



PHARMACOLEARN

It's Important to Know!

MPG103T — Phytochemistry

M.Pharm Semester 1 · Pharmacognosy

4 hrs/week, 4 credits, 60 hours, 100 marks

Complete notes — all 5 units in one PDF

Unit 1 — Biosynthetic Pathways and Radio Tracing Techniques

Unit 2 — Drug Discovery and Development

Unit 3 — Extraction and Phytochemical Studies

Unit 4 — Phytochemical Fingerprinting

Unit 5 — Structure Elucidation by Spectroscopic Techniques

Free download from Pharmacolearn

Visit this subject on pharmacolearn.fyi

<https://pharmacolearn.fyi/m-pharm/pharmacognosy/mpg-103t-phytochemistry/>

Share freely — please keep the Pharmacolearn link intact, so students always find the latest notes.

Unit 1: Biosynthetic Pathways and Radio Tracing Techniques

Free notes at pharmacolearn.fyi

Phytochemistry, subject code MPG103T, is the third core theory subject of M.Pharm Pharmacognosy, Semester 1. It carries 4 credits and 60 hours of theory, taught at 4 hours per week, for 100 marks. Its scope, as the syllabus states it, is to learn knowledge of natural product drug discovery and to be able to isolate, identify and extract the phytoconstituents. This subject asks the chemical questions underneath pharmacognosy: how does the plant build its active constituents, how were they turned into medicines, how are they got out of the plant material and separated, and how is their identity and structure proved? Five units answer in order. Unit 1 covers biosynthetic pathways and radio tracing techniques, following the named alkaloids, glycosides, steroids, a coumarin and terpenoids from building blocks to finished constituent, with their importance in the herbal industries. Unit 2 covers drug discovery and development: the history of herbs as sources of drugs, lead selection and optimization through artemisinin and the andrographolides, the product discovery process and drug registration, and clinical studies with the phases of trials and protocol design. Unit 3 covers extraction and phytochemical studies: choice of method and solvent, successive and exhaustive extraction, microwave assisted extraction and the other modern methods, fractionation, and separation including preparative HPLC and flash column chromatography. Unit 4 covers phytochemical fingerprinting: HPTLC and LC-MS and GC-MS in the characterization of herbal extracts, and the general subject of structure elucidation. Unit 5 applies UV, IR, MS and NMR (^1H and ^{13}C) to nine named compounds: carvone, citral and menthol; luteolin and kaempferol; nicotine and caffeine; and glycyrrhizin. One discipline runs through the notes: a pathway is proved by labelled atoms, an extract is known by its fingerprint, and a structure is proved when independent spectra agree.

1.1 Primary and secondary metabolism: where phytoconstituents come from

Every phytoconstituent in this unit is a product of plant metabolism, and the first distinction is between primary and secondary metabolism. Primary metabolism is the chemistry every living plant cell runs to live and grow: photosynthesis, the interconversion of sugars, and the making of amino acids, nucleotides, fatty acids and the polymers built from them. Its products are universal, because the cell cannot live without them. Secondary metabolism branches off from it and makes the specialised compounds, the secondary metabolites, that give a species its chemical character: alkaloids, glycosides, terpenoids, steroids, flavonoids and the rest of the pharmacognosy catalogue. They serve the plant in its world, deterring herbivores and microbes, attracting pollinators and screening ultraviolet light, and they are distributed unevenly, in one family or species, one organ, one stage of growth. That unevenness is why they matter to pharmacy: a compound made by almost no other organism, with a strong effect on animal physiology, is a drug lead waiting to be noticed.

