveries
- Natural processes involve many factors, e.g. most diseases have multifactorial causes
- Use in-vivo tests, e.g. zebra fish, *C. elegans*
- Use all the scientific tools as an extension of our senses, e.g. measure metabolite profiles, proteome and transcriptome

*Systems biology!*
Studies of traditional medicines

- evidence based use
- novel drugs

• Mode of action
• Toxicity
• Markers for activity
• Quality assurance

• Synergism, prodrugs
• Variability, contaminations
• Synergism, prodrugs
• How to define quality
What is synergism?

Two or more agents working together to produce a result not obtainable by any of the agents independently.


1+1 > 2

the basis of life
Isobologram

Synergy: IC50 values (μg/ml) of combination of ginkgolides A and B in PAF-induced *in vitro* thrombocyte aggregation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3 : 1</td>
<td>2.40</td>
<td>4.41</td>
<td>1.42</td>
</tr>
<tr>
<td>2 : 1</td>
<td>2.20</td>
<td>3.60</td>
<td>1.72</td>
</tr>
<tr>
<td>1 : 1</td>
<td>1.80</td>
<td>2.21</td>
<td>2.12</td>
</tr>
<tr>
<td>1 : 2</td>
<td>1.55</td>
<td>1.27</td>
<td>2.43</td>
</tr>
<tr>
<td>1 : 3</td>
<td>1.40</td>
<td>0.88</td>
<td>2.57</td>
</tr>
<tr>
<td>1 : 10</td>
<td>1.30</td>
<td>0.29</td>
<td>2.79</td>
</tr>
</tbody>
</table>

Synergism: How to prove?

• Loss of activity in bioassay guided fractionation and recovery of activity after pooling inactive fractions:
  – compounds still unknown.

• Isobolograms showing activities of different combinations of two compounds:
  – which compounds to test?

• Systems biology, correlating compounds with activity:
  – combine activity and metabolomics data.
Systems Biology

*not hypothesis, but observation based*

- Organism under different conditions
- Measure as many parameters as possible
  - Metabolome
  - Proteome
  - Transcriptome
  - Physiological data
- Use e.g. multivariate analysis to find any differences, correlations, etc.
- Hypothesis based on observations
- Datamining
Systems biology
Measure as many parameters as possible
All “Omics” are tools for observing living systems on all possible levels:

SYSTEMS BIOLOGY
Key technology: metabolomics, the chemical characterization of a phenotype

Aim Metabolomics: Identification and quantification of all metabolites in an organism

Life is chemistry at work!
### Comparison metabolomic tools

<table>
<thead>
<tr>
<th></th>
<th>LC-MS</th>
<th>GC-MS</th>
<th>TLC</th>
<th>MS-MS</th>
<th>NMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample prep</td>
<td>-</td>
<td>--</td>
<td>++</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Reproducible</td>
<td>--</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Absolute qnt</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+++</td>
</tr>
<tr>
<td>Relative qnt</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>Identity</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Compound No</td>
<td>++</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Sensitive</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+++</td>
<td>-</td>
</tr>
</tbody>
</table>
The best panacea?

Wine!
Wine as a complex system

• What compounds make a wine to be a good one?
• NMR-based metabolomics of 150 wines ranked for quality on scale 1-4 (4 = best)

Classification wines
taste panel – NMR data of EtOAc extracts

- Lactic acid
- Succinic acid
- Threonine
- Alanine
- Caffeic acid
- Gallic acid
- Vanillic acid
- Proline
- 2,3-Butanediol
- GABA
- Malate
- Quercetin
- Catechin
Induction antibiotic production in *Actinomyces* species (HK Kim et al.)

- Induction silenced genes by medium manipulation, or
- Induction by methyljasmonate

• Measure metabolome and biological activity

Collaboration with prof. G. Van Wezel, IBL
TLC biogram and NMR MBT 3 extracts. Control, MJ and SA treated extracts.
Elicitation *Actinomyces* strain BS10

Anthranilic acid

Actinomycin C3
**Jasmonate Induction Actinomyces species: actinomycin C3 and anthranilic acid.**

