NMR spectroscopy in the structure elucidation of natural products- the Hassalidins

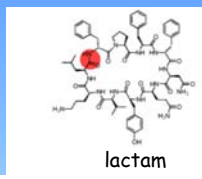
Peter Schmieder

- Growing resistance against antibiotics is an increasing risk to public health
- New compounds or new scaffolds for lead compounds are necessary
- Marine organisms and cyanobacteria are a valuable source of such new scaffolds
- A significant part of the active compounds are peptides, many are depsipeptides

The normal "rules" for peptide as we know them from peptides and proteins produced at the ribosome, however, do not apply.

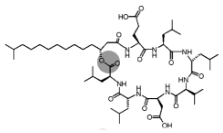
Many of the peptides are heavily modified and there is basically no restriction regarding the chemical composition: "anything goes"

Many of the peptides are macrocycles



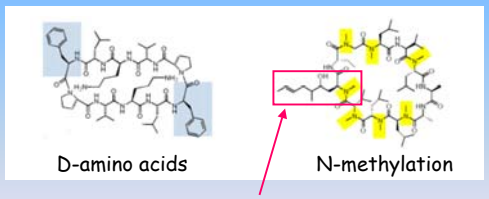
lactam

imine



lactone (depsipeptide)

The stereochemistry in the peptides is unpredictable and the amino acids can be chemically modified, N-methylation being a very common modification. Unusual amino acids can be present.

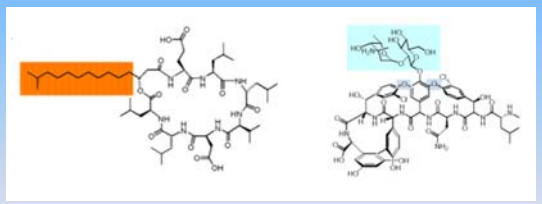


D-amino acids

N-methylation

MeBmt = (4R)-4-[(E)-2-butenyl]-4,N-dimethyl-L-threonine

In addition, lipid chains or carbohydrates may be attached to the peptide



Because of the chemical diversity the elucidation of the structure is a completely different task as compared to the structure determination of peptides and proteins with proteinogenic amino acids.
There, a signal found has to belong to the well defined set of the 20 spin systems of the amino acids that can occur in the protein.
Here, an extended set of NMR experiments has to be recoded to obtain the desired assignments.

How are these diverse structures produced by nature?

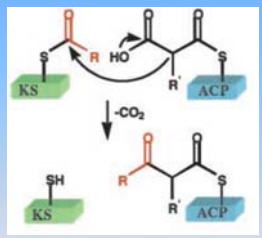
They are mainly produced at two types of of modular multienzym complexes that work in a very similar ways

Polyketidesyntase (PKS)

and

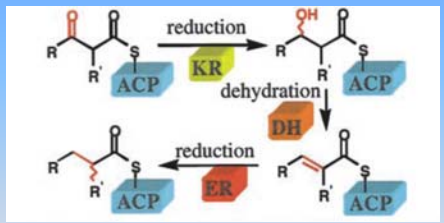
Non-ribosomal peptide synthase (NRPS)

The formation of a carbon-carbon bond in the PKS is accomplished via a Claisen-like reaction



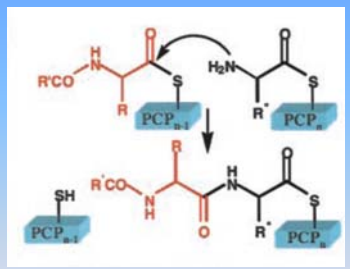
KS = Ketosynthase, ACP = Acyl carrier protein

After formation of the carbon-carbon bond various modifications can take place



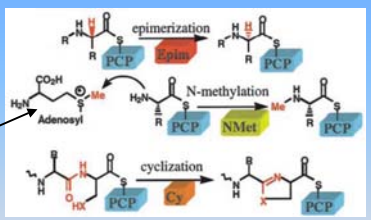
KR = Ketoreduction, DH = dehydration, ER = reduction
ACP = Acyl carrier protein