Biosynthesis is the step by step enzymatic building of these compounds inside the plant, and a pathway is the ordered sequence of intermediates and enzymes from a primary starting material to the finished metabolite. The pathways draw on a small set of shared building blocks: acetyl and malonyl coenzyme A; the aromatic amino acids phenylalanine, tyrosine and tryptophan; the five carbon isoprene units IPP and DMAPP; and amino acids such as ornithine and lysine used directly as skeletons. Four great routes organise the unit, and every named compound belongs to one of them or to a combination. The shikimate pathway supplies aromatic compounds: phenylpropanoids, coumarins, flavonoids and the aromatic halves of many alkaloids. The acetate malonate route supplies polyketide compounds, and the acetate mevalonate

route, with its plastid partner the MEP pathway, supplies the isoprene units behind all terpenoids and steroids. The amino acid pathways supply the alkaloids, whose nitrogen and much of whose skeleton come straight from an amino acid. The map below shows the four routes.

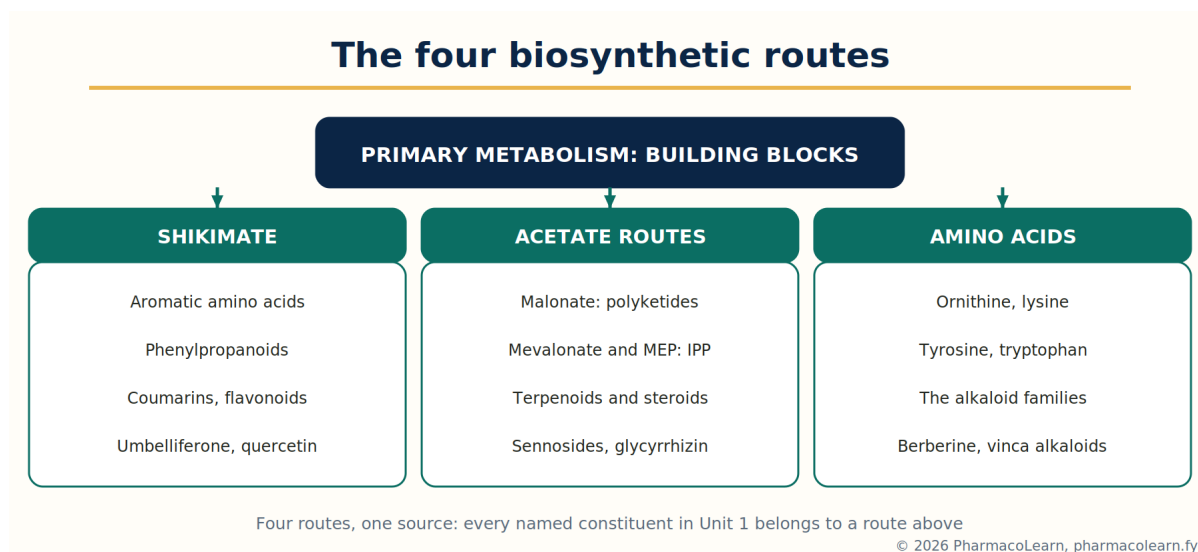


Figure: The four great biosynthetic routes of this unit. Primary metabolism supplies the building blocks; the named syllabus compounds are the end products of the routes shown.

Three working ideas recur in every pathway answer. First, the building block idea: a complex product is assembled from small units joined regularly, so its structure carries their pattern, the alternating oxygenation of a polyketide, the five carbon repeats of a terpenoid, the C6-C3 unit of a phenylpropanoid. Second, the intermediate idea: between building block and product stands a sequence of intermediates, each made by one enzyme, and the pathway is that list in order. Third, the evidence idea: a textbook pathway is a proved claim, and the proof, for most classical pathways, came from radio tracing, the subject of Section 1.5.

1.2 The shikimate pathway and the phenylpropanoids

The shikimate pathway is the plant route to aromatic compounds, named for shikimic acid, its best known intermediate, first isolated from the Japanese star anise (*Illicium*, shikimi in Japanese). It operates in plants and microorganisms and not in animals, which is why its products, the aromatic amino acids, are essential nutrients for animals. In outline, the pathway condenses phosphoenolpyruvate (from glycolysis) with erythrose 4-phosphate (from the pentose phosphate pathway) and proceeds through shikimic acid to chorismic acid, the branch point of aromatic biosynthesis, from which the plant makes phenylalanine, tyrosine and tryptophan.

