PharmacoLearn

MPG103T Phytochemistry, M.Pharm Semester 1

Unit 5: Structure Elucidation by Spectroscopic Techniques

Free notes at pharmacolearn.fyi

5.1 How the four techniques combine on one compound

This unit applies the general method of Section 4.5 to nine named compounds, using UV, IR, MS and NMR (1H and 13C), in the syllabus's four groups: the monoterpenes carvone, citral and menthol; the flavonoids luteolin and kaempferol; the alkaloids nicotine and caffeine; and glycyrrhizin, the large triterpenoid saponin, proved by zones. For every compound the answer pattern is the same: formula and unsaturation from MS; chromophore from UV; functional groups from IR; the hydrogen map from 1H NMR; the carbon map from 13C NMR; and the assembly, checked against all four. The workflow figure shows that division of labour.

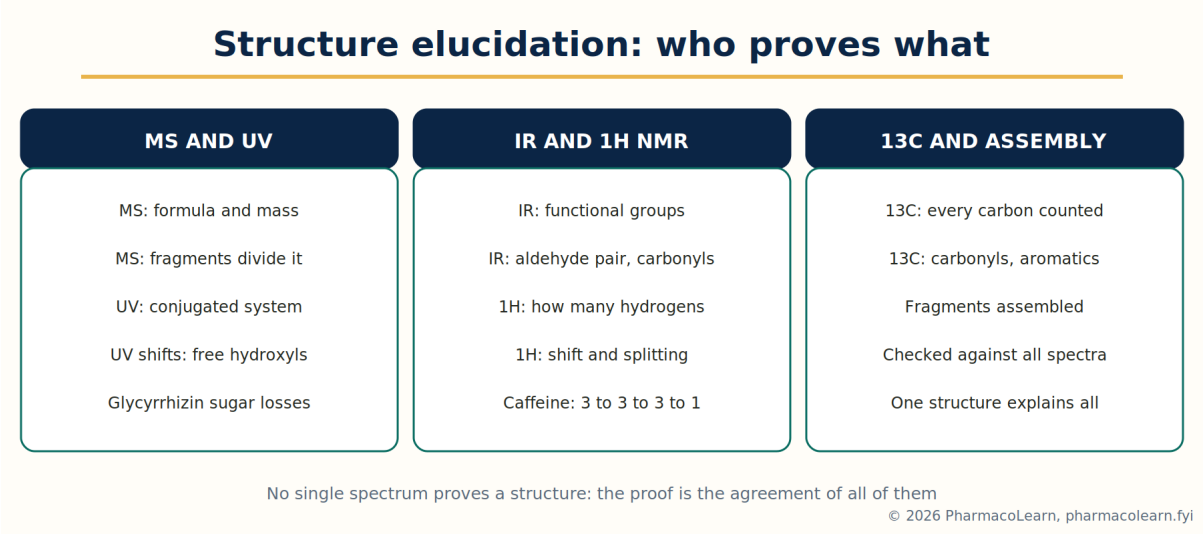

Figure: The structure elucidation workflow of this unit. No single spectrum proves a structure; the proof is the agreement of all five readings.

A discipline note governs the sections below: the spectra are taught by their standard textbook features, functional group regions, characteristic fragmentation and standard chemical shift regions. Exact values vary with solvent, concentration and instrument in real work, and an answer earns its marks by assigning the right feature to the right technique, not by memorising invented precision.

5.2 Carvone

Carvone is the monoterpene ketone of spearmint and caraway oils, C10H14O, a six membered ring carrying a conjugated enone system and an isopropenyl side chain, its five degrees of unsaturation accounted for by the ring, two double bonds and the carbonyl. The ultraviolet spectrum shows the conjugated enone absorption, separating carvone at once from a saturated or isolated ene ketone. The infrared spectrum confirms the conjugated carbonyl, at lower frequency than a saturated ketone, with alkene absorptions completing the picture. The mass spectrum gives the molecular ion and the fragmentation of a cyclic enone, cleavage beside the carbonyl and loss of side chain fragments. The 1H NMR spectrum is the core of the proof: the ring enone proton and the terminal =CH2 pair of the isopropenyl group; allylic ring signals; and two methyl singlets of different kinds, the methyl on the ring double bond and the isopropenyl methyl. The 13C NMR spectrum shows ten carbons, the ketone carbonyl far downfield and four olefinic carbons. Assembled, the spectra leave one structure: a cyclohexenone bearing a methyl on its

double bond and an isopropenyl group on the saturated carbon. Its two enantiomers, the spearmint and caraway forms, are identical in these spectra and are distinguished by optical rotation and chiral methods.