MIC antibiotics *B. subtilis* with 2.2 mM AA

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>MIC (ug/ml)</th>
<th>MIC with anthranilic acid</th>
<th>Effect on MIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streptomycin (sulfate)</td>
<td>31.2</td>
<td>5</td>
<td>6x Lower</td>
</tr>
<tr>
<td>Neomycin (sulfate)</td>
<td>62.5</td>
<td>15.6</td>
<td>5x Lower</td>
</tr>
<tr>
<td>Penicillin G</td>
<td>0.0062</td>
<td>0.0031</td>
<td>6x Lower</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>0.0153</td>
<td>0.0153</td>
<td>No effect</td>
</tr>
<tr>
<td>Vancomycin (HCl)</td>
<td>0.156</td>
<td>0.156</td>
<td>No effect</td>
</tr>
<tr>
<td>Nalidixic acid</td>
<td>3.1</td>
<td>6.2</td>
<td>2x Increase</td>
</tr>
<tr>
<td>Actinomycin C</td>
<td>0.078</td>
<td>0.039</td>
<td>2x Lower</td>
</tr>
</tbody>
</table>
My Dream!

• To find a new drug from a plant
My nightmare?

The biologically active compound I finally found, cannot be produced in sufficient amounts.

When you can read this, you have excellent eyes.
It is feasible to grow plant cells in large scale bioreactors.

Economy:
- 1500 $/kg at 0.3 g/l/14d
- 430 $/kg at 3 g/l/14 d
Successful industrial plant biotechnology processes:

- Taxol (US/Germany, Korea, Japan)
- Shikonin (Japan)
- Ginseng roots (Japan, Korea)

However, for most important pharmaceuticals production too low to compete with the plant
Catharanthus roseus (Apocynaceae) source of terpenoid indole alkaloids

- Ajmalicine
  - improving cerebral blood circulation
- Vinblastine, vincristine
  - antitumor

Metabolic engineering or synthetic biology?
Plastid shikimate pathway
TRYPTOPHAN

Cytosol Mevalonate terpenoid pathway
C_{15}, C_{30}

Plastid Deoxyxylose terpenoid pathway
C_{10}, C_{20}, C_{40}
GERANIOL

GERANIOL

TDC

STR

STR

SGD

AJMALICINE

CATHARANTHINE

VINDOLINE

VINBLASTINE
Transgenic *Catharanthus roseus* cell cultures

- Stable cell lines overexpressing *TDC*- and/or *STR*- genes
- *TDC*-overexpression only increased tryptamine
- *STR*-overexpression increased alkaloid production (200-300 mg/l)
- Overexpression ORCA regulatory genes upregulated a series of enzymes, but not alkaloid levels
Overexpression ORCA3 in Catharanthus roseus cells

Johan Memelink and co-workers
Hairy root culture of *Weigelia* expressing *TDC* and *STR*-genes

- secologanin: below detection limit
- tryptamine: 20 μg/g DW
- ajmalicine: 1.4 μg/g DW
- serpentine: 0.2 μg/g DW

*Hallard et al. 2000*
Plant products in microorganism

- Advantage easy growth
- Intermediates should be present, or otherwise feeding is required
- Product not toxic for organism
- Few plant genes known
- Short pathways
Feed tryptamine and juice of *Symphoricarpos albus* berries (contain sugar and 2% secologanin) yield 2 g/l alkaloid per 3 days. *Geerlings et al. 2001*
EU-project SMARTCELL

Biosynthesis secologanin

Integration of transcriptomics, proteomics and metabolomics data

Successful expression whole pathway in plant

Conclusions

- Metabolic engineering is feasible
- Too little knowledge about pathways
- Few genes known
- Unpredictable results of overexpression single genes
- Effect on total flux generally limited
- Time consuming elucidation pathways
- Economical constraints: bioreactor or farmer?
Artemisia annua – artemisinin
poor farmer or big pharma?
Blueprint of the cell factory has many levels, but do we know the actual process?
Metabolism has 4 dimensions: 3 of space and 1 of time

Metabolomics is at best like a picture of low resolution, but not like the high resolution film needed to see where what is happening, fluxomics!
A plant consist of many micro- if not even nano-metabolomes
Plants consist of many micro- and nano-metabolomes.

Pyrrolizidine alkaloids in *Jacobaea vulgaris* plants.

Leaf tissues:
- Epidermis
- Palisade mesophyll
- Spongy mesophyll
Phenylpropanoids

Pyrrolizidine alkaloids

Nuringtyas et al. In press
Plastid MEP pathway

Plastid Trp pathway

Intermediates X & Y

Vacuole Secologanin? Loganin?

