



Basics of Primary and Secondary Metabolites

March 27, 2017

Young Hae Choi

Natural Products Laboratory
Institute of Biology, Leiden University
Leiden, The Netherlands



Difficulty of Nat. Prod. Res.: Diversity of metabolites

- Metabolites in organisms
 - Constitutive metabolites : primary and secondary metabolites
 - Non constitutive metabolites : phytoalexin
 - Exogenous metabolites : by other organism or chemicals
- about **250,000** natural products are known
- more than **3,500** metabolites in single plant
- about **4000** new ones are reported every year
- about **15%** of all plants have been studied to some extent for their constituents



NADES Workshop

- The 1st day (Introduction)

Natural Products Research: What to do and Where to go?

Green Technology: DES, ILs and SFE

Basic Phytochemistry

- The 2nd day (Application of NADES)

Sample Preparation and Analysis

Application of NADES to Natural Products Extraction

- The 3rd day (Practice)

NADES extraction of Flavonoid from *Sophora japonica* flowers



The 1st day

- **Natural Products Research**
- **Green Technology**
- **Basic Phytochemistry**

Now, you have samples for metabolomics!

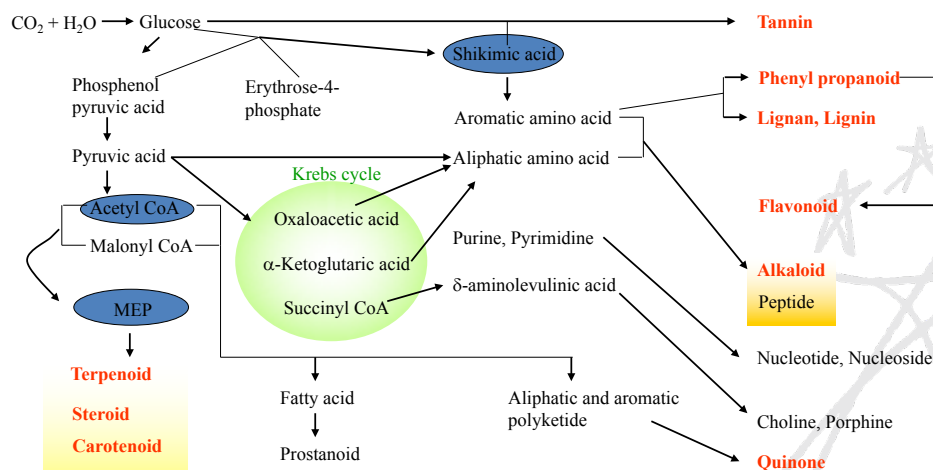


Basics of plant metabolites

- **The 1st step of research** to know about metabolites present in living organisms
- **Chemical characteristics** of group of metabolites (carbohydrates, amino acids, flavonoids, terpenoids, alkaloids...) detected by diverse analytical methods (UV, MS, NMR)
- **Common metabolites (usual suspect)** detected in plants
- **Exceptional metabolites** which show different analytical features from others

Are you afraid of huge number of metabolites?

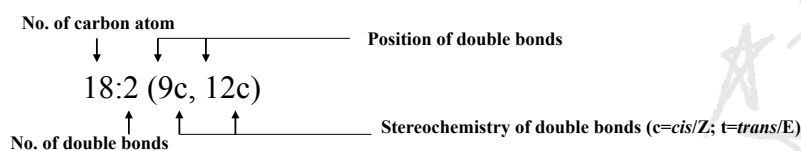
Plant Metabolites





Fatty Acid and Lipid (1)

- Simple and complex
- Simple lipid: fatty acid, glyceride (ester of fatty acid and glycerol), wax (ester of higher fatty acid and higher alcohol)
- Complexed: glycolipid (sphingo, glycerol), phospholipid (sphingo, glycerol)



Saturated

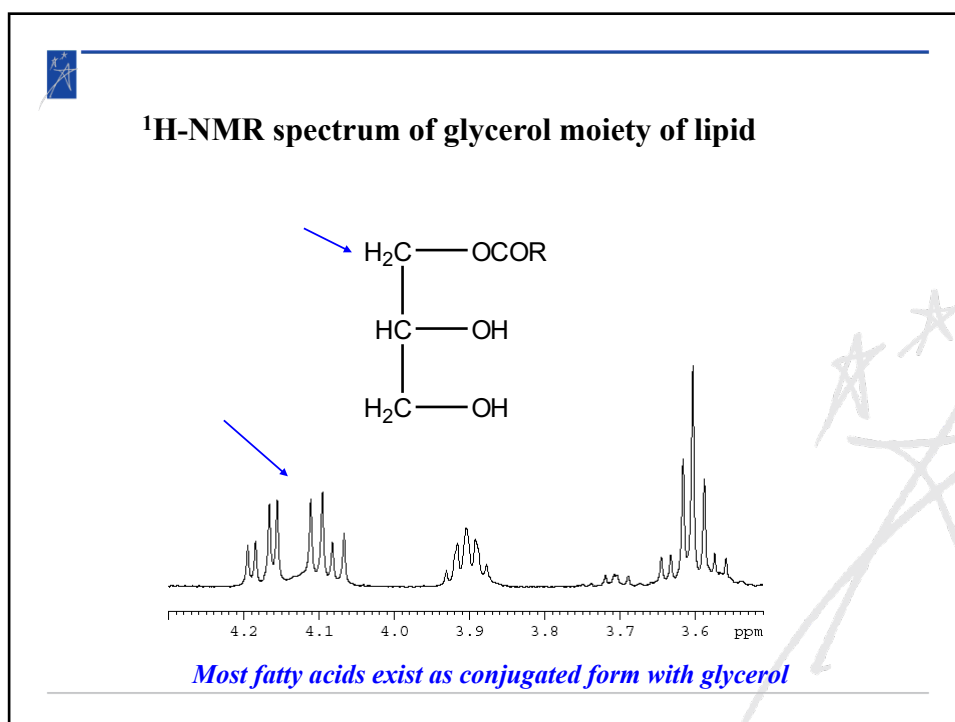
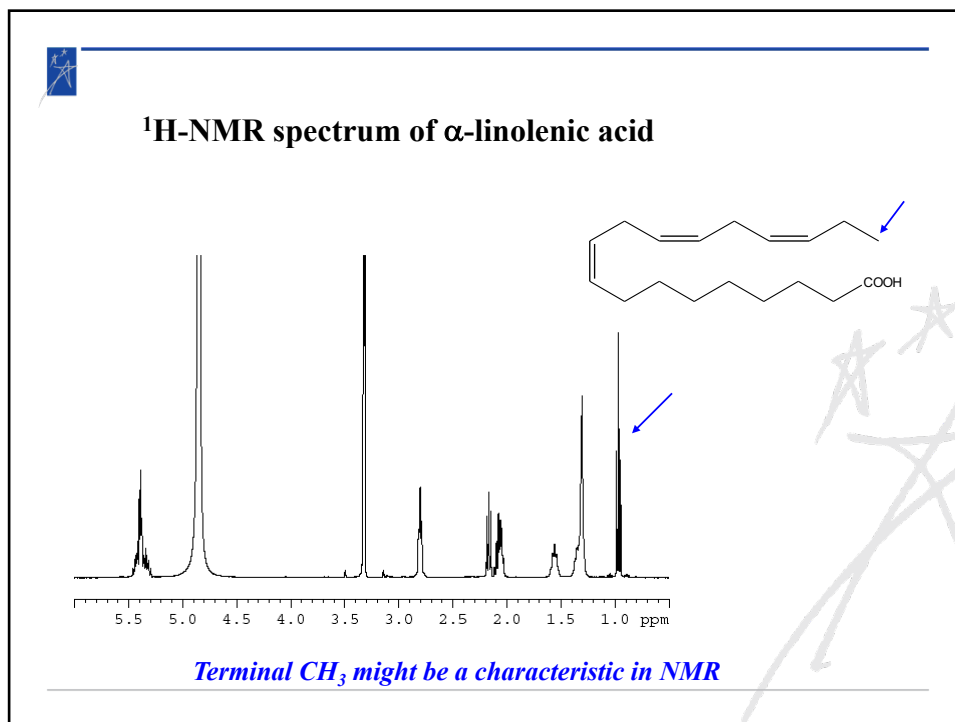
butyric	$\text{CH}_3(\text{CH}_2)_2\text{COOH}$ (4:0)	stearic	$\text{CH}_3(\text{CH}_2)_{16}\text{COOH}$ (18:0)
caproic	$\text{CH}_3(\text{CH}_2)_4\text{COOH}$ (6:0)	arachidic	$\text{CH}_3(\text{CH}_2)_{18}\text{COOH}$ (20:0)
caprylic	$\text{CH}_3(\text{CH}_2)_6\text{COOH}$ (8:0)	behenic	$\text{CH}_3(\text{CH}_2)_{20}\text{COOH}$ (22:0)
capric	$\text{CH}_3(\text{CH}_2)_8\text{COOH}$ (10:0)	lignoceric	$\text{CH}_3(\text{CH}_2)_{22}\text{COOH}$ (24:0)
lauric	$\text{CH}_3(\text{CH}_2)_{10}\text{COOH}$ (12:0)	cerotic	$\text{CH}_3(\text{CH}_2)_{24}\text{COOH}$ (26:0)
myristic	$\text{CH}_3(\text{CH}_2)_{12}\text{COOH}$ (14:0)	montanic	$\text{CH}_3(\text{CH}_2)_{26}\text{COOH}$ (28:0)
palmitic	$\text{CH}_3(\text{CH}_2)_{14}\text{COOH}$ (16:0)	melissic	$\text{CH}_3(\text{CH}_2)_{28}\text{COOH}$ (30:0)



Fatty Acid and Lipid (2)

Unsaturated

palmitoleic	$\text{CH}_3(\text{CH}_2)_5\text{CH}=\text{CH}(\text{CH}_2)_7\text{COOH}$	<ul style="list-style-type: none"> • Abundant in the seeds • Choline, α-linolenic acid • Difficult to identify each lipid by NMR → GC-MS
oleic	$\text{CH}_3(\text{CH}_2)_7\text{CH}=\text{CH}(\text{CH}_2)_7\text{COOH}$	
<i>cis</i> -vaccenic	$\text{CH}_3(\text{CH}_2)_3\text{CH}=\text{CH}(\text{CH}_2)_9\text{COOH}$	
linoleic	$\text{CH}_3(\text{CH}_2)_4\text{CH}=\text{CHCH}_2\text{CH}=\text{CH}(\text{CH}_2)_7\text{COOH}$	
α -linolenic	$\text{CH}_3\text{CH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CH}(\text{CH}_2)_7\text{COOH}$	
γ -linolenic	$\text{CH}_3(\text{CH}_2)_4\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CH}(\text{CH}_2)_4\text{COOH}$	
gadoleic	$\text{CH}_3(\text{CH}_2)_9\text{CH}=\text{CH}(\text{CH}_2)_7\text{COOH}$	
arachidonic	$\text{CH}_3(\text{CH}_2)_4\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CH}(\text{CH}_2)_3\text{COOH}$	
eicosapentaenoic (EPA)	$\text{CH}_3\text{CH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CH}(\text{CH}_2)_3\text{COOH}$	
erucic	$\text{CH}_3(\text{CH}_2)_7\text{CH}=\text{CH}(\text{CH}_2)_{11}\text{COOH}$	
docosapentaenoic (DPA)	$\text{CH}_3\text{CH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CH}(\text{CH}_2)_5\text{COOH}$	
docosahexaenoic (DHA)	$\text{CH}_3\text{CH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CH}(\text{CH}_2)_2\text{COOH}$	
nervonic	$\text{CH}_3(\text{CH}_2)_7\text{CH}=\text{CH}(\text{CH}_2)_{13}\text{COOH}$	