The formation of a peptide bond in the NRPS



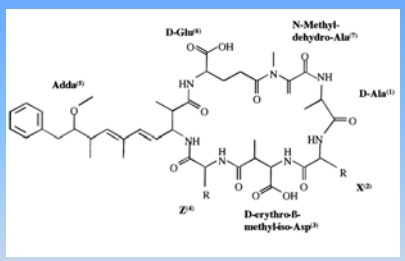
PCP = Peptide carrier protein

After formation of the peptide bond various modification can take place



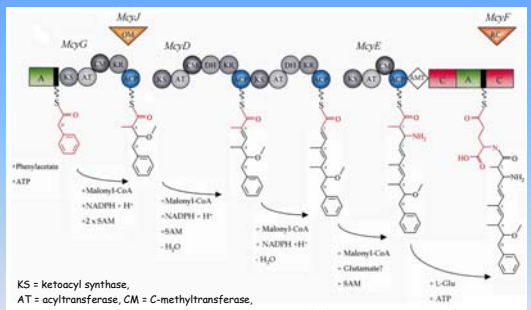
Epim = Epimerization domain, N-Met = N-methylation domain, Cy = cyclisation domain, PCP = peptide carrier protein

Synthesis of Microcystins



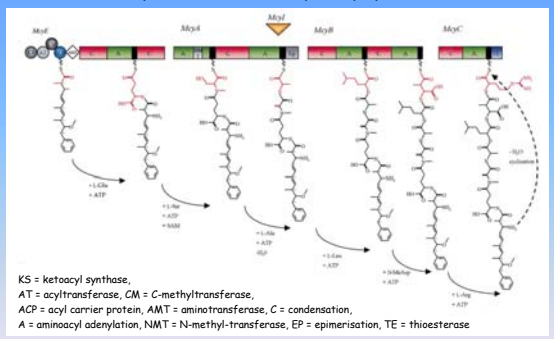
Microcystins are a familiy of bioactive peptides, Microcystin LR has $X^{(2)}$ = Leucine and $Z^{(4)}$ = Arginine

Synthesis of the unnatural amino acid Adda



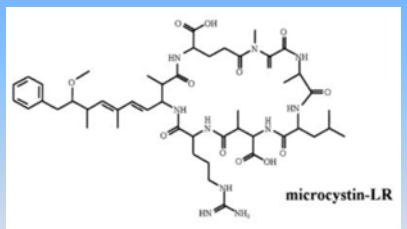
KS = ketoacyl synthase, AT = acyltransferase, CM = C-methyltransferase, OM = O-methyltransferase, KR = ketoacyl reductase, DH = dehydratase, AMT = aminotransferase, RC = racemase, ACP = acyl carrier protein, C = condensation, A = aminoacyl adenylation

Synthesis of the cyclic peptide



KS = ketoacyl synthase, AT = acyltransferase, CM = C-methyltransferase, ACP = acyl carrier protein, AMT = aminotransferase, C = condensation, A = aminoacyl adenylation, NMT = N-methyl-transferase, EP = epimerisation, TE = thioesterase

The final result is the active cyclic peptide



NMR spectroscopy in the structure elucidation of natural products

After the isolation and the detection of biological activity the novel compounds have to be characterized and their structure elucidated.

In principle, NMR-spectroscopy is ideally suited for the structure elucidation of unknown compounds since it can detect each atom separately and can detect through-bond connectivities.

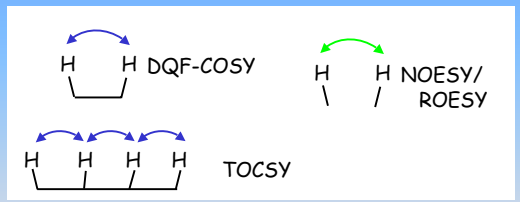
It therefore corresponds well to a chemists view of a molecule.