5.3 Citral

Citral is the acyclic monoterpene aldehyde of lemongrass oil, C10H16O, occurring as the geometric isomer pair geranial (citral a, E) and neral (citral b, Z); a natural sample shows doubled signals near the double bond carrying the isomerism. The ultraviolet spectrum shows the conjugated enal system. The infrared spectrum is decisive for the functional group: the aldehyde carbonyl together with the aldehyde C-H absorption, the characteristic pair no ketone shows, beside the alkene absorptions of the second, isolated double bond. The mass spectrum gives the molecular ion for the formula with four degrees of unsaturation and the fragmentation of an acyclic chain, including allylic cleavage. The 1H NMR spectrum reads the chain directly: the aldehyde proton far downfield as a doublet, coupled to the neighbouring vinyl proton; two olefinic protons in total; the methylene chain; and three methyl groups, the gem dimethyl pair and the methyl on the conjugated double bond, whose shift differs between the isomers and is the routine handle for the geranial to neral ratio. The 13C NMR spectrum shows the aldehyde carbon at the carbonyl extreme, four olefinic carbons and the aliphatic remainder, doubled in part for the pair. Citral teaches the acyclic method: integration, coupling along the chain and the two carbonyl readings, IR and 13C, carry the structure.

5.4 Menthol

Menthol is the monoterpene alcohol of peppermint oil, C10H20O, saturated apart from its cyclohexane ring, and the deliberate contrast of the group: a proof with no chromophore and no double bond to help. The ultraviolet spectrum is featureless in the working region, and stating that absence is part of the answer, since it excludes every unsaturated alternative at the formula's single degree of unsaturation, which the ring alone accounts for. The infrared spectrum gives the functional group directly, the broad hydroxyl absorption with C-O absorption and no carbonyl band, fixing an alcohol with the formula. The mass spectrum gives the molecular ion, characteristically weak for an alcohol, with water loss and isopropyl loss supporting a substituted cyclohexanol. The 1H NMR spectrum carries the proof: the carbinol proton as a multiplet whose coupling pattern reports its axial or equatorial position and therefore the stereochemistry against the isomers; the isopropyl group as two methyl doublets and a methine; the ring methyl as a further doublet; and the ring envelope. The 13C NMR spectrum shows ten carbons, all aliphatic, one oxygen bearing. Menthol's lesson is stereochemical: the constitution is proved readily, and the distinction from neomenthol and isomenthol rests on coupling constants and physical constants, the honest limit of the four techniques.

5.5 Luteolin

Luteolin is a flavone, C15H10O6, with hydroxyl groups at the 5, 7, 3 prime and 4 prime positions, and its elucidation is the flavonoid method in full. The ultraviolet spectrum is the flavone's two band pattern, band II from the A ring benzoyl system and band I from the B ring cinnamoyl system, and the shift reagents interrogate the hydroxyls in the classical manner: sodium methoxide moves band I for a free 4 prime hydroxyl, aluminium chloride and its acid comparison report the 5 hydroxyl beside the 4 keto group and any ortho dihydroxyl system, sodium acetate reports a free 7 hydroxyl, and boric acid with sodium acetate confirms the B ring's ortho dihydroxyl grouping, luteolin's catechol unit. The infrared spectrum shows the conjugated flavone carbonyl, lowered by conjugation and hydrogen bonding, with phenolic hydroxyl and aromatic bands. The mass spectrum gives the molecular ion and the retro Diels-Alder type cleavage of the C ring, whose A ring and B ring fragments carry each ring's hydroxylation count separately, the most informative fragmentation in flavonoid work. The 1H NMR spectrum shows the H-3 singlet; the meta coupled A ring protons H-6 and H-8 as small doublets; the B ring's three proton pattern of a 3,4

disubstituted ring; and the strongly downfield chelated 5 hydroxyl proton, confirming the 5 hydroxy 4 keto arrangement. The 13C NMR spectrum shows fifteen carbons, the flavone carbonyl among them. Every technique answers a different question and the answers interlock, which is the definition of a flavonoid proof.