GC analysis of fatty acids (methylation): **lipidomics**

1. Add 1 mL of **0.5 M NaOH solution** (10 g/500 mL in MeOH) to extract
2. React for 30 min at 75 °C
3. Transfer to 10 mL test tube
4. Extract with *n*-hexane (2 mL x 3) and transfer to 25 mL evaporating flask
5. Evaporate *n*-hexane extract
6. Add 0.5 mL of **BF₃ in MeOH** and transfer to 1.5 mL-glass vial
7. React for 20 min at 75 °C
8. Add 0.5 mL of *n*-hexane
9. Injection of 1 µL of *n*-hexane solution to GC
10. GC condition is the same to nonpolar metabolites analysis

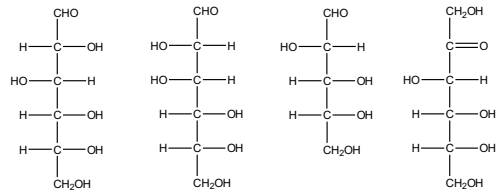


Carbohydrate (1)

- Polyhydroxy aldehyde or polyhydroxy ketone Stereoisomer having C_n(H₂O)_n
- Aldose (Aldehyde functional group), Ketose (Ketone functional group)
- Pentose: C5, Hexose: C6
- Sugars
 - Monosaccharides (aldoses and ketoses)
 - Trioses, Tetroses, Hexoses
 - Oligosaccharides
 - Disaccharides, Trisaccharides
 - Sugar derivatives
 - Alcohols, Acids, Esters, Glycosides
 - Polysaccharides (glycans)
 - Hexosans (Glucans, Fructans, Galactans, Mannans, Glucomannans, Galactomannans)
 - Pentosans (Xylans, Arabans)
 - Glucouronans (Glucouranans, Galactouronans)



Carbohydrate (2)

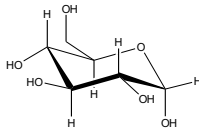
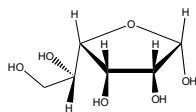
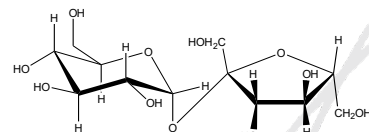


D-Glucose

D-Mannose

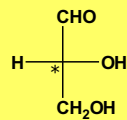
D-Arabinose

D-fructose

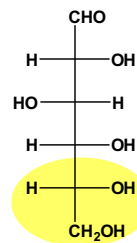
 α -D-Glucopyranose α -D-Glucofuranose β -D-Fructofuranosyl-(2 \rightarrow 1)- α -D-glucopyranoside
(Sucrose)

Carbohydrate (3)

D- and L- form: Fischer-Rosanoff's rule



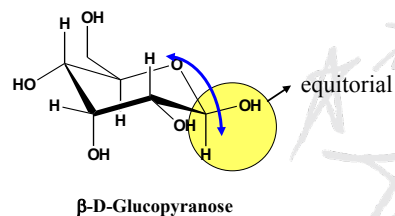
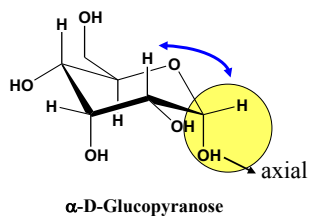
(+)D-glyceraldehyde



D-glucose : natural form

Carbohydrate (4)

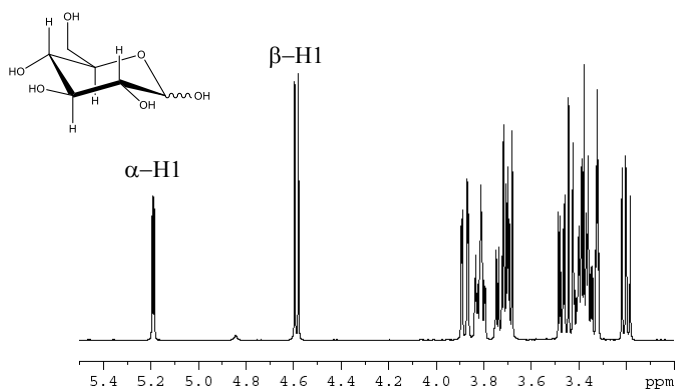
Anomeric configuration : α - and β form



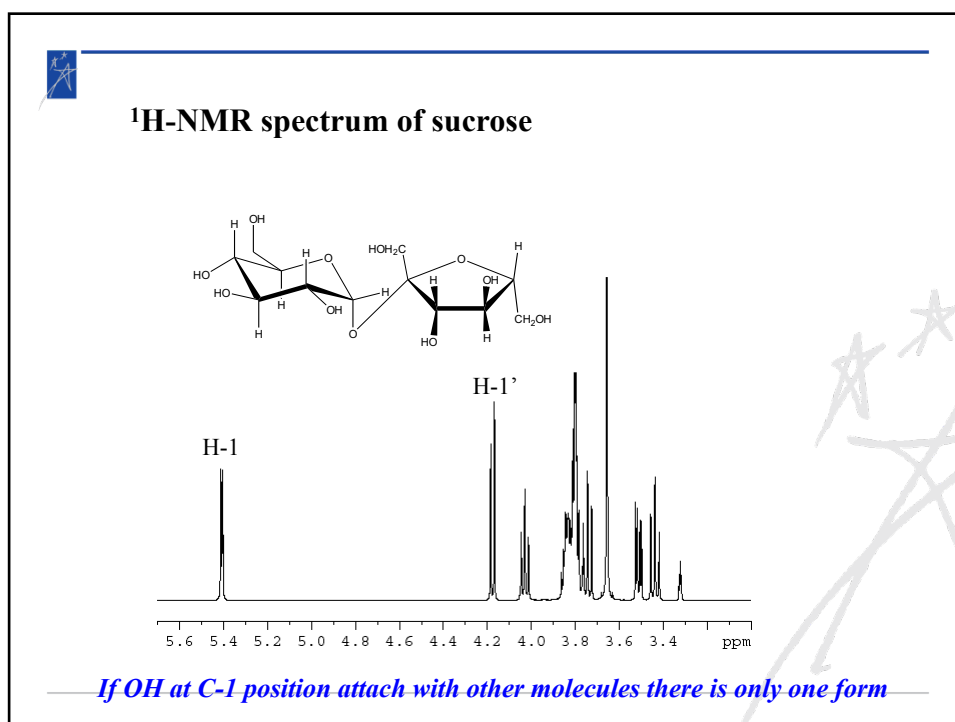
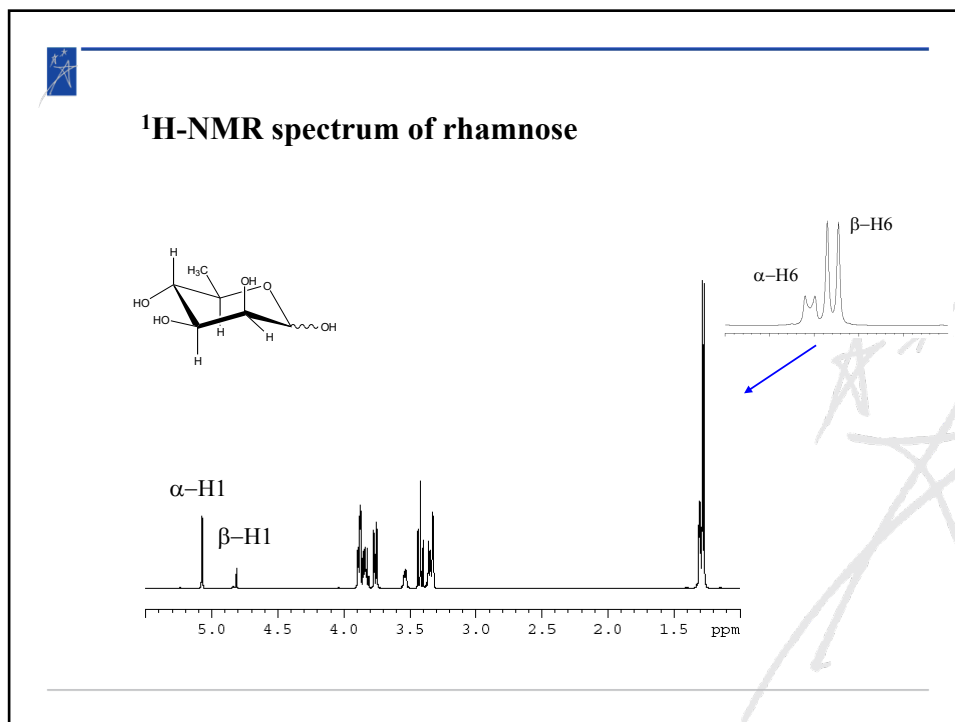
α form- 60° : $J=2-3$ Hz
 β form- 180° : $J=7-8$ Hz

Ratio of glucose (α : β)= 36:64

$^1\text{H-NMR}$ spectrum of glucose



Monosaccharide has two forms in water solution





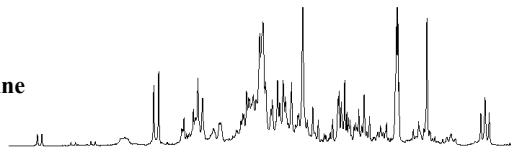
Carbohydrate (5)

- Low sensitivity in LC-MS (derivatization needed for increasing sensitivity)
- Good sensitivity and resolution in GC-MS after TMS-derivatization
- Several signals in GC-MS analysis
- Oligosaccharides make NMR signals broadening
- In NMR analysis CPMG pulse needed