It is, however, quite insensitive and a fair amount of material is therefore required for a structure elucidation, i.e. > 1 mg of a 1000 Da compound.

It misses some of the atoms because of their physical properties, e.g. oxygen. Sometimes detection is difficult because of low natural abundance, e.g. nitrogen. It can not easily clarify stereochemistry.

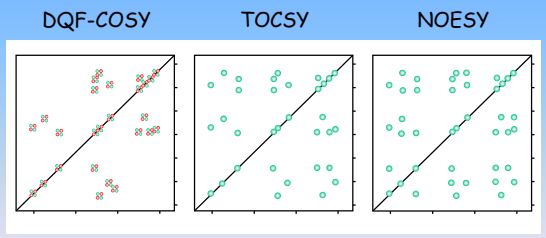
It will therefore need support from other methods, namely mass spectroscopy and chiral chromatography.

Homonuclear NMR



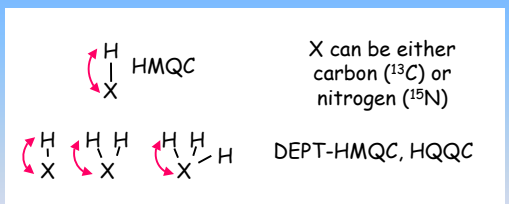
Homonuclear NMR

provides information on spin systems and thus allows to group together parts of a molecule

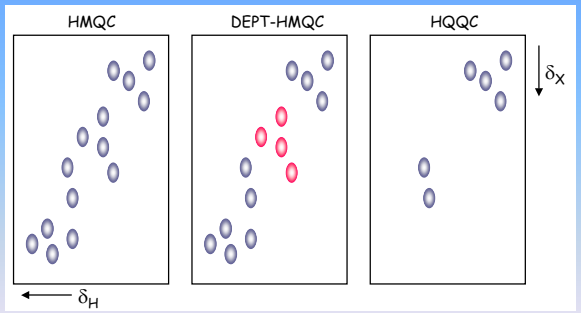


Heteronuclear NMR (1)

provides „fingerprints" of a molecule and can provide information on multiplicity

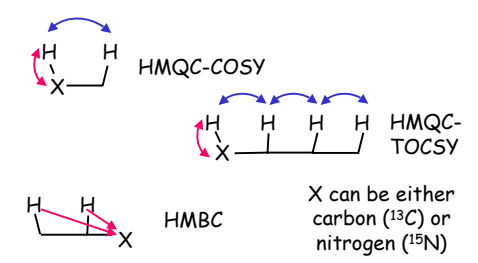


Information on the presence of CH, CH_2 and CH_3 groups can thus easily be obtained

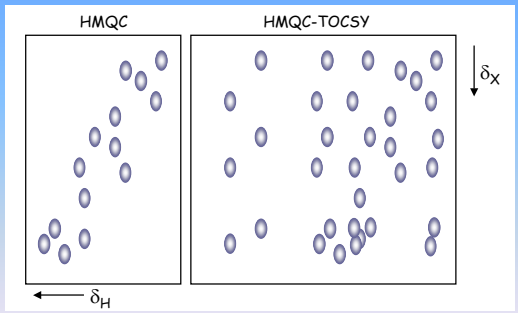


Heteronuclear NMR (2)

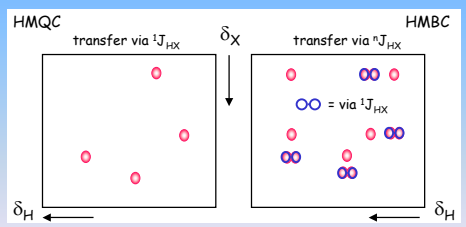
provides superior resolution due to the good dispersion and also access to heteronuclei not bound to protons



A HMQC-TOCSY combines the information on spin systems with the dispersion of the heteronuclear spectra



An HMBC provides correlations to non-proton bearing heteronuclei and thus information to connect spin systems together. It can thus also provide sequential information in peptidic compounds



In the end all spectra can be viewed as pieces of a puzzle, yielding a more or less complete picture when all pieces are assembled without contradictions.