For pharmacognosy the decisive branch is phenylalanine. Phenylalanine ammonia lyase removes ammonia from phenylalanine to give cinnamic acid, opening the phenylpropanoid pathway, the source of the C6-C3 family: hydroxycinnamic acids such as caffeic and ferulic acid, the coumarins (Section 1.12), the lignans, and, by joining a phenylpropanoid unit with acetate derived units, the flavonoids. Umbelliferone is a direct phenylpropanoid product, formed by hydroxylation and lactone ring closure of a cinnamic acid derivative. The flavonoids, quercetin among them, are mixed origin: their B ring and three carbon bridge come from phenylalanine, their A ring from three acetate malonate units, and examiners reward the student who can say which part came from which route. Tryptophan feeds the indole alkaloids (Sections 1.8 and 1.9) and tyrosine the isoquinoline alkaloids such as berberine, so the shikimate pathway stands behind much of the alkaloid list as well.

1.3 The acetate pathways: polyketides, mevalonate and the MEP route

Acetate, carried as acetyl coenzyme A and its carboxylated partner malonyl coenzyme A, is the second great building block system, and it works in two distinct ways that must not be confused. In the acetate malonate or polyketide mode, acetyl and malonyl units are joined head to tail into a growing chain, exactly as fatty acids are built, but the chain is then folded and cyclised, often with its oxygen atoms retained in a regular alternating pattern, to give aromatic and phenolic products: the A ring of the flavonoids, many anthraquinones including the sennosides of Section 1.10 in their ring system, and a wide range of phenolic compounds. The retained, regularly spaced oxygenation is the structural signature of a polyketide origin.

In the second mode, three acetate units are assembled into mevalonic acid, and mevalonate is converted, with loss of one carbon, into the five carbon units IPP and DMAPP. This acetate mevalonate pathway is the classical route to terpenoids and steroids. Plants also run a second route to the same units, the methylerythritol phosphate (MEP) pathway, also called the non mevalonate or DOXP pathway, building IPP and DMAPP from pyruvate and glyceraldehyde 3-phosphate in the plastids. The working rule is compartmentation: the cytosolic mevalonate pathway chiefly supplies sesquiterpenoids, triterpenoids and steroids, while the plastid MEP pathway chiefly supplies monoterpenoids, diterpenoids and carotenoids. The MEP route is the standing lesson that a pathway map is never finished: plastid terpenoids labelled in a pattern mevalonate could not explain, and following that anomaly revealed the second route.

Once IPP and DMAPP exist, terpenoid biosynthesis is an assembly line of chain lengthening and cyclisation. DMAPP and IPP join to geranyl diphosphate (C10), the monoterpene precursor behind menthol, carvone and citral. Another IPP gives farnesyl diphosphate (C15), behind artemisinin. Another gives geranylgeranyl diphosphate (C20), behind taxol and the andrographolides. Two farnesyl units joined tail to tail give squalene (C30), oxidised to 2,3-oxidosqualene and cyclised, in plants through cycloartenol, to the steroidal skeleton behind hecogenin, the guggulsterones, the withanolides, digitoxin's aglycone and the triterpenoid saponins. The table below fixes the chain lengths, the fastest check on any terpenoid answer.