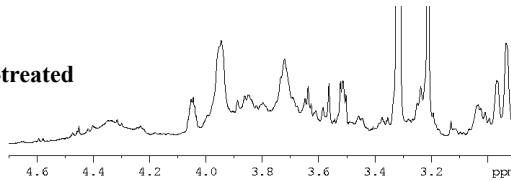


$^1\text{H-NMR}$ spectrum of cannabis cell line

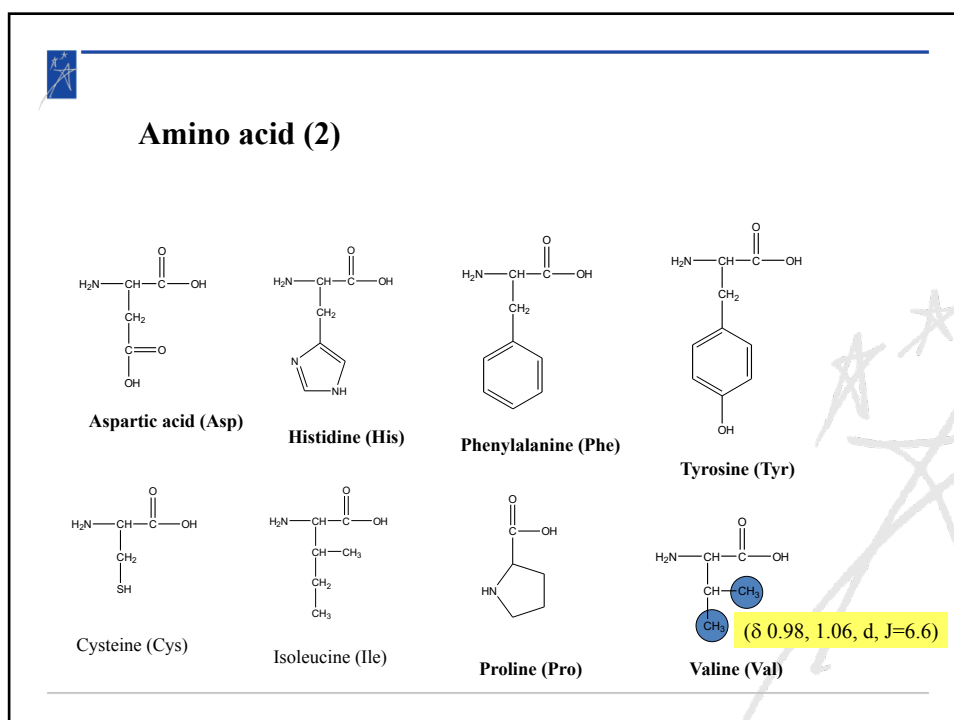
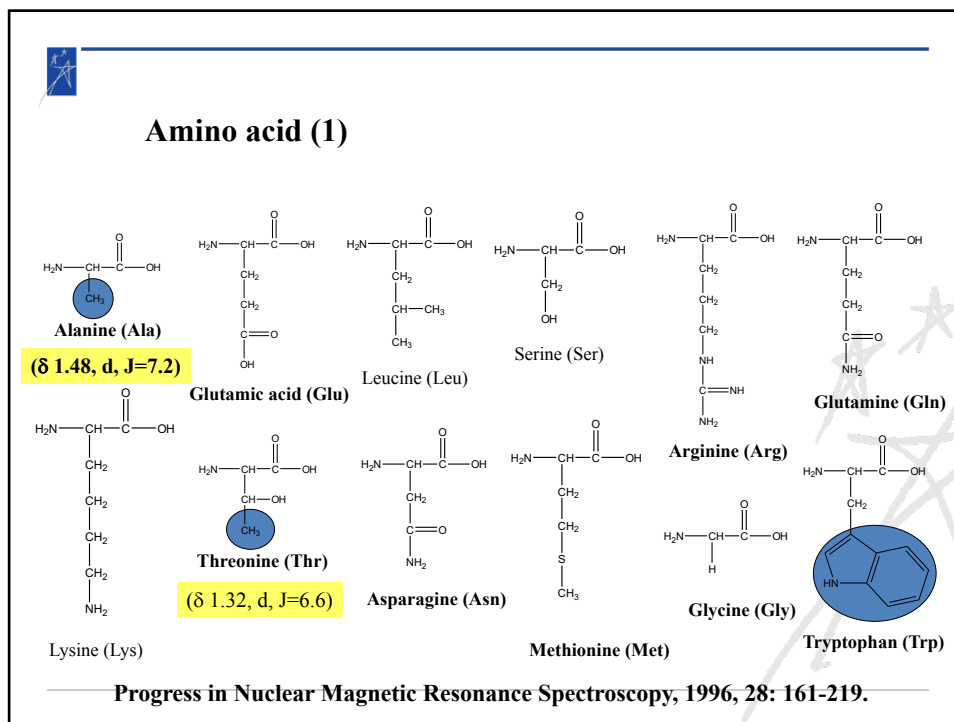
Normal cell line

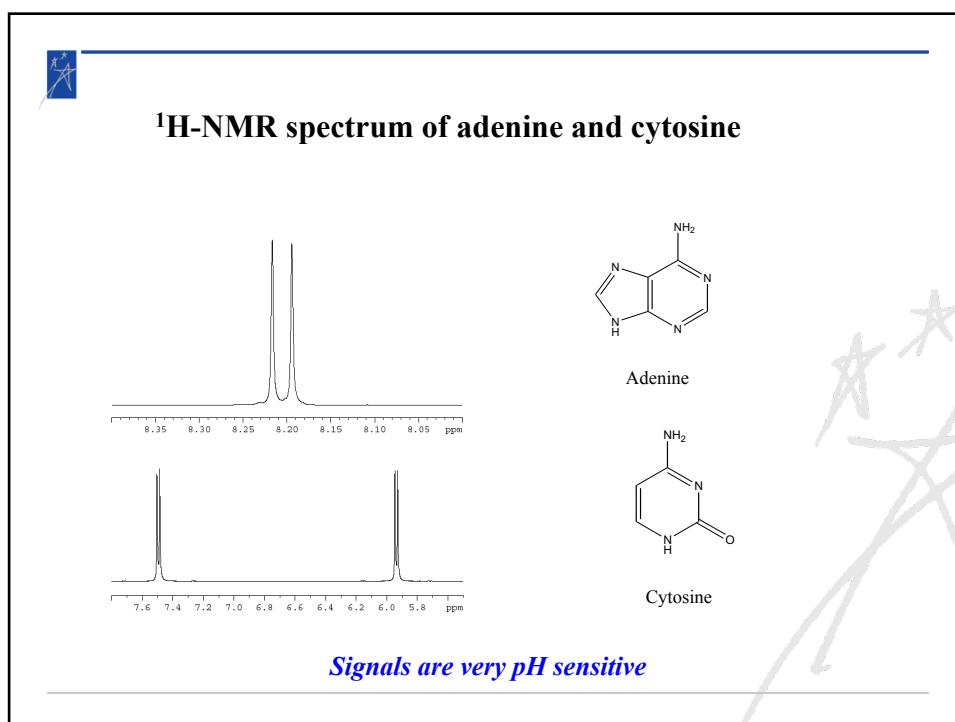
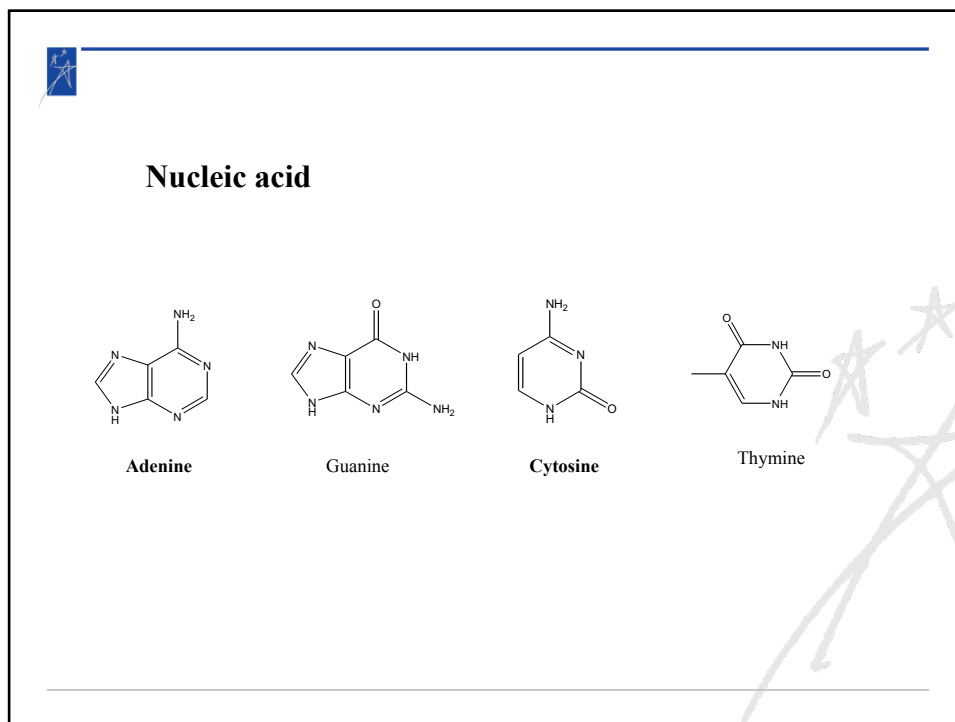


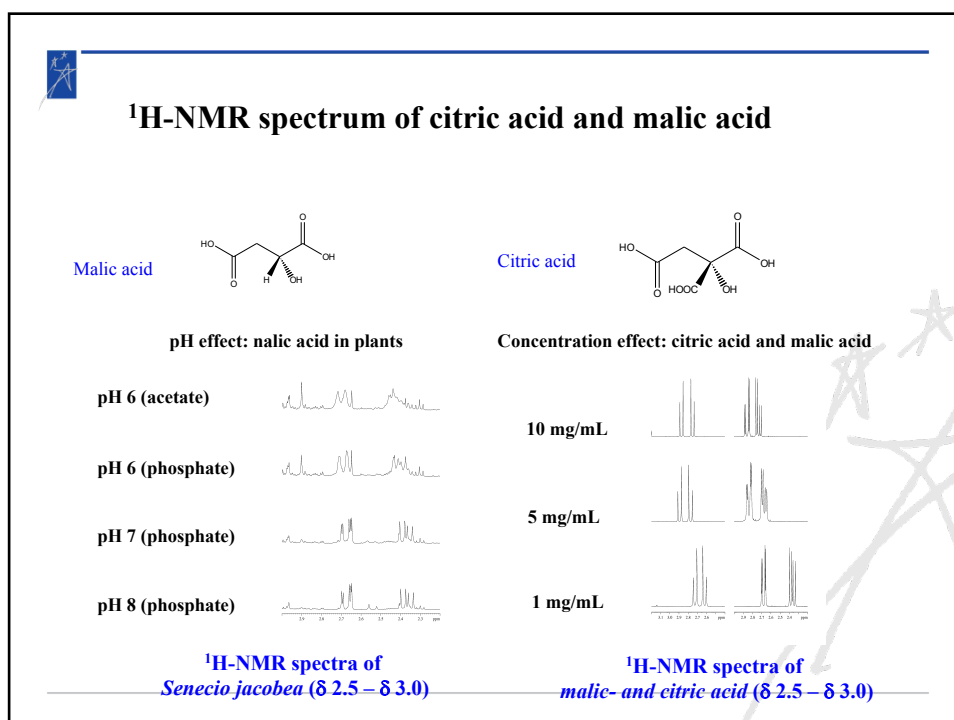
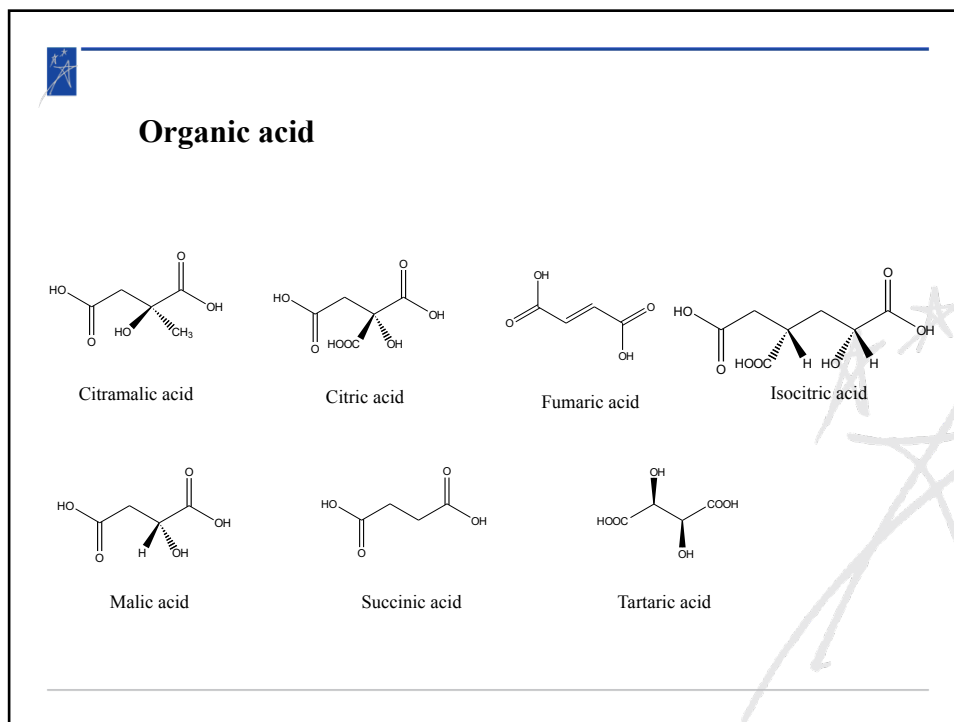
Polysaccharide-treated