The results should not only be consistent in terms of NMR but also in terms of mass spectroscopy. A high resolution mass can be quite helpful but with a molecular weight higher than 800 the results can be ambiguous here as well

The Hassalidins

The Hassalidins

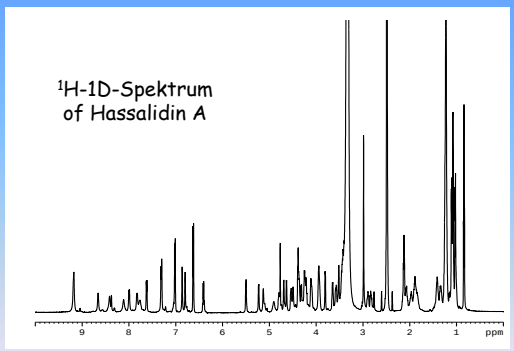
Two related depsipeptides (A and B) were isolated from a cyanobacterium *Hassallia sp.*

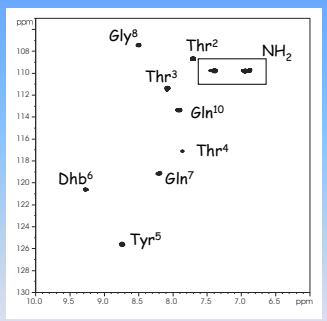
Mass spectroscopy showed that they are previously unknown compounds.

Both exhibit antifungal activity

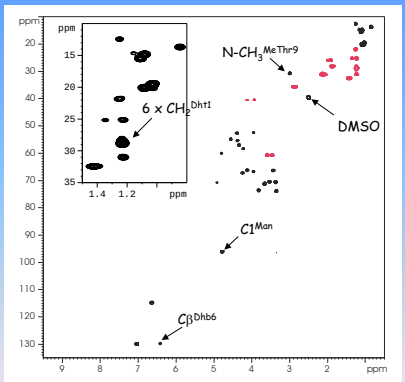
MS and chiral chromatography revealed that both are depsipeptides that contain a lipid chain and carbohydrates (mannose in Hassalidin A, mannose and rhamnose in Hassalidin B)

^{1}H-1D-Spektrum of Hassalidin A



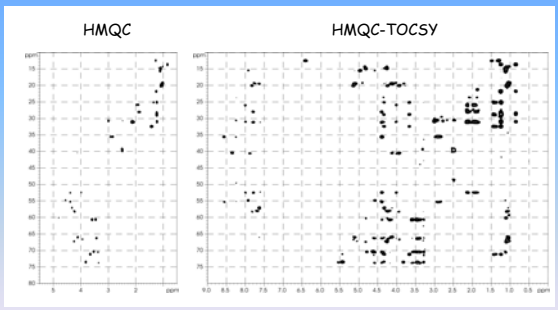


^{1}H,^{15}N-correlation as a first finger-print

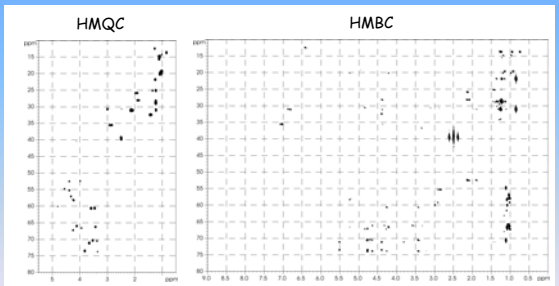


^{1}H,^{13}C-correlation as a second finger-print, including CH_2 separated from the rest