5.6 Kaempferol

Kaempferol is a flavonol, C15H10O6, luteolin's isomer and its ideal teaching partner: the two share a formula and differ in architecture, kaempferol carrying its hydroxyls at 3, 5, 7 and 4 prime, one on the C ring and a single one on the B ring. The ultraviolet spectrum again shows the two flavonoid bands, with the flavonol character in band I's position and the shift work: sodium methoxide reports the free 4 prime hydroxyl, aluminium chloride the chelation systems at 3 hydroxyl with 4 keto and at 5 hydroxyl with 4 keto, sodium acetate the free 7 hydroxyl, while the boric acid test that answered for luteolin's catechol B ring is negative. The infrared spectrum matches the family. The mass spectrum's retro Diels-Alder fragments divide the hydroxyls between the rings differently from luteolin's, the direct spectral expression of the isomerism. The 1H NMR spectrum is the cleanest discriminator: the B ring is para substituted and shows the symmetrical four proton pattern of two doublets, against luteolin's three proton pattern; the A ring again shows the H-6 and H-8 pair; and there is no H-3 singlet, because the 3 position carries the hydroxyl, the absence itself being a proof step. The 13C NMR spectrum shows the flavonol carbon set, C-2 and C-3 both oxygen bearing. Taught together, the pair show why elucidation needs every technique: formula and mass call them identical, and only the shifts, the fragmentation division and the proton patterns separate them.

5.7 Nicotine

Nicotine is the tobacco alkaloid, C10H14N2, a pyridine ring joined to an N-methylpyrrolidine ring, and its two nitrogens of different kinds, aromatic and tertiary amine, organise its spectra. The ultraviolet spectrum is the pyridine chromophore, a modest absorption confirming an aromatic nitrogen heterocycle, supporting evidence only. The infrared spectrum shows aromatic and aliphatic C-H and C=C and C=N absorptions, with no carbonyl and no N-H, excluding amide and primary or secondary amine alternatives. The mass spectrum gives the molecular ion and cleavage at the bond joining the rings, the pyrrolidine derived fragment carrying the base peak in the standard account, as expected where charge settles on the saturated nitrogen ring. The 1H NMR spectrum divides into the two rings: four aromatic protons in the pyridine pattern, fixing the 3 substitution; the pyrrolidine envelope, with the junction methine as the link; and the N-methyl as a sharp three proton singlet. The 13C NMR spectrum shows ten carbons, five heteroaromatic, four aliphatic ring carbons and the N-methyl carbon. The assembly builds a 3 substituted pyridine and an N-methylpyrrolidine separately and joins them at the methine, which is nicotine.

5.8 Caffeine

Caffeine is the purine alkaloid of tea and coffee, C8H10N4O2, a trimethylxanthine: a fused purine dione skeleton carrying three N-methyl groups. The ultraviolet spectrum shows the xanthine chromophore's strong absorption, also the basis of caffeine's routine ultraviolet assay. The infrared spectrum shows the two carbonyls as strong absorptions, with C=C and C=N bands and no O-H or N-H absorption, an absence excluding xanthine itself and the dimethyl relatives with free NH. The mass spectrum gives the molecular ion, with fragmentation of the purine dione ring system. The 1H NMR spectrum is the most economical in the unit: four signals, three N-methyl singlets of three protons each and one lone ring proton, the C-8 hydrogen, as a singlet; integration alone, 3:3:3:1, nearly writes the structure. The 13C NMR spectrum shows eight carbons: two carbonyls, three fused ring carbons and the three N-methyl carbons. Caffeine teaches the value of counting: in a small skeleton the NMR census of methyls and the single ring proton prove what a larger molecule would need two dimensional work to show.