Vacuole STR, tryptamine, secologanin, strictosidine, ajmalicine

Vascular cell

Epidermal cell

Idioblast

Simplified model intra- and intercellular transport

C. roseus alkaloid biosynthesis
Intra- and Intercellular Transport

• Diffusion
  • Concentration
  • Mass transfer aqueous-lipid, lipid-aqueous to pass membranes
  • pH gradients (vacuole pH ca. 4.5, cytosol pH ca. 6.5)

• Selective transporters through membranes (ABC transporters)
  • Excretion
  • Uptake
Alkaloid transport into vacuoles of *Catharanthus roseus*

Roytrakul and Verpoorte, Phytochem. Rev. 6(2007)383
Logistics cell factory

- Production machinery (enzymes) should be in place
- Co-factors and energy (ATP) should be available to keep the assemblage belt running
- Assemblage of the product fully depends on availability of precursors
- Precursors need to be delivered at the right moment, on the right place in sufficient quantity
- Product must be stored before future use

Organization of precursor delivery is a crucial factor in the plant cell factory
Plant cells are like a car factory, everything must be at the right moment on the right place in the right quantity.
Engineering the cell factory

Metabolic Engineering
- A few genes
- Biosynthetic genes
- Regulatory genes

Original cell or organism

Synthetic biology
- Large number of genes
- Biosynthetic genes
- Transporter genes
- Regulatory genes
- RNAi to block competitive pathways

“Minimal cell”

Technologies: molecular biology and metabolite analysis
Did you ever asked yourself any of the following questions?

- Up to 30% flavonoids in flowers?
- Why is a plant extract a viscous liquid?
- How are non-water soluble compounds like terpenoids, cellulose biosynthesized?
- How do plants survive in the desert?
- How do organisms survive at low temperatures?
- How does a dry seed gets alive?
Down to the lowest level: the cell content

$^1$H-NMR plant extract: overall picture

- Alkaloids
- Flavonoids
- Phenylpropanoids
- Phenoloics
- Tannins
- Sugars
- Amino acids
- Glycosides
- Organic acids
- Terpenoids
- Steroids
Down to the lowest level: the cell content

Our hypothesis:
Everywhere in living systems
Natural Deep Eutectic Solvents (NADES) occur and form a third liquid phase of intermediate polarity

- Ionic liquids are formed by mixing an acid and a base (e.g. choline and malic acid)
- Deep eutectic solvents are formed by mixing two solids (e.g. glucose and malic acid)

Deep eutectic solvent (DES):

Mixing two crystalline compounds results in a lowering of the melting point, in case of DES to a temperature below room temperature

- Vapor pressure virtually zero
- Low risks for explosions and non-flammability
- Friendly for environment

- Highly viscose
- Not volatile
- Recovery of compounds
Ingredients and NADES (mole/mole)
Some examples of Natural Deep Eutectic Solvents (NADES)

<table>
<thead>
<tr>
<th>Glucose-Choline chloride-Water</th>
<th>1:1:1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fructose-Choline chloride-Water</td>
<td>1:1:1</td>
</tr>
<tr>
<td>Sucrose-Choline chloride-Water</td>
<td>1:1:1</td>
</tr>
<tr>
<td>Glucose-Fructose</td>
<td>1:1</td>
</tr>
<tr>
<td>Fructose-Sucrose</td>
<td>1:1</td>
</tr>
<tr>
<td>Glucose-Sucrose</td>
<td>1:1</td>
</tr>
<tr>
<td>Sucrose-Glucose-Fructose</td>
<td>1:1:1</td>
</tr>
</tbody>
</table>

| Malic acid-Glucose          | 1:1 |
| Malic acid-Fructose         | 1:1 |
| Malic acid-Sucrose          | 1:1 |
| Citric acid-Glucose         | 2:1 |
| Citric acid-Trehalose       | 2:1 |
| Citric acid-Sucrose         | 1:1 |
| Maleic acid-Glucose         | 4:1 |
| Maleic acid-Sucrose         | 1:1 |
### Some examples of deep eutectic solvents