Information on the spin-systems via the HMQC-TOCSY

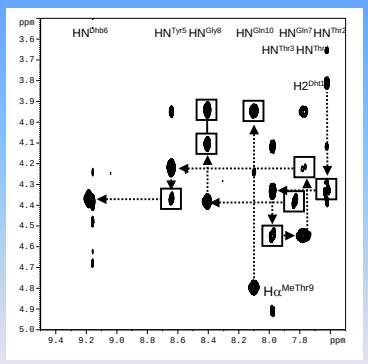


Information on the quarntenary carbons via the HMBC



After assembling the "pieces of the puzzle" NMR revealed the presence of four threonins (one N-methylated), two glutamins, a dehydroaminobutyric acid, a glycin and a tyrosin.

Mass spectroscopy confirmed the composition of amino acids.



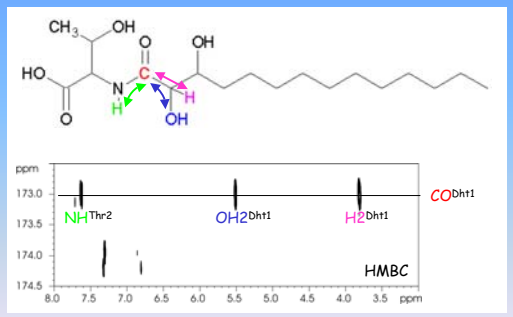
The sequence of amino acids was established using a conventional "sequential walk" The HMBC did not yield signals that were strong enough

The „core" of the molecule was thus assembled.

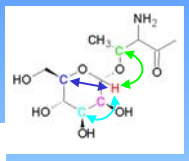
The question on the attachment of the lipid and the carbohydrate(s) to the peptide scaffold were more elaborate to answer.

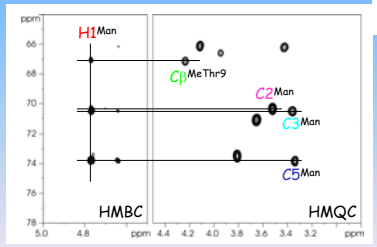
Key experiment is the HMBC, i.e. the H,C long-range correlation

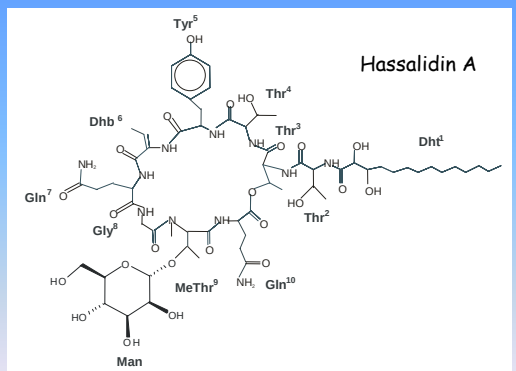
Attachment of the fatty acid to the peptide



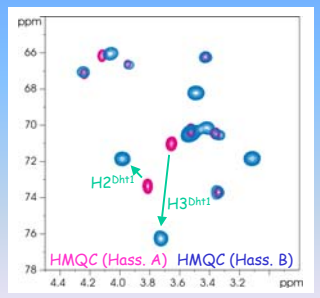
Attachment of the mannose to the peptide





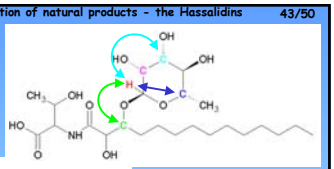


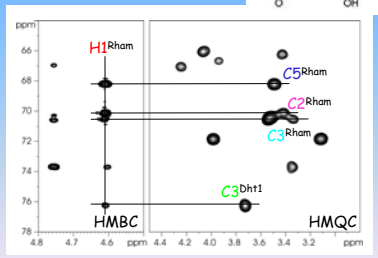
Attachment of the rhamnose to the peptide