Oligo- or polysaccharide-rich sample cause signal broadening

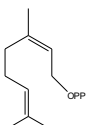




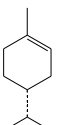


Monoterpenoid (C10)

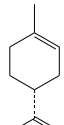
- Component of plant essential oil
- GC-MS targeted analysis



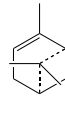
nerylphosphosphate



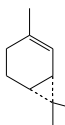
menthane



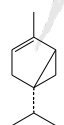
limonene



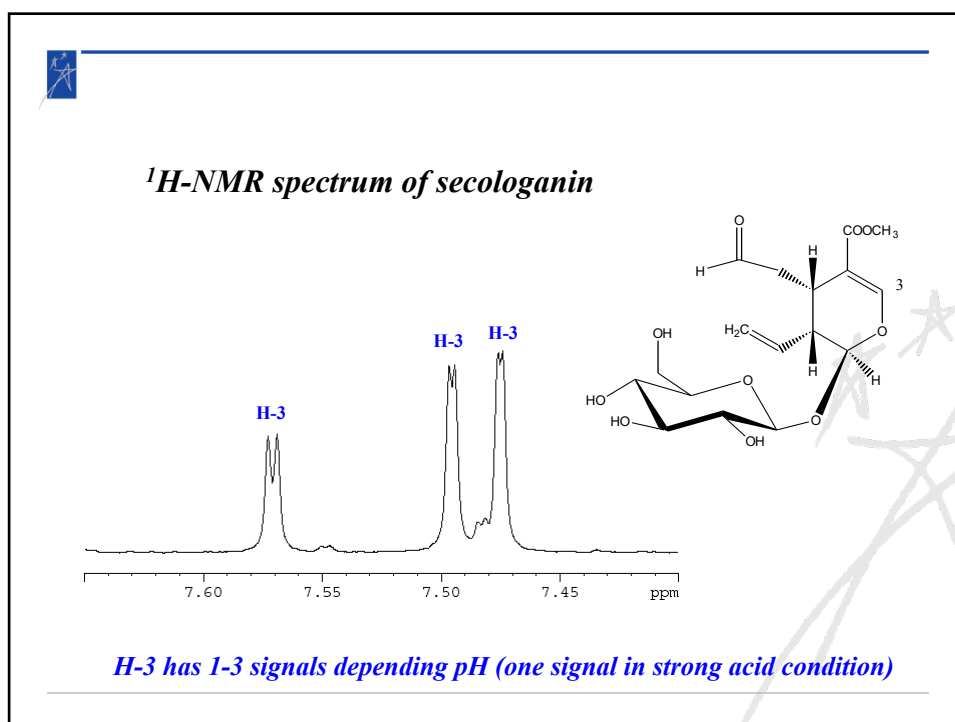
α -pinene



carnene

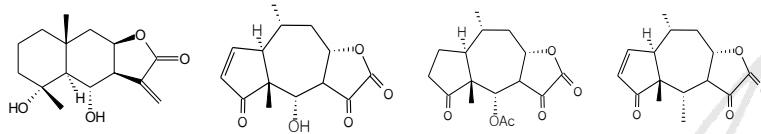


thujene

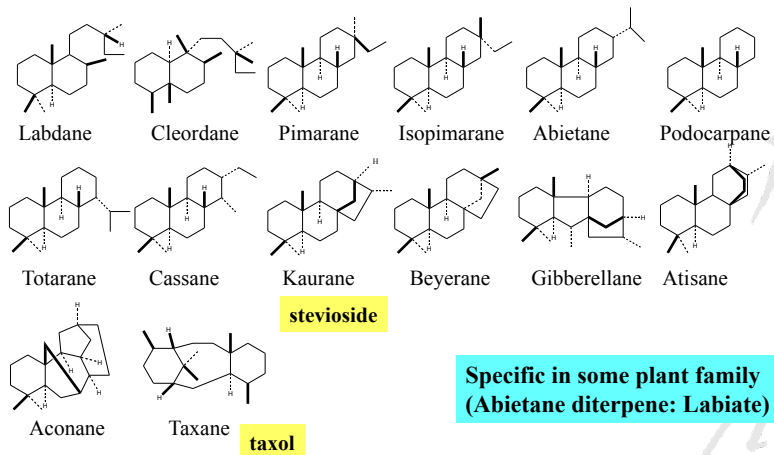


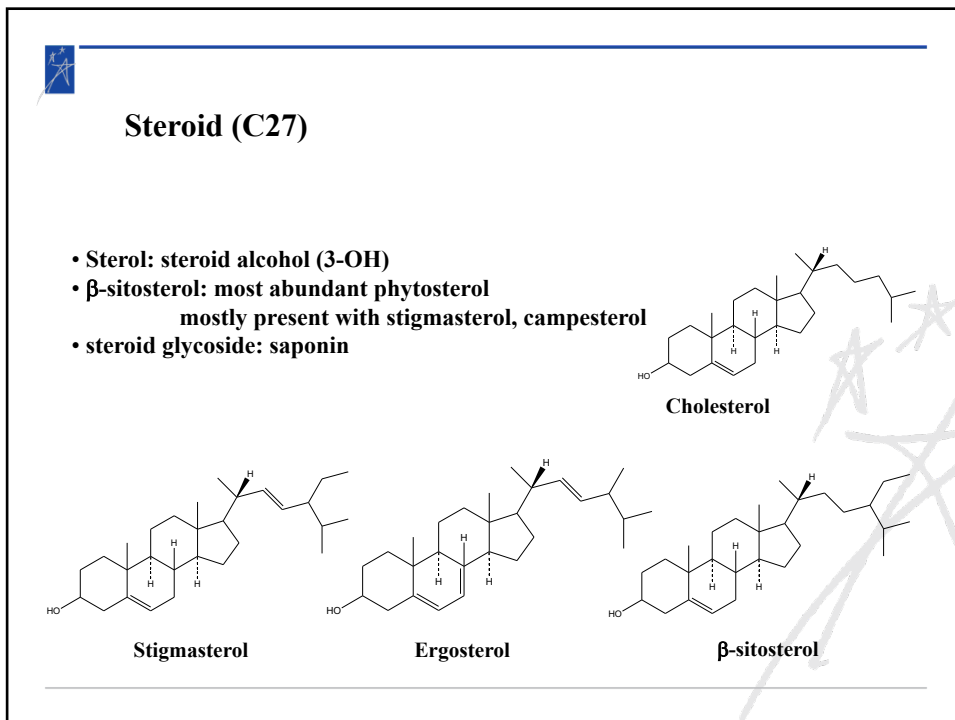
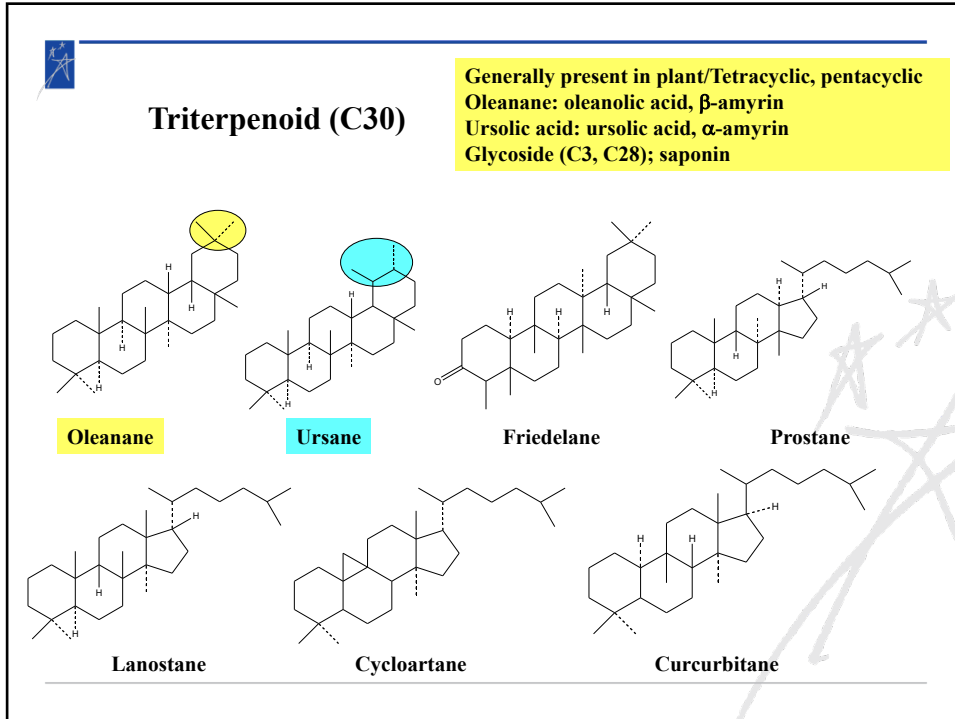
Sesquiterpenoid (C15)

- Most diverse group of compounds (more than 100 skeletons)
- Cytotoxic: sesquiterpene lactone
- Phytoalexins in Solanaceae



Diterpenoid (C20)

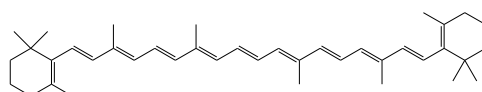






Carotenoid (C40)

- 8 isoprenoid
- color (yellow, red)
- trans (natural form)/cis (light)
- abundant in carrot