Precursor	Carbons	Class built from it	Syllabus examples
Geranyl diphosphate (GPP)	10	Monoterpenoids	Menthol, carvone, citral
Farnesyl diphosphate (FPP)	15	Sesquiterpenoids	Artemisinin
Geranylgeranyl diphosphate (GGPP)	20	Diterpenoids	Taxol core, andrographolide
Squalene, then oxidosqualene	30	Triterpenoids and steroids	Hecogenin, guggulsterone, withanolides, glycyrrhizin, bacosides, cucurbitacins, digitoxin aglycone

1.4 The amino acid pathways to the alkaloids

Alkaloids are nitrogen containing, typically basic and potent secondary metabolites, and biosynthetically they are the amino acid family's products. Ornithine gives the tropane alkaloids and supplies the pyrrolidine ring of nicotine. Lysine gives the piperidine and quinolizidine alkaloids and the piperidine ring of piperine's partner unit. Tyrosine is the most productive: two tyrosine derived units join in the benzylisoquinoline pathway, which produces berberine and, by rearrangement, morphine's family; tyrosine also supplies ephedrine's skeleton in part. Tryptophan, through tryptamine, gives the indole alkaloids: joined with a monoterpene unit it produces strictosidine, the universal precursor of the monoterpene indole alkaloids, from which quinine's family, strychnine and the vinca alkaloids descend. Caffeine's purine skeleton comes from purine metabolism, not from an amino acid directly, and two syllabus names are mixed: piperine joins a lysine derived piperidine with a phenylpropanoid acid, and ephedrine a phenylalanine derived unit with a

three carbon unit. The origins map below collects the routes.

The pattern of an alkaloid pathway answer is constant. Name the amino acid precursor and the first committed intermediate where one is standard, strictosidine for the monoterpene indole alkaloids, reticuline for the benzyloquinolines. Then follow the skeleton building steps in outline, cyclisation, oxidation, methylation and rearrangement, to the finished alkaloid. Decarboxylation of the amino acid to its amine is a recurring first step, and Pictet-Spengler type condensation a recurring ring forming step.

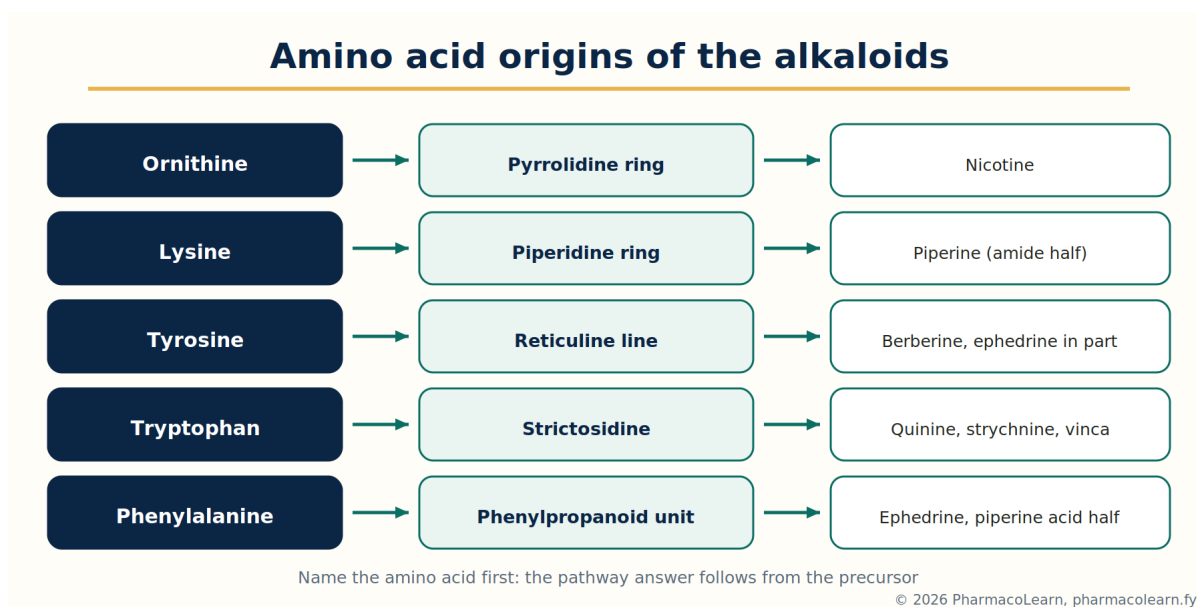


Figure: The amino acid origins of the syllabus alkaloids. Tryptophan plus a monoterpene unit, through strictosidine, is the single most productive line in the unit.