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose-Choline chloride-Water</td>
<td>1:1:1</td>
<td></td>
</tr>
<tr>
<td>Fructose-Choline chloride-Water</td>
<td>1:1:1</td>
<td></td>
</tr>
<tr>
<td>Sucrose-Choline chloride-Water</td>
<td>1:1:1</td>
<td></td>
</tr>
<tr>
<td>Glucose-Fructose</td>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>Fructose-Sucrose</td>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>Glucose-Sucrose</td>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>Sucrose-Glucose-Fructose</td>
<td>1:1:1</td>
<td></td>
</tr>
<tr>
<td>Malic acid-Glucose</td>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>Malic acid-Fructose</td>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>Malic acid-Sucrose</td>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>Citric acid-Glucose</td>
<td>2:1</td>
<td></td>
</tr>
<tr>
<td>Citric acid-Trehalose</td>
<td>2:1</td>
<td></td>
</tr>
<tr>
<td>Citric acid-Sucrose</td>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>Maleic acid-Glucose</td>
<td>4:1</td>
<td></td>
</tr>
<tr>
<td>Maleic acid-Sucrose</td>
<td>1:1</td>
<td></td>
</tr>
</tbody>
</table>
NADES extractions safflower

Dai et al. In press
## Solubility macromolecules (mg/ml)

<table>
<thead>
<tr>
<th>NADES</th>
<th>Starch</th>
<th>Gluten</th>
<th>DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>LGH</td>
<td>-</td>
<td>4.8</td>
<td>286.6</td>
</tr>
<tr>
<td>GCH</td>
<td>15.8</td>
<td>0.2</td>
<td>2.5</td>
</tr>
<tr>
<td>PCH</td>
<td>11.5</td>
<td>0.3</td>
<td>7.7</td>
</tr>
<tr>
<td>PMH</td>
<td></td>
<td>4.3</td>
<td>173.7</td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td>1.5</td>
<td>252.1</td>
</tr>
<tr>
<td>SoCH</td>
<td>-</td>
<td>0.03</td>
<td>2.8</td>
</tr>
</tbody>
</table>
Laccase activity in malic acid-choline Cl (1:1)
0 % water, 2: 25% water, 3: 50% water.
Ocurrence in Plants?

- Plants secrete non-volatile saps
  - To attract insects: nectar

\[ ^1 \text{H NMR Cleome hassleriana} \text{ nectar. s: sucrose, g: glucose, f: fructose} \]
Barley Seed Germination

NMR Aleurone extract: Sucrose-choline 1:1
NADES may explain

- Biosynthesis of water insoluble compounds
- High level of accumulation of poorly water soluble compounds
- How lichen can survive drought
- How cacti and resurrection plants survive
- How a seed can germinate after 30,000 years in the permafrost
- ...........

You see it, when you understand it.  
*Johan Cruijff*
NADES beginning of life?

• Self organizing structures, liquid crystals
• Different chemistry than in water
• Intermediate polarity between water and lipids
• Water miscible, but remain stable upon dehydration
• Strongly retain water
• Liquid in large temperature range, even far below 0°C
After this existential question back to basics, experiments for at home!

How do you make the best caipirinha?
How do you make the best caipirinha?

- Sugar is a solvent!
- Sequence of lime extraction is important
- First sugar or cachaca/wodka?
- We measured clear difference of the caipirinha’s metabolome as measured by NMR
- But what tastes best?

You may send me the results: verpoort@chem.leidenuniv.nl
What is the conclusion?

• Everything is connected with everything from macro- to nanoscale
• Communication on all these levels via chemistry, sound, light, …. 
• Plants are superorganisms, as they include many organisms like symbionts and endophytes 
• Natural products chemistry is the key to functional genomics and systems biology
Multidisciplinary or interdisciplinary?

• You have your own specific expertise of your discipline.
• You can team up with other disciplines to do an interdisciplinary project.
• Be an expert!
What does my crystal ball tell me?

Natural products chemistry is the key for exploring nature in a systemic way, leading to understanding and exploiting nature to our benefit.

Learn from Nature
Learn from our ancestors!
Collaboration makes the impossible, possible
Perspectives systems biology, systems chemistry and biodiscovery

• Nature has still many useful undiscovered compounds, enzymes, genes, designs
  (e.g. recently discovered: RNAi, artemisinin, taxol, MEP terpenoid pathway, thermophiles)

• Observation based approaches will be the way to discover leads for novel products
Anecdotic examples of important observations

• Discovery of penicillin
  – Growth inhibition zones with two microorganisms on one plate → Antibiotics

• Discovery of vinblastine and vincristine
  – Testing plant for antidiabetes, observing effect on leucocytes → Antitumor medicines

• Discovery Omeprazole (Losec) no 1 best sold drug worldwide for some 15 years
  – Not active but pro-drug to treat ulcers

• Viagra, a failed antihypertension drug → k€€€
Discovery is by chance

There are no navigators for research
You need to be a good observer with an open mind