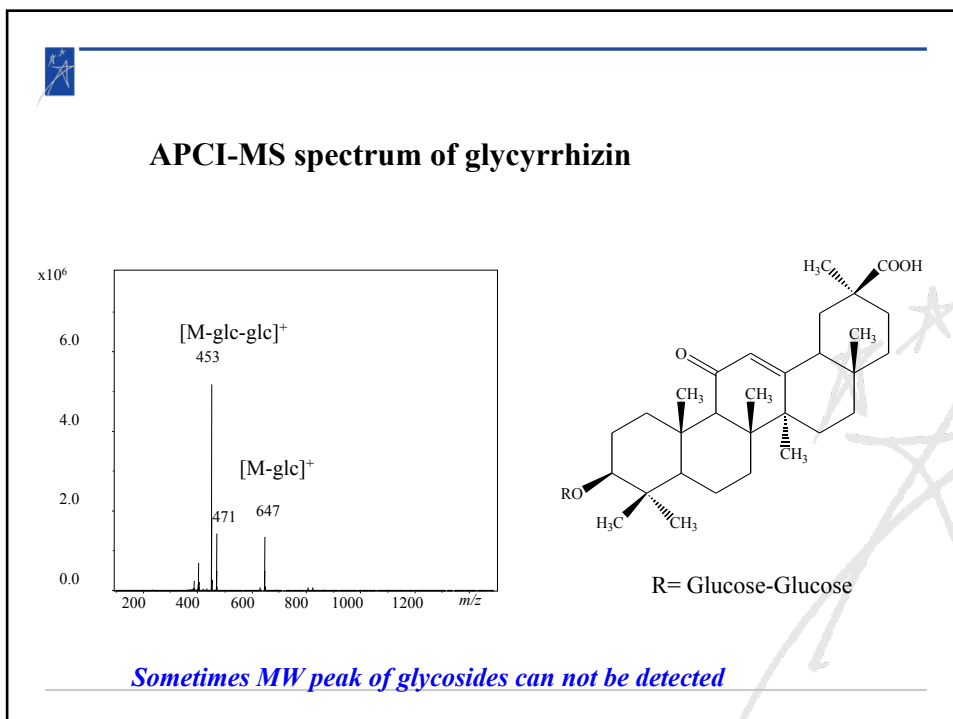
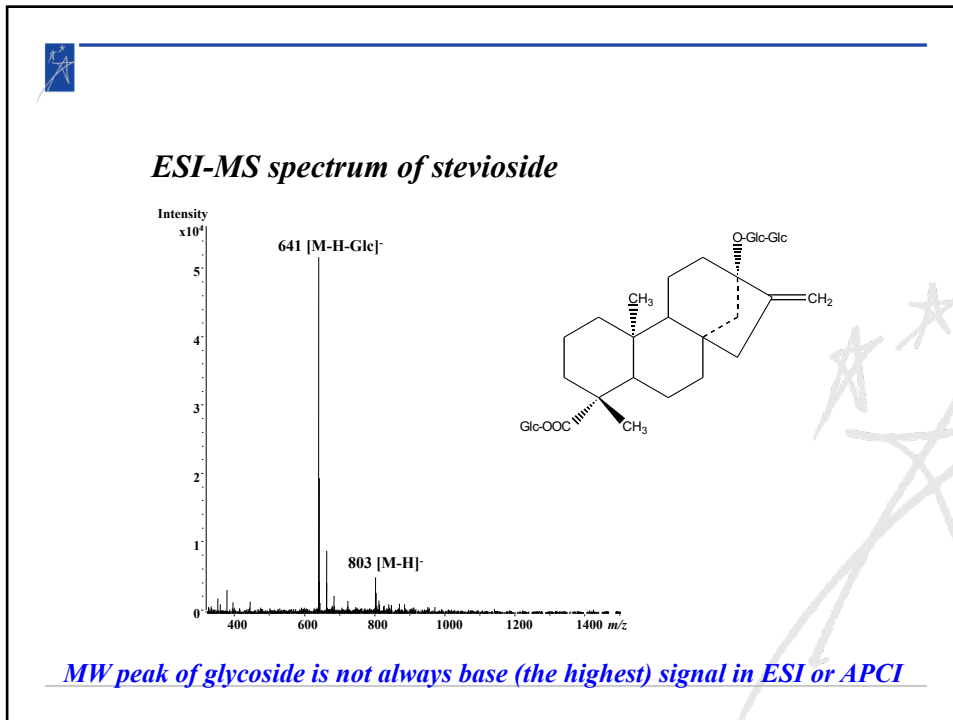


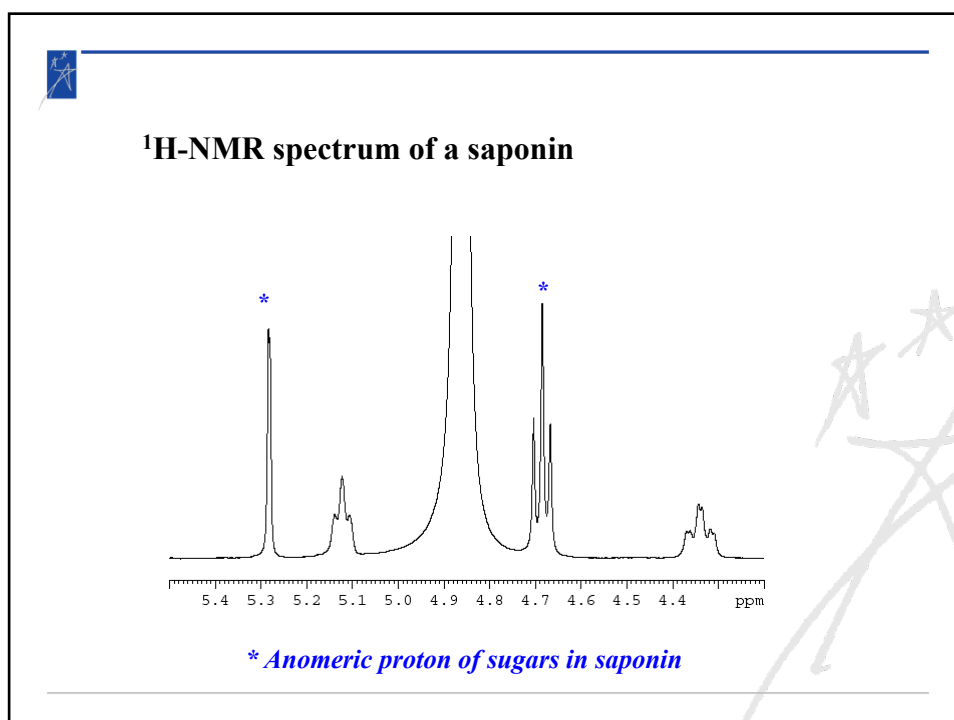
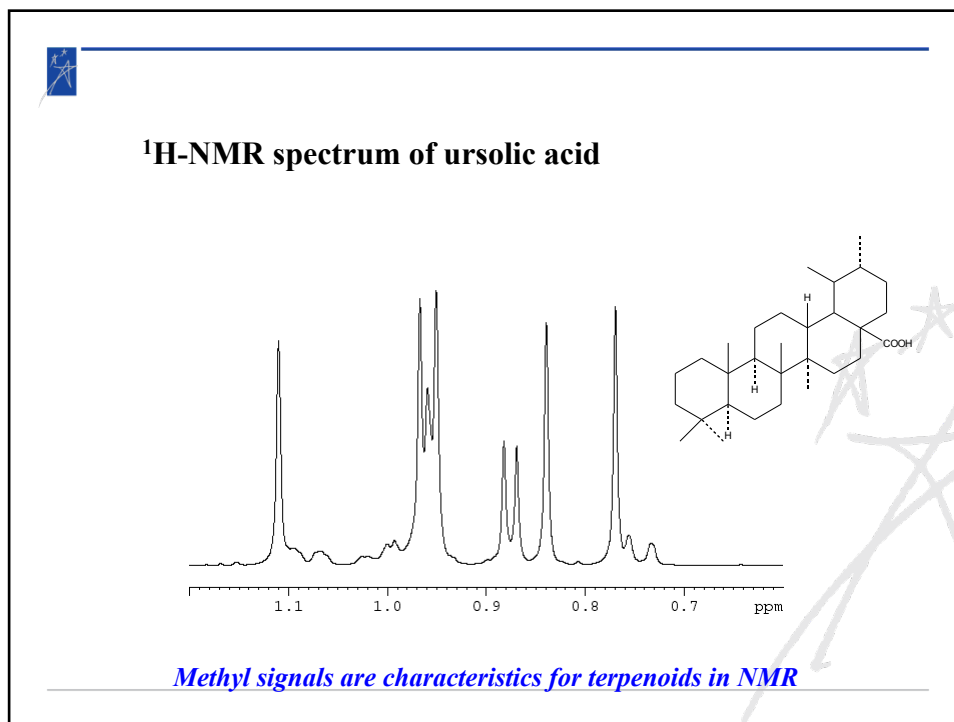
β -carotene



Analysis of terpenoid and steroid

- MS-based method is more powerful than NMR for individual terpenoid or steroid
 - GC-MS for aglycones and LC-MS for glycosides
 - Methyl and anomeric proton of sugar are characteristic features in NMR spectra
 - Many terpenoids exist as glycosides
- triterpenoids and steroids glycosides : saponin

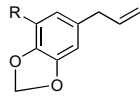




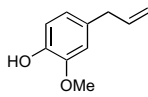


Phenylpropanoid

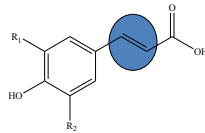
- C₆-C₃ (phenyl + propane)
- C₃ : carboxyl acid, aldehyde, alcohol or olefin
- High vapor pressure: Essential oil
- Chlorogenic acid : Ester of quinic acid and caffeic acid
- Cinnamic acid : Cinnamomum, volatile
- Coniferin, syringin, anethole, eugenol, safrole, myristicin



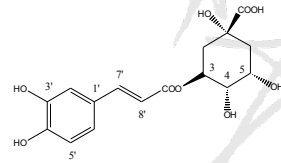
R=H, safrole
R=ome, myristicin



eugenol



R₁ = OCH₃, R₂ = H, Ferulic acid
R₁ = OH, R₂ = H, Caffeic acid
R₁ = H, R₂ = H, Coumaric acid
R₁ = OCH₃, R₂ = OCH₃, Sinapic acid

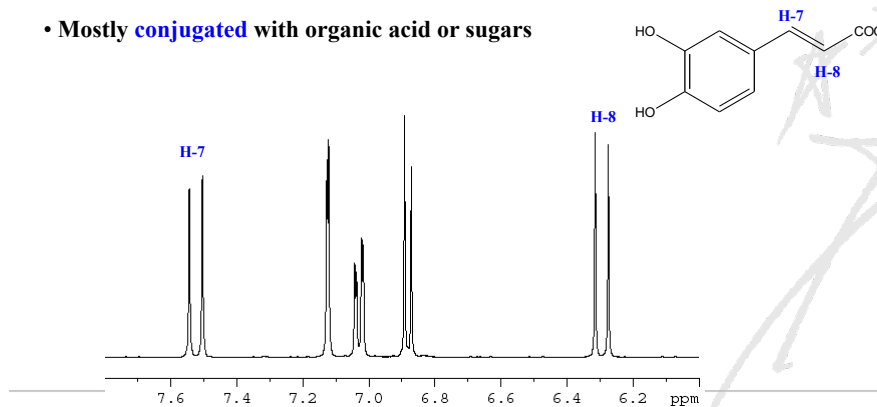


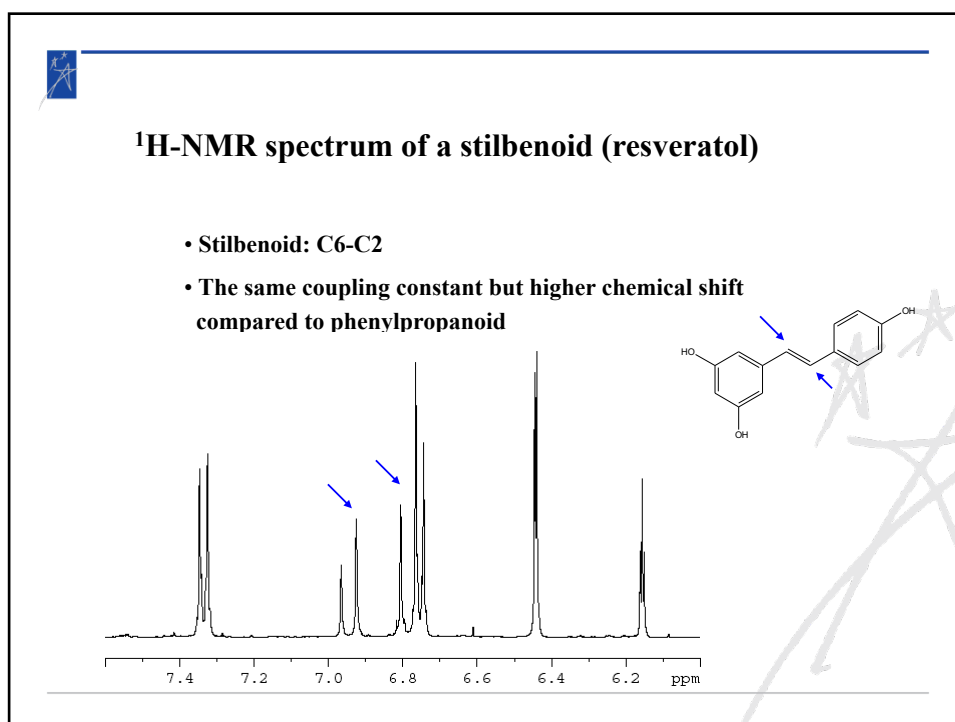
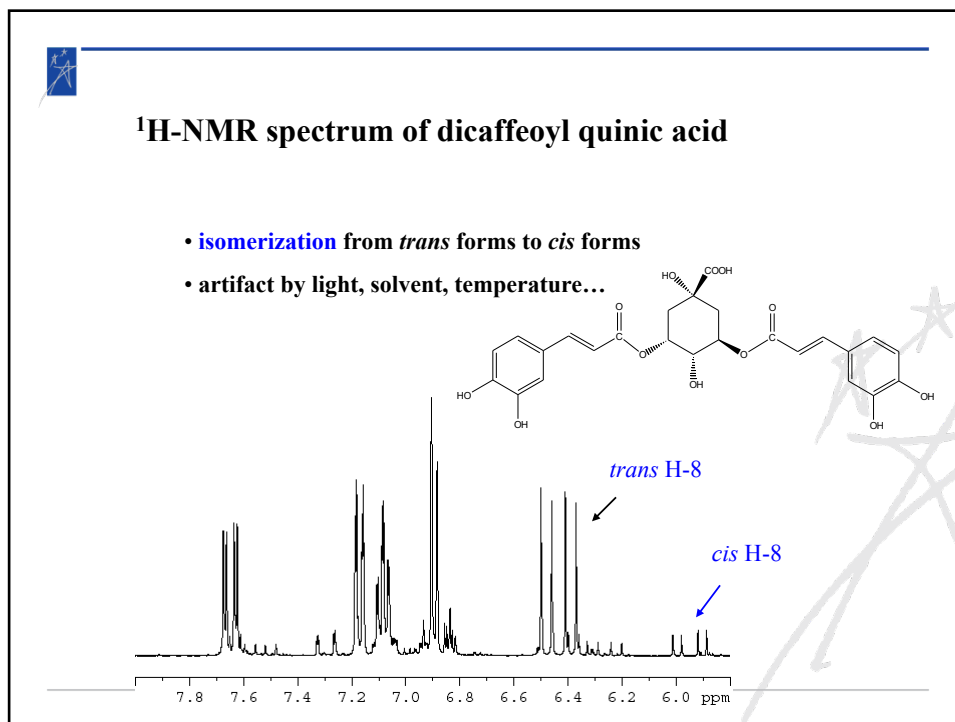
chlorogenic acid