1.5 Radio tracing techniques: how pathways are proved

A biosynthetic pathway cannot be seen: its intermediates are present in tiny amounts, turn over rapidly, and sit inside a living tissue running thousands of other reactions. Radio tracing, or radiotracer technique, is the classical method that made pathways provable, and the syllabus places it in this unit's title because the pathways taught around it rest on its evidence. A precursor is prepared with one or more atoms replaced by a radioactive isotope, most often carbon-14, with tritium, phosphorus-32 and sulfur-35 used for suitable questions. The labelled precursor is fed to the plant, by root or cut stem uptake, injection, leaf application or, for photosynthesis labelling, labelled carbon dioxide. After a chosen time the product is isolated, purified rigorously to constant radioactivity, and counted, usually by liquid scintillation counting. If the product is radioactive, the precursor's atoms reached it. The decisive refinement is degradation: the product is broken down step by step into defined fragments and each is counted. Radioactivity in exactly the positions the proposed pathway predicts, and absent where it predicts absence, proves the route and the order of its steps.

The standard working methods around the feeding experiment complete the technique. Autoradiography locates radioactivity in a tissue or on a chromatogram. Isotope dilution, adding unlabelled material to trap a fleeting intermediate, tests whether a suspected intermediate lies on the path. Double labelling, feeding two isotopes at once, tests whether part of a molecule is incorporated intact. Feeding suspected advanced intermediates, which should label the product more efficiently than an early precursor if the order is right, orders the steps. Stable isotopes, carbon-13 above all, are the modern partner: the label is followed not by counting but by NMR and mass spectrometry, which read its position directly, with no radiation hazard. The tracer workflow below shows the sequence of a classical proof.

```

Prepare the labelled precursor (14C or 3H at a known position)
|
Feed to the intact plant, tissue or culture; allow metabolism
|
Harvest, extract and isolate the target constituent
|
Purify to constant specific radioactivity
|
Measure incorporation (scintillation counting)
|
Degrade the product step by step; count each fragment
|
Label in the predicted positions only: pathway step confirmed

```

Radio tracing: proving a pathway



Purification to constant activity and degradation make the proof

© 2026 PharmacoLearn, pharmacolearn.fyi

Figure: The radio tracing workflow. Purification to constant activity and the degradation step are what separate a proof from a suggestion.

The interpretation rules are as examinable as the method. Incorporation must be specific, not the general labelling that results when a precursor is broken down to labelled carbon dioxide or acetate and rebuilt into everything; position specific degradation is the guard. A low incorporation does not refute a pathway, because the fed precursor is diluted by the plant's own pool. Failure to incorporate does not exclude a compound, which may simply not reach the site of synthesis. The result is strongest when several precursors label the product in mutually consistent positions. By this logic the mevalonate and shikimate routes, and the central places of strictosidine and reticuline, were established.

1.6 The alkaloid set: the pattern for biosynthesis, isolation, characterisation and purification

The syllabus asks for the named constituents under one combined head: biosynthesis, isolation, characterisation and purification, with special reference to their importance in the herbal industries. That head sets the pattern for Sections 1.6 to 1.12. For biosynthesis, give the precursor and route of Sections 1.2 to 1.4. For isolation, the general alkaloid method is the backbone: the drug is extracted with alcohol or dilute acid, and the alkaloids are liberated by making the aqueous extract alkaline, whereupon the free bases pass into an immiscible organic solvent; the bases are drawn back into dilute acid, the solution washed, basified and re-extracted, an acid base shuttle that separates bases from neutral and acidic companions. Individual alkaloids are separated by fractional crystallisation of their salts, column