© 2026 PharmacoLearn | pharmacolearn.fyi

5.9 Glycyrrhizin

Glycyrrhizin, met in Unit 1 as liquorice's triterpenoid saponin, is the unit's large structure and its proof by zones. Its composition is glycyrrhetinic acid, an oleanane triterpene bearing a ketone, a hydroxyl and a carboxylic acid, joined to a disaccharide of two glucuronic acid units. The mass spectrum, by soft ionisation methods of the LC-MS family rather than electron impact, gives the molecular mass and the diagnostic losses: the glucuronic acid units lost in sequence down to the aglycone ion, establishing the sugar chain and its length. The ultraviolet spectrum reports the aglycone's conjugated enone, the one real chromophore in the molecule. The infrared spectrum is correspondingly rich: broad hydroxyl and carboxylic acid absorption, carbonyl absorptions, and the strong C-O envelope of the sugar units. The 1H NMR spectrum divides into zones: the enone olefinic proton; the anomeric protons of the two sugar units, whose couplings report their configurations; the carbinol proton at the hydroxyl position; and, most characteristically, the cluster of sharp methyl singlets of the oleanane skeleton upfield, the triterpene fingerprint counting the skeleton's methyls. The 13C NMR spectrum completes the census: ketone and carboxyl carbons at the downfield extreme, the enone pair, the anomeric and sugar carbons, and the full triterpene set. The assembly states the method for every large glycoside: prove the aglycone from its zone, the sugars from theirs, the joining by the anomeric signals and mass losses, and check the carbon and mass count against the formula.

5.10 The nine compounds in one table, and Unit 5 in one view

Compound	Class	Decisive spectral evidence
Carvone	Monoterpene enone	UV enone; IR conjugated carbonyl; 1H NMR two methyl singlets and isopropenyl =CH2
Citral	Acyclic monoterpene aldehyde	IR aldehyde C-H pair; 1H NMR aldehyde proton; doubled signals of the geranial and neral pair
Menthol	Saturated monoterpene alcohol	UV silent; IR hydroxyl with no carbonyl; 1H NMR carbinol proton and isopropyl doublets
Luteolin	Flavone	UV two bands with catechol shift responses; MS retro Diels-Alder division; 1H NMR H-3 singlet and three proton B ring
Kaempferol	Flavonol	UV flavonol shifts, boric acid negative; 1H NMR para substituted B ring and no H-3
Nicotine	Pyridine pyrrolidine alkaloid	MS base fragment from the pyrrolidine ring; 1H NMR pyridine four protons and N-methyl singlet
Caffeine	Purine alkaloid (trimethylxanthine)	IR two carbonyls, no N-H; 1H NMR three N-methyl singlets and one ring proton, 3:3:3:1
Glycyrrhizin	Triterpenoid saponin glycoside	MS sequential glucuronic acid losses; UV enone; 1H NMR methyl singlet cluster and anomeric protons

Unit 5 in one view: MS gives the formula and the fragments, UV the conjugated system, IR the functional groups, 1H NMR the hydrogen map and 13C NMR the carbon census, and the nine compounds are chosen so that each skill is decisive somewhere: conjugation in carvone, the aldehyde pair in citral, honest absence in menthol, shift chemistry in the two flavonoid isomers, nitrogen's marks in nicotine and caffeine, and proof by zones in glycyrrhizin. A structure is proved when the proposed molecule explains every reading at once.

Frequently Asked Questions

Q: Name the four great biosynthetic routes and the building blocks each one uses.

The shikimate pathway proceeds through shikimic and chorismic acids to the aromatic amino acids, feeding phenylpropanoids, coumarins and the aromatic parts of alkaloids. The acetate malonate route joins acetyl and malonyl coenzyme A into polyketide chains, feeding phenolic and anthraquinone systems. The acetate mevalonate route, with the plastid MEP pathway beside it, builds the five carbon units IPP and DMAPP, feeding every terpenoid and steroid. The amino acid routes build the alkaloids from ornithine, lysine, tyrosine and tryptophan.

Q: How does radio tracing prove a biosynthetic pathway?

A precursor labelled with a radioactive isotope, usually carbon-14, at a known position is fed to the plant. The target constituent is isolated and purified to constant specific radioactivity; if it is radioactive, the precursor's atoms reached it. The product is degraded step by step and each fragment counted: label in exactly the positions the proposed pathway predicts, and absent where it predicts absence, proves the route and the order of its steps. Carbon-13 stable isotope labelling, read by NMR and MS, is the modern partner method.

Q: Trace the origins of quinine, strychnine and the vinca alkaloids to their common precursor.

All three are monoterpene indole alkaloids. Tryptophan gives tryptamine, which condenses with secologanin to strictosidine, the family's universal precursor. Quinine arises from corynanthe type intermediates by rearrangement and opening of the indole system to build its quinuclidine ring. Strychnine arises by rearrangements folding the monoterpene unit into its cage skeleton. The vinca alkaloids are dimers: catharanthine and, through tabersonine, vindoline are coupled oxidatively to give vinblastine and vincristine.