¹H-NMR spectrum of caffeic acid

- NMR: most powerful method for structure elucidation
- Large coupling constant (16Hz): *trans* form
- Mostly conjugated with organic acid or sugars

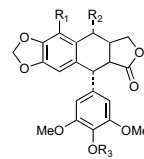
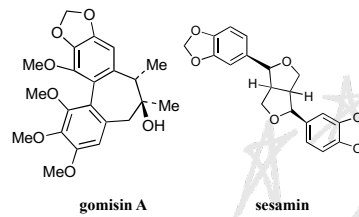






Lignan

- Oxidative coupling of C6-C3 : C₁₈
- Generally β coupling
- podophyllotoxin : *Podophyllum*, *Juniperus* species
Precursor of etoposide, teniposide
- gomisin: *Schisandra* species, hepatoprotective
- enterodiol, enterolactone: from human and animal, 1980
- pinoresinol, sesamin: polarity (+) (*Sesamum indicum*)
- neolignan: not β coupling, magnol, honokiol
- norlignan: C₁₇

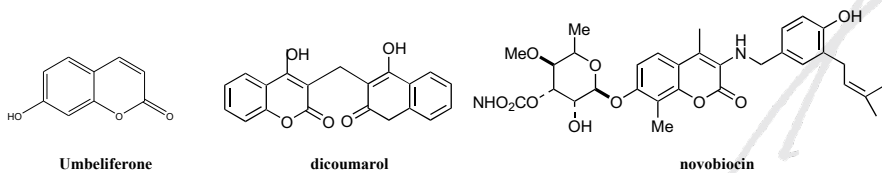
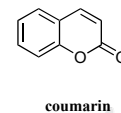


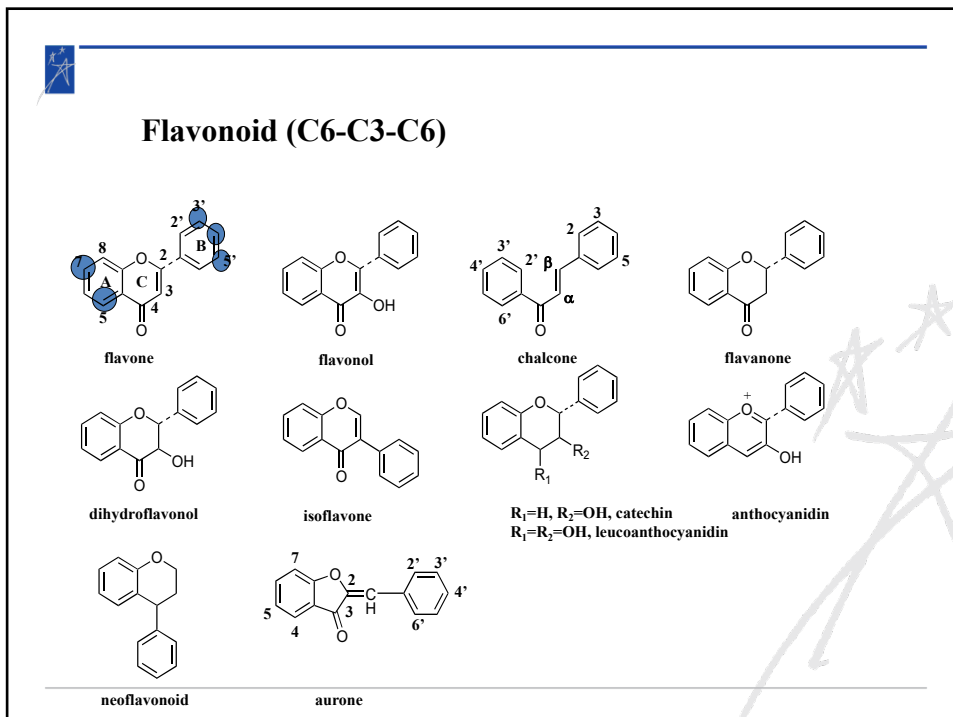
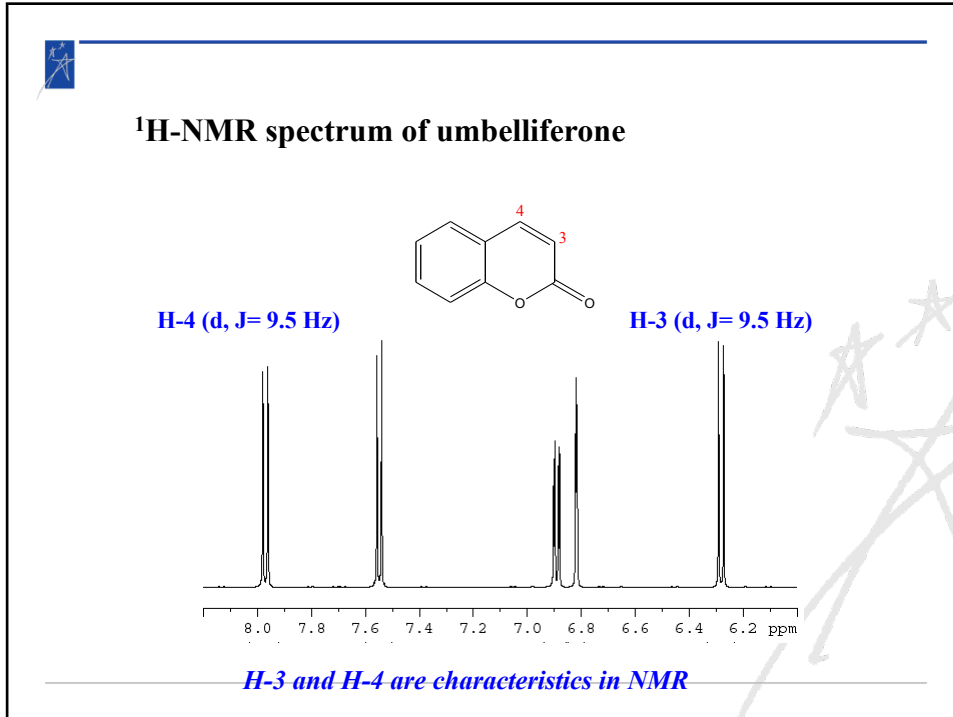
R₁=H, R₂=OH, R₃=Me, podophyllotoxin
 R₁=OH, R₂=R₃=H, α-peltatin
 R₁=OH, R₂=OH, R₃=Me, β-peltatin



Coumarin

- 2H-1-Benzopyran-2-one
- Umbeliferae, Ranunculaceae, Leguminosae, Compositae
- simple coumarin, furanocoumarin, pyroanocoumarin, biscoumarin
- Open cycle by alkali
- dicoumarol: *Medicago sativa*, *Melilotus officinalis*
inhibition vitamin K inhibition
- novobiocin (gram negative infection, *Streptomyces niveus*), umbelliferone, scopoletin

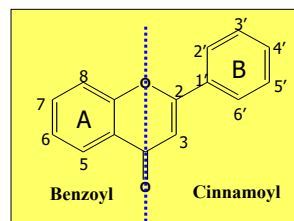






UV spectrum of flavone & flavonol

- 240-400nm region
 - Band I: 300-380 nm (B ring)
 - Band II: 240-280 nm (A ring)
1. A ring: more OH group, Band II bathochromic shift
 2. B ring: more OH group, Band I bathochromic shift



Flavone	A-ring pattern	Band II
flavone	-	250
5-OH flavone	5	268
7-OH	7	252
5,7-OH flavone	5,7	268
baicalein	5,6,7	274
norwogonin	5,7,8	281

Flavonol	B-ring pattern	Band I
galangin	-	359
kaempferol	4'	367
quercetin	3',4'	370
myricetin	3',4',5'	374

Flavonoid type	Band I
Flavone	304-350 nm
Flavonol (3-OH)	352-385 nm
Flavonol (sub. 3-OH)	328-357 nm



UV shift reagent for flavonoid

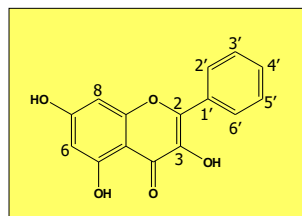
1. NaOMe: 3-OH (band I 40-65 nm bathochromic shift)
4'-OH (band I 40-65 nm shift, intensity decrease)
2. NaOAc: 7-OH (band II 5-20 bathochromic shift)
4'-OH, no 3,7-OH (band I shoulder)
3. NaOAc/H₃BO₃: B ring ortho diOH (band I 12-30nm bathochromic shift)
4. AlCl₃ & AlCl₃/HCl :
3 or 5-OH (acid stable complex), ortho-diOH (acid unstable complex)
-> B ring diOH: AlCl₃ & AlCl₃/HCl 30-40 nm shift
triOH 20 nm shift
5-OH: 35-55 nm
3-OH: 60 nm
3,5-diOH: 50-60 nm

(Ref: Systemic identification of Flavonoids, T.J.Mabry, K.R.Markham & M.B.Thomas, Springer-Verlag, 1970)