chromatography and recrystallisation. For glycosides the method differs, because they are polar and often fragile: extraction is with water or aqueous alcohol, hydrolysing enzymes are inactivated early, tannins and pigments are removed, and purification is by crystallisation and chromatography, with care against acid and enzymatic hydrolysis throughout. For steroids, saponins and terpenoids, extraction follows polarity: non polar solvents for sapogenins after hydrolysis of their saponins, alcohol for the saponins themselves. Characterisation across all groups rests on one ladder: physical constants, class colour and precipitation reactions, chromatographic behaviour against reference standards, and the spectroscopic proof of Unit 5. Purification is judged by a single chromatographic spot or peak and constant physical constants.

1.7 Ephedrine and quinine

Ephedrine is the phenethylamine alkaloid of Ephedra species (Ma Huang), a sympathomimetic long used in pharmacy. Biosynthetically it is mixed origin: its phenylpropane skeleton derives from phenylalanine by way of cinnamic acid and a benzaldehyde stage, and the side chain is completed by condensation with a three carbon unit from pyruvate, followed by transamination and N methylation, the joining of a shikimate derived unit with a carbohydrate derived unit. Isolation follows the general alkaloid method, with the practical note that ephedrine is comparatively water soluble and steam volatile for an alkaloid. Ephedra extracts are standardised on their ephedrine alkaloids, and the same potency places them under strict regulatory control in many markets.

Quinine is the cinchona alkaloid of Cinchona bark, the historic antimalarial whose isolation is one of the founding events of Unit 2. Biosynthetically it is a monoterpene indole alkaloid with a rearranged skeleton: tryptamine condenses with secologanin to strictosidine, the pathway proceeds through corynanthe type intermediates, and the indole system is then rearranged and opened to build the quinuclidine ring. That rearrangement is the examinable surprise: quinine looks nothing like a simple indole alkaloid, yet tracer work places it firmly in the strictosidine family. Isolation is industrial alkaloid extraction from bark, the total cinchona alkaloids being separated into quinine, quinidine, cinchonine and cinchonidine by fractional crystallisation of their salts. Its stereoisomer quinidine is itself a medicine, a reminder that separation of near identical alkaloids is a core phytochemical skill.

1.8 Strychnine, piperine and berberine

Strychnine, from the seeds of *Strychnos nux-vomica*, is the most complex structure in this section. It is a monoterpene indole alkaloid: tryptamine and secologanin give strictosidine, and a long sequence of rearrangements, through intermediates of the Wieland-Gumlich type, folds the monoterpene unit around the indole nitrogen into strychnine's cage of fused rings. Its extreme potency as a convulsant is why it is taught: *Strychnos* seeds must be standardised and handled as the toxic drugs they are. Isolation follows the alkaloid shuttle, strychnine being separated from its companion brucine, the dimethoxy analogue occurring with it, by fractional crystallisation of salts.

Piperine is the pungent alkaloid of black pepper, *Piper nigrum*, and the unit's clearest mixed origin alkaloid. One half is piperic acid, a phenylpropanoid derived acid carrying a methylenedioxy group; the other half is piperidine, from lysine; the two are joined as an amide. Piperine is a standardised marker for pepper extracts and a well known inhibitor of drug metabolising enzymes and efflux transport in the gut, added to some formulations as a bioavailability enhancer and an interaction caution in its own right. Isolation is by alcohol extraction of the ground fruit and crystallisation from the concentrate.

Berberine is the yellow isoquinoline alkaloid of *Berberis* and related genera and of drugs such as daruharidra and coptis. It is a benzylisoquinoline alkaloid: two tyrosine derived units build reticuline, and berberine is formed from that line by the berberine bridge reaction, in which an enzyme converts the N methyl group of a reticuline type intermediate into the methylene bridge closing berberine's extra ring,

followed by oxidation to the fully aromatic, quaternary, intensely yellow berberinium system. That permanent charge explains its characterisation: berberine is handled as a salt, is strongly coloured, and is standardised by chromatographic assay against berberine reference standard.