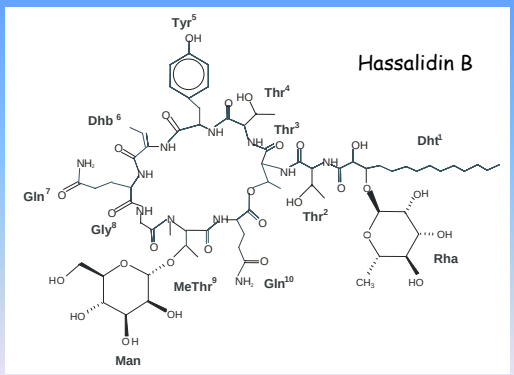
Chemical shift turns out to be the perfect indicator for structural changes

Attachment of the rhamnose to the fatty acid



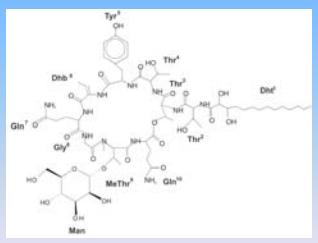


Hassalidin B



Stereochemistry

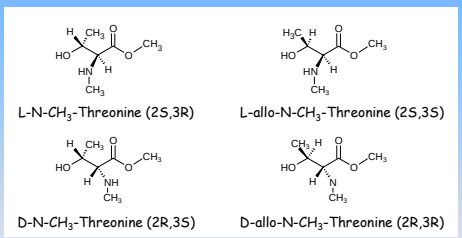
The stereochemistry of the amino acids was determined using GC-MS. This was not done in a sequence specific manner.



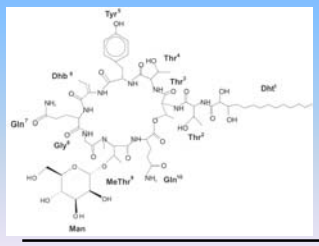
D-Thr
L-Thr
D-allo-Thr
D-Tyr
D-Gln
L-Gln
Gly
Dhb
MeThr (no reference)

Stereochemistry

The stereochemistry of the N-Methyl-Threonine will be solved by synthesis of reference material



The remaining stereochemistry will be solved by peptide synthesis of the core considering all 12 remaining possibilities. Partial hydrolysis would be tricky because of the sequence of three Thrs in a row.



D-Thr or L-Thr or D-allo-Thr (6 possibilities)

D-Gln or L-Gln (2 possibilities)

Paper:
T.Neuhof, P. Schmieder, K. Preussel, R. Dieckmann, H. Pham, F. Bartl, H. von Döhren, Hassallidin A, a glycosylated lipopeptide with antifungal activity from the cyanobacterium Hassallia sp., *J Nat Prod.* (2005) **68**, 695-700.

Patent:
T.Neuhof, R. Dieckmann, H. von Döhren, K. Preussel, M. Seibold , P. Schmieder "Lipopeptides having pharmaceutical activity"; Europäisches Patent, EP05004582, Anm.Tag 02.03.2005

Literature:

D.E. Cane, C.T. Walsh, C. Khosla, Harnessing the biosynthetic code: combinations, permutations, and mutations, *Science* (1998) **282**, 63-68.

D.E. Cane, C.T. Walsh, The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases, *Chem Biol.* (1999) **6**, R319-25.

S.A. Sieber, M.A. Marahiel, Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics, *Chem Rev.* (2005) **105**, 715-38.

D. Tillett, E. Dittmann, M. Erhard, H. von Döhren, T. Börner, B.A. Neilan, Structural organization of microcystin biosynthesis in Microcystis aeruginosa PCC7806: an integrated peptide-polyketide synthetase system, *Chem Biol.* (2000) **7**, 753-764.

acknowledgement

FMP/TU Berlin
Zerrin Fidan

TU Berlin
Torsten Neuhof
Hans von Döhren

Bruker
Wolfgang Bermel

Timo Schmiederer
Roderich Süssmuth