¹H-NMR spectrum of flavonoid

A-ring proton: H-6, H-8
5,7-diOH: 6.0-6.5 ppm, d, J=2.5 Hz

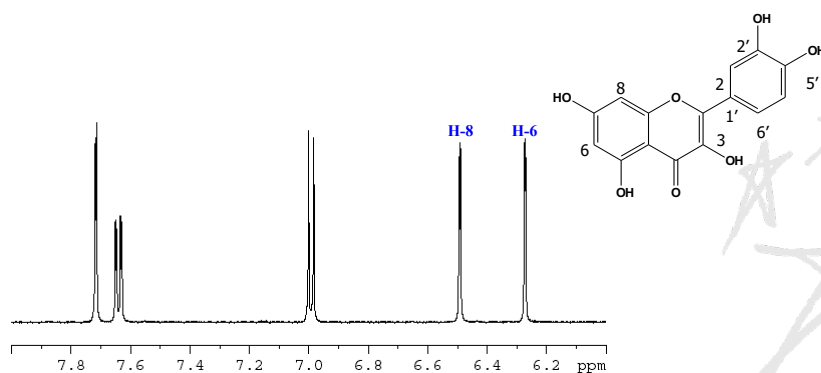
compound	H-6	H-8
5,7-diOH flavone, isoflavone	6.0-6.2	6.3-6.5
5-OH, 7-glycosyl flavone, isoflavone	6.2-6.4	6.5-6.9
5,7-diOH flavanone, dihydroflavonol	5.75-5.95	5.9-6.1
5-OH, 7-glycosyl flavanone, dihydroflavonol	5.9-6.1	6.1-6.4



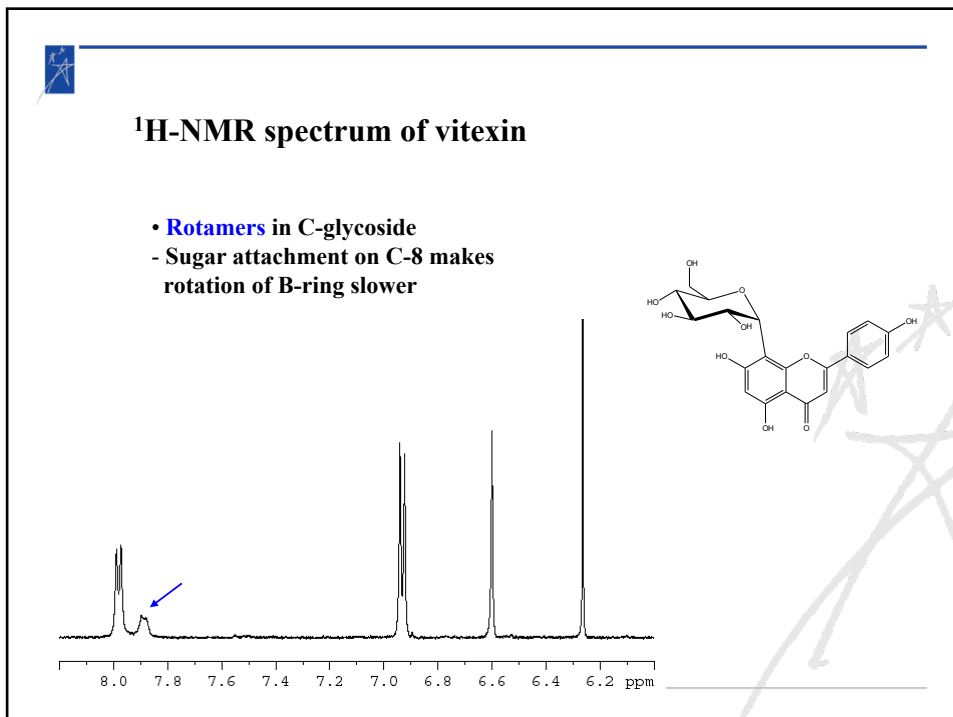
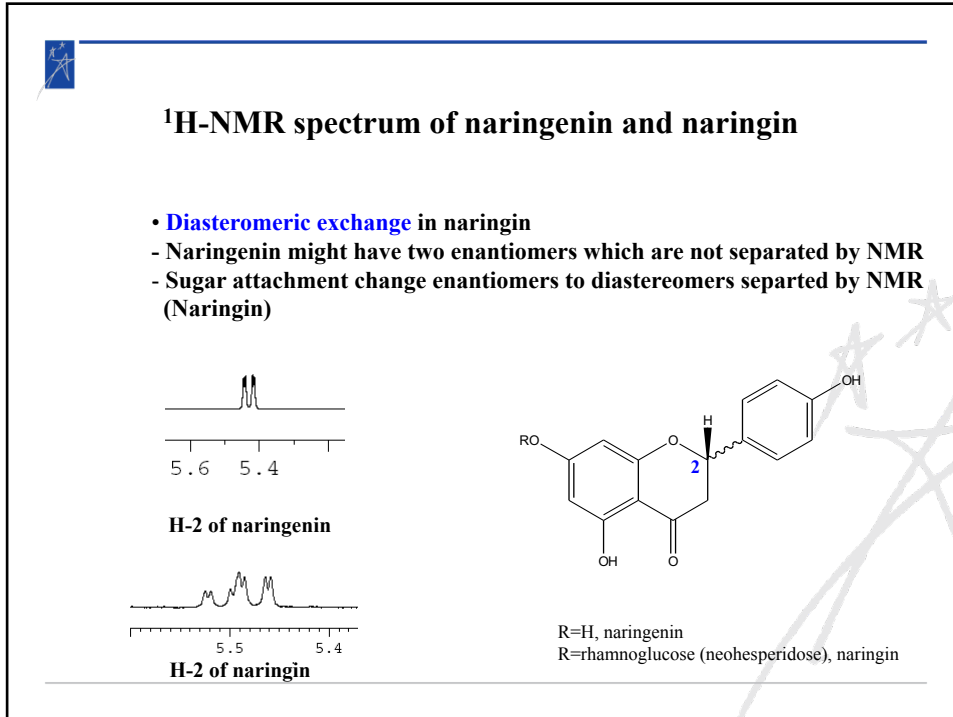
B-ring proton: downfield than A-ring, 6.7-7.9 ppm

C-ring proton: H-3 6.3 ppm

¹H-NMR spectrum of quercetin



H-6 and H-8 are characteristics in NMR

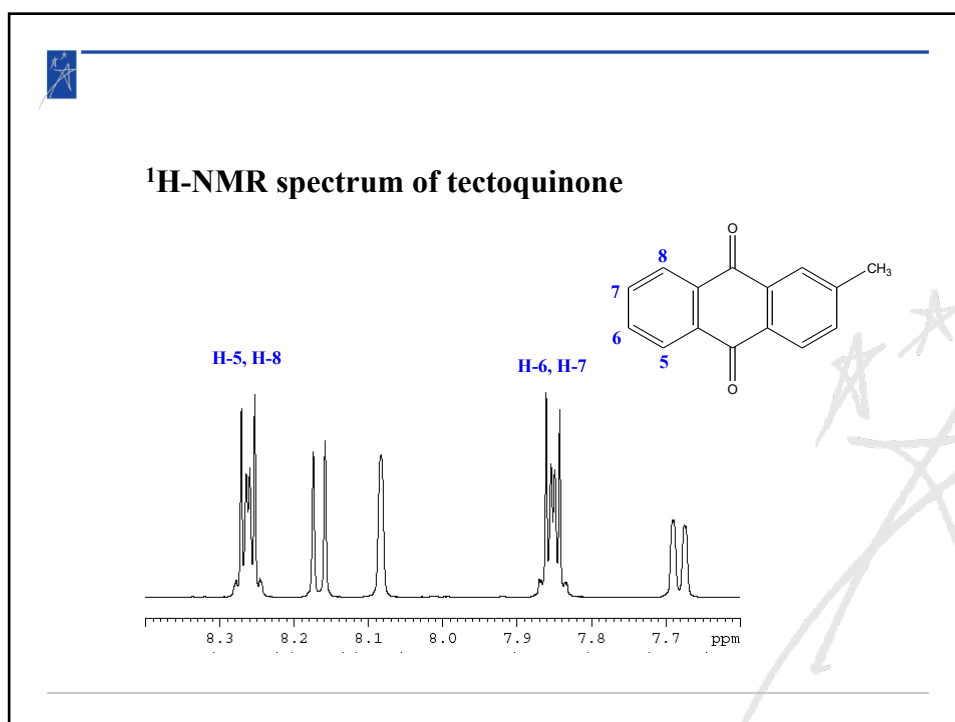


Quinone

- Animal, plant, microorganism
- Anthraquinone, naphthoquinone, benzoquinone
- Ubiquinone, vitamin K, mitomycin, sennoside, alkannin, alizarin
- Diverse color: 1,4-quinone (yellow), 1,2-quinone (red), OH makes color darker

The image shows four chemical structures with their corresponding $^1\text{H-NMR}$ chemical shifts:

- Benzoquinone:** Two doublets at 6.72 ppm.
- Naphthoquinone:** Two doublets at 7.7m and 8.1m ppm, and two doublets at 6.95 and 8.07m ppm.
- Anthraquinone:** No chemical shifts are indicated.
- Emodin:** No chemical shifts are indicated.



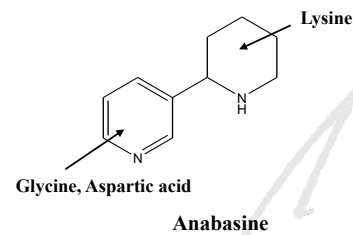
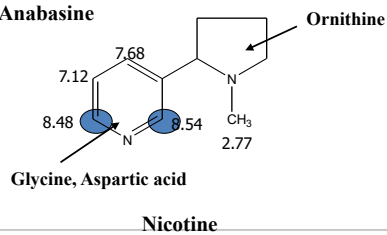


Alkaloid (1)

- Heterocyclic Nitrogen containing compound from plant, Basic, Intensive biological activity
- There is no exact definition
- Now, just nitrogen containing natural products
- Very plant-specific

Pyridine alkaloid

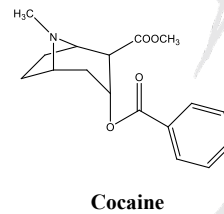
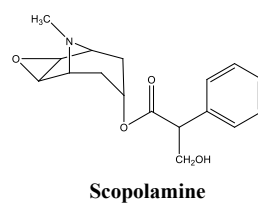
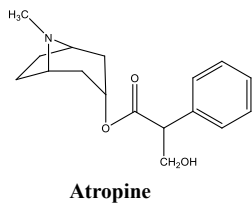
- Nicotine
- 5% in dried leaves of *Nicotiana tabacum*
- Insecticidal
- Anabasine



Alkaloid (2)

Tropane alkaloid

- Atropine: 1833, Solanaceae *Atropa belladonna*, isolated by Mein
- Cocaine: from Coca leaves
- Scopolamine

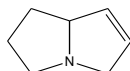




Alkaloid (3)

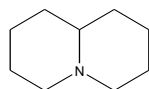
Pyrrolizidine alkaloid

- Cytotoxic, carcinogenic
- Compositae senecio species
- Senecionine, monocrotaline



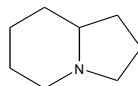
Quinolizidine alkaloid

- Lupinane group alkaloid
- Matrine, nuphridine, sparteine



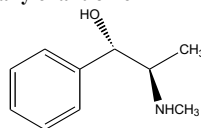
Indolizidine alkaloid

- minor
- Elaecarpine, tylophorine



Benzylamine alkaloid

- Ephedrine group alkaloid
- 1887, Nagai isolated from Ephedra species
- many enantiomer



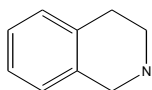
Ephedrine



Alkaloid (4)