Q: State the biosynthetic classes of digitoxin, glycyrrhizin, the sennosides and the bacosides.

Digitoxin is a cardiac glycoside whose aglycone is steroidal, built by the mevalonate route through squalene, with deoxy sugars attached. Glycyrrhizin is a triterpenoid saponin: oxidosqualene cyclises to the oleanane skeleton, oxidised to glycyrrhetinic acid and glycosylated with two glucuronic acid units. The sennosides are anthraquinone (dianthrone) glycosides of polyketide origin, from the acetate malonate route. The bacosides are dammarane type triterpenoid saponins, again from oxidosqualene.

Q: Outline the lead selection and optimization process, using artemisinin as the example.

An active extract is dereplicated against known compounds, then fractionated under bioassay guidance to the pure active, whose potency, selectivity, solubility, stability and tractability decide its selection as lead. Optimization makes analogues and maps structure activity relationships, improving those properties together. Artemisinin was isolated from Artemisia annua by low temperature extraction after the traditional fever remedy was screened; its endoperoxide bridge is the pharmacophore, and semisynthetic analogues, dihydroartemisinin, artemether and artesunate, fixed its poor solubility and short duration, with cultivation and engineered yeast fermentation solving supply.

Q: State the four phases of clinical trials and the question each phase answers.

Phase I asks whether the product is safe and tolerable and how the body handles it, in small numbers of healthy volunteers or, where required, patients. Phase II asks whether it works in the target disease and at what dose, in controlled patient studies. Phase III asks whether the benefit is confirmed at scale against the standard, in large randomised controlled trials, and supplies the evidence for registration. Phase IV, after approval, watches the product in ordinary practice for long term and rare effects.

Q: How are the solvent and the method chosen for an extraction?

The solvent is chosen by the target constituent's polarity, like dissolves like: non polar solvents for oils, waxes and the least polar terpenoids, medium polarity solvents for free bases and aglycones, and water, ethanol and their mixtures for glycosides, saponins and salts, modified by stability, selectivity, safety and purpose. The method is chosen by the heat the constituent can bear: maceration and percolation cold for thermolabile compounds, reflux and Soxhlet for heat stable exhaustive work, distillation for volatile oils, and the modern methods where speed and solvent saving matter.

Q: Distinguish successive extraction from exhaustive extraction, and name the modern extraction the syllabus emphasises.

Exhaustive extraction is extraction continued until the drug yields effectively all of the extractable constituent, as in a Soxhlet run to completion. Successive extraction extracts the same marc in turn with solvents of increasing polarity, petroleum ether, then chloroform or ether, then alcohol, then water, each to exhaustion, so the extracts arrive sorted by constituent class. The modern method the syllabus names is microwave assisted extraction, in which microwaves heat the water inside the cells, the cells rupture, and the constituents are released in minutes with far less solvent.

Q: What is a phytochemical fingerprint, and what do HPTLC, LC-MS and GC-MS each contribute to characterising an extract?

A fingerprint is the whole characteristic pattern an extract gives under a fixed method, compared against an authenticated reference for identity, consistency and adulteration. HPTLC gives the routine fingerprint and marker assay on one plate, sample and reference side by side. LC-MS identifies and profiles the non volatile constituents by mass and fragmentation. GC-MS names the volatile constituents peak by peak. Together they characterise the whole extract.

Q: How do UV, IR, MS and NMR divide the work in the structure elucidation of a phytoconstituent?

Mass spectrometry gives the molecular mass and formula, and its fragmentation reports the molecule's breakable parts, such as the retro Diels-Alder division of a flavonoid or the sugar losses of glycyrrhizin. Ultraviolet spectroscopy reports the conjugated system, with shift reagents locating free hydroxyls in flavonoids. Infrared spectroscopy identifies the functional groups, the aldehyde C-H pair of citral, the twin carbonyls of caffeine. 1H NMR maps the hydrogens by number, shift and splitting, and 13C NMR censuses every carbon. The structure is proved when one molecule explains all four readings at once, as in separating the isomers luteolin and kaempferol.

More free M.Pharm notes and unit wise PDFs at pharmacolearn.fyi
Free video lectures on the PharmacoLearn YouTube channel