1.9 Taxol and the vinca alkaloids

Taxol, paclitaxel, is listed by the syllabus with the alkaloids, and the honest classification note belongs in every good answer: paclitaxel is a taxane diterpenoid, built from geranylgeranyl diphosphate through the taxadiene skeleton, carrying a nitrogen containing side chain attached as an ester, which is why teaching lists group it loosely with the alkaloids; it is not an amino acid derived alkaloid of the Section 1.4 pattern. Its biosynthesis runs from GGPP to taxa-4(5),11(12)-diene, then through extensive oxygenation of the core, attachment of the phenylalanine derived side chain, and final acylations. The industry story is supply: the original bark source, *Taxus brevifolia*, could not supply the drug sustainably, and production moved to semisynthesis from baccatin type precursors from renewable yew needles and to plant cell culture. Isolation is the separation of one diterpenoid from hundreds of close taxane relatives by repeated chromatography, and characterisation is wholly spectroscopic.

The vinca alkaloids, vincristine and vinblastine from *Catharanthus roseus*, close the list with the unit's most instructive biosynthesis. Both are dimeric monoterpene indole alkaloids: the strictosidine route builds two monomers, catharanthine and, through tabersonine, vindoline, and oxidative coupling of the two gives the dimer anhydrovinblastine, from which vinblastine and, by oxidation of one methyl group to an aldehyde, vincristine are obtained. One small change separates the two drugs and their clinical profiles differ accordingly, so the separation and assay of the dimers, present in the leaf in very small amounts beside far larger amounts of their monomers, is an industrial chromatography problem of the first rank. *Catharanthus* joins this unit to Unit 2: a common ornamental plant, a tracer proved pathway, and two of the most important anticancer medicines taken from a plant.

1.10 The glycoside set: digitoxin, glycyrrhizin, sennosides, bacosides and quercetin

A glycoside is a constituent in which a sugar part, the glycone, is joined to a non sugar part, the aglycone or genin, and the pharmacology usually lives in the aglycone while the sugar governs solubility and transport. The five syllabus names span four chemical families, and the table fixes their classes before the paragraphs take each in turn.

Constituent	Source drug	Class	Aglycone or active part
Digitoxin	<i>Digitalis purpurea</i> leaf	Cardiac glycoside (cardenolide)	Digitoxigenin, a steroidal aglycone with an unsaturated five membered lactone
Glycyrrhizin	Liquorice root, <i>Glycyrrhiza glabra</i>	Triterpenoid saponin glycoside	Glycyrrhetic acid, an oleanane triterpene
Sennosides	Senna leaflets and pods, <i>Senna alexandrina</i>	Anthraquinone glycosides (dianthrone)	Sennidins, joined rhein type anthrone units
Bacosides	<i>Bacopa monnieri</i> herb	Triterpenoid saponin glycosides (dammarane type)	Jujubogenin and pseudojujubogenin type aglycones
Quercetin	Widespread in leaves, fruits and flowers, richly in onion and <i>Sophora</i> buds as its glycosides	Flavonol (flavonoid), occurring in plants chiefly as glycosides such as rutin	Quercetin itself is the aglycone

Digitoxin biosynthesis is steroidal: the aglycone is built from the acetate mevalonate route through squalene and the sterol line, with the cardenolide's unsaturated five membered lactone formed at the side

chain, and three digitoxose units attached by glycosylation; the digitoxoses are deoxy sugars, a characterisation point examiners use. Isolation is the classical careful extraction: rapid enzyme control, defatting, aqueous alcoholic extraction, and purification distinguishing the leaf's primary glycosides from secondary glycosides such as digitoxin, formed by loss of a terminal glucose.

Glycyrrhizin, the sweet principle of liquorice, is built from oxidosqualene cyclised to the oleanane skeleton of beta amyrin, oxidised to glycyrrhetic acid and glycosylated with two glucuronic acid units. It returns in Unit 5 as a full structure elucidation example. Industrially it is extracted from the root with water or aqueous alcohol, isolated through its salts, and liquorice extracts are standardised on it.

Sennosides A and B are dianthrone glycosides: their anthraquinone ring system is polyketide in origin, the aglycone units are rhein anthrone type halves joined into the dianthrone, and glucose units complete the glycosides. Senna is standardised on the sennosides, isolated by aqueous alcoholic extraction protected against oxidation and hydrolysis, and characterised by the Borntrager type colour reaction together with chromatographic assay.

Bacosides are the dammarane type triterpenoid saponins of *Bacopa monnieri*, brahmi, built from oxidosqualene and glycosylated; bacopa extracts are standardised on total bacosides, the standardisation of a mixture of close saponins rather than of one crystal.

Quercetin carries the classification note flagged in the verification record. Quercetin itself is an aglycone, the flavonol at the centre of a large family; in the plant it occurs overwhelmingly in glycosidic form, as rutin (quercetin with rutinose, from *Sophora* buds and buckwheat), quercitrin and other glycosides, which is why the syllabus lists it with the glycosides. Its biosynthesis is the mixed shikimate and acetate pattern of Section 1.2: the B ring from phenylalanine through the phenylpropanoid route, the A ring from three malonate units, joined in the chalcone and cyclised to the flavonoid skeleton, then hydroxylated to quercetin's pattern. Unit 5 returns to its close relatives luteolin and kaempferol.

1.11 The steroid set: hecogenin, guggulsterone and withanolides

The syllabus word steroids here covers the steroidal and triterpenoid products of the squalene line, and the three names show its three industrial faces. **Hecogenin** is a steroidal sapogenin, the aglycone of the saponins of *Agave* and related plants, built from the sterol line by spiroketal ring formation at the side chain. Its importance is as an industrial starting material: sapogenins such as hecogenin and diosgenin were the raw materials from which steroid hormones and corticosteroids were semisynthesised, because the plant supplies the finished steroid nucleus. Isolation runs through hydrolysis of the saponins and crystallisation of the sapogenin.

Guggulsterone, met in the previous subject as the marker of guggulipid, is the steroidal ketone pair, the E and Z isomers, from the oleo gum resin of *Commiphora wightii*. Its origin is the sterol line with oxidative trimming to the pregnane type skeleton. Guggul extracts are standardised on total guggulsterones, the E and Z pair separated and measured chromatographically, a neat example of isomer aware standardisation.

Withanolides are the steroidal lactones of *Withania somnifera*, ashwagandha root, a family rather than one compound: a steroidal skeleton carrying a six membered unsaturated lactone in its side chain, withaferin A and the withanosides among the named members, built from the sterol line by side chain oxidation and lactonisation. Ashwagandha extracts are standardised on total withanolides, a defined sum of named members measured by chromatography, not an imagined single active.

1.12 Umbelliferone and the cucurbitacins

Umbelliferone is the syllabus coumarin, 7-hydroxycoumarin, widespread in the Apiaceae and named from the umbelliferous plants. Its biosynthesis is the phenylpropanoid pathway in miniature: phenylalanine gives cinnamic acid, hydroxylation gives p coumaric acid, ortho hydroxylation and isomerisation set up the ring closure, and lactonisation closes the coumarin ring. Coumarins are characterised by their fluorescence under ultraviolet light, strengthened in alkali, used in their chromatographic detection.

The cucurbitacins are the terpenoids of the unit: intensely bitter, highly oxygenated triterpenoids of the Cucurbitaceae, built from oxidosqualene through the cucurbitadienol skeleton and then extensively oxidised. Their bitterness at extremely low concentration and their cytotoxicity make them chemical defence compounds for the plant, safety markers for cucurbit drugs and foods, and research leads at once.

1.13 Importance in the herbal industries, and Unit 1