Isoquinoline alkaloid

- Morphine, berberine
- Mostly benzylisoquinoline



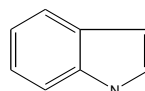
Piperidine alkaloids

- Arecoline
- *Areca catechu*



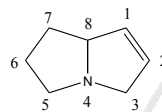
Indole alkaloid

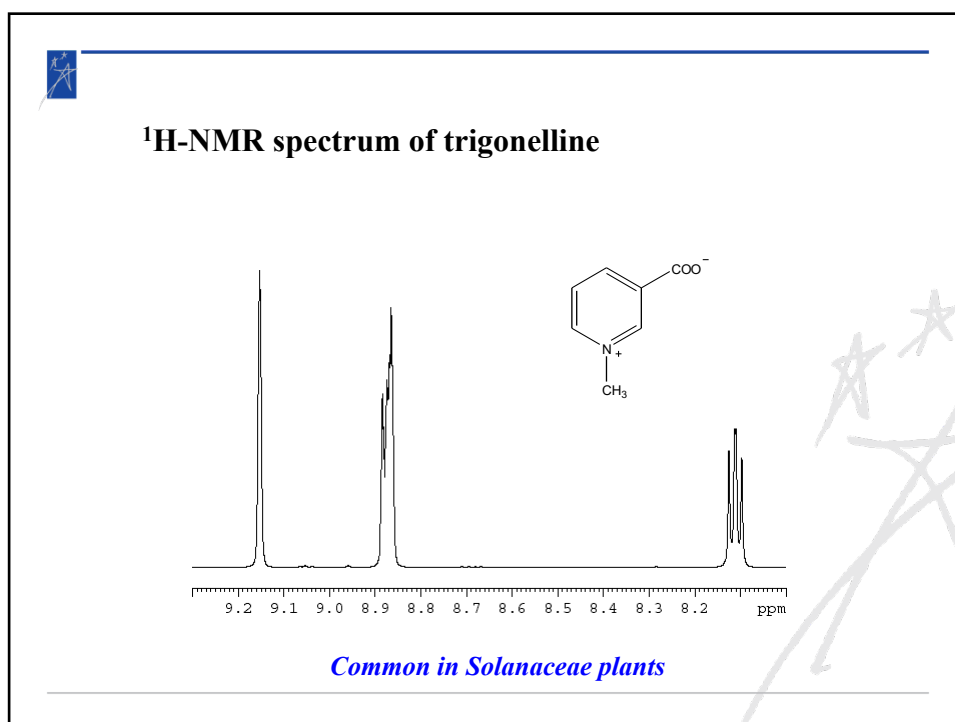
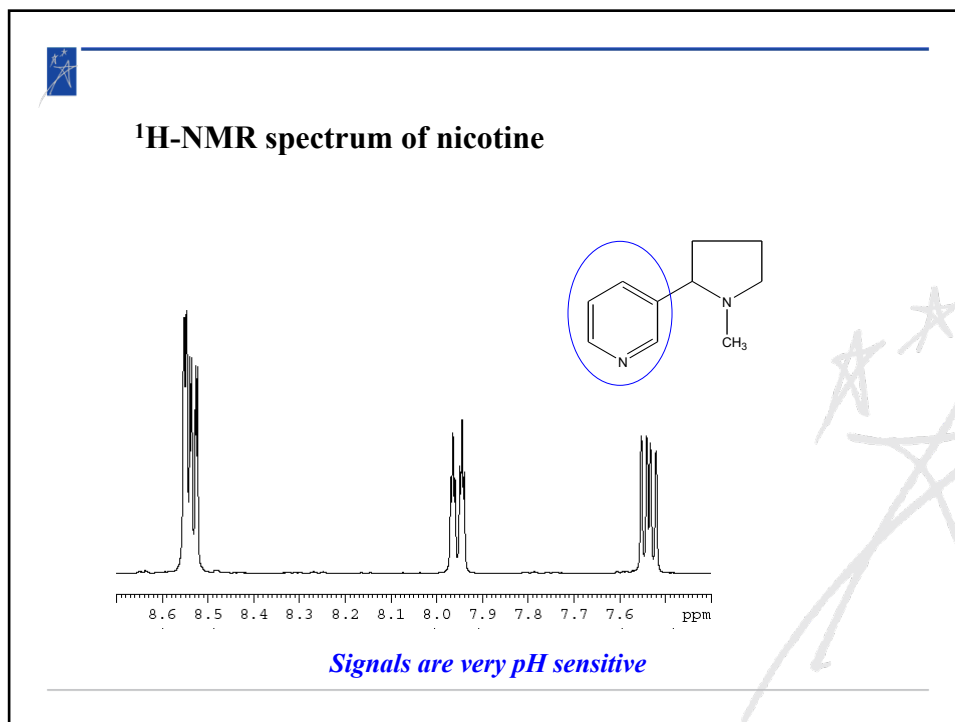
- Biosynthesized from Tryptophan, secologanin
- Reserpine (*Rauwolfia serpentina*)
- Yohimbine, vincristine, vinblastine

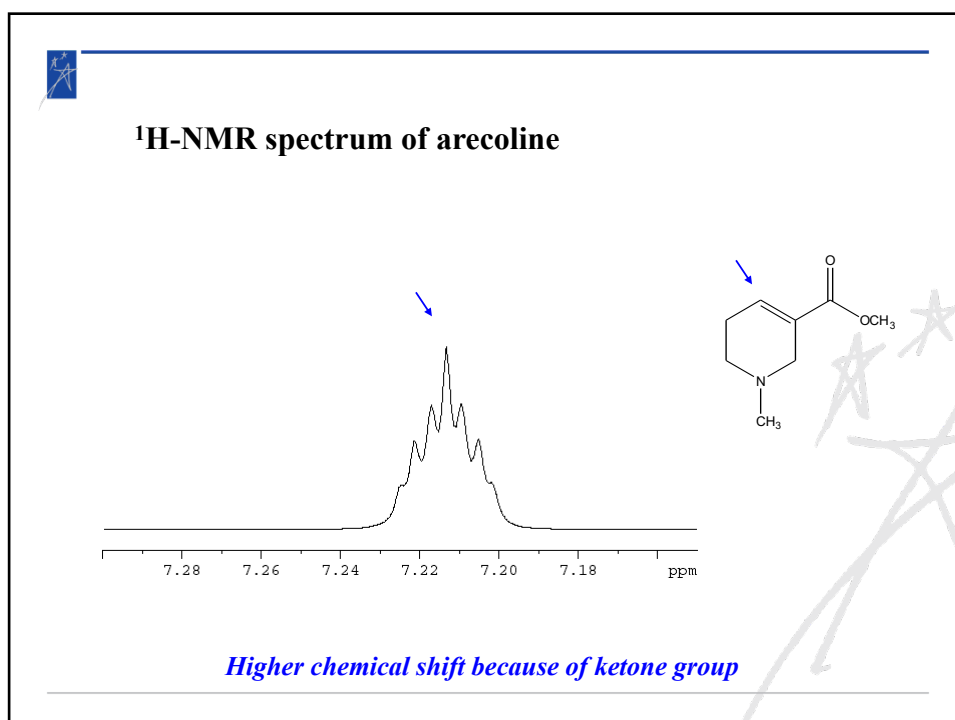
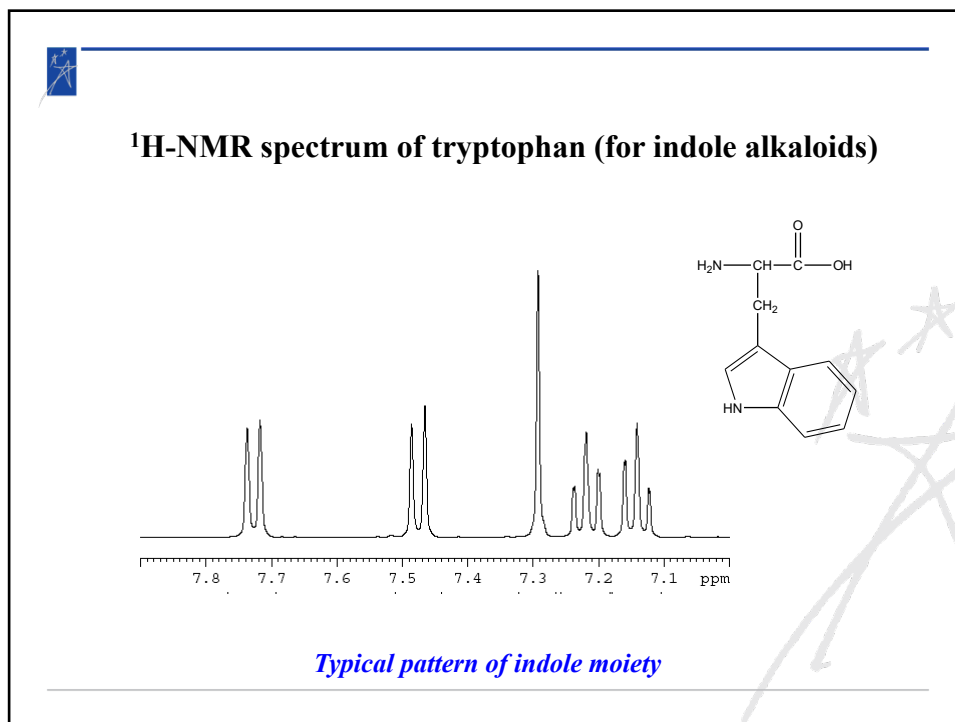


Pyrrolizidine Alkaloid

- Widespread occurrence
- Diverse range of biological activities
- Hepatotoxic



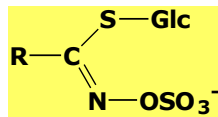






Glucosinolates

- Mostly occur in **Cruciferae** (Arabidopsis, Brassica)
- β -thioglucoside-N-hydroxysulfates



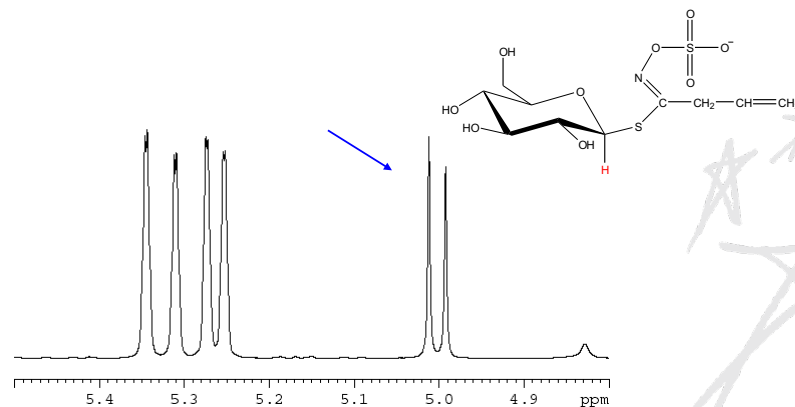
Sulfur-linked *b-D*-glucopyranose

R = side chain

- **aliphatic** (ω -methylthioalkyl) – most abundant
- **aromatic** (benzyl)
- **heterocyclic** (indole)



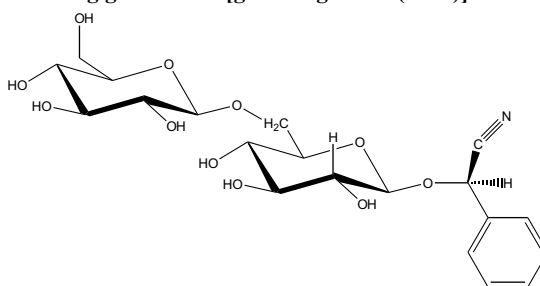
$^1\text{H-NMR}$ spectrum of sinigrin



H-1 of glucosinolate has big coupling constants (10 Hz)

Cyanogenic glycoside

- Producing HCN gas by treatment of dil. acid or alkali
- Mostly monoglucoside
- Diglycoside: amygdalin, vicianin, lucumin, linustatin, neolinustatin containing gentiobiose [glucose-glucose (1→6)]



Amygdalin

Tomorrow

- Sample Preparation and Analysis
- Application of NADES to Natural Products Research
- Preparation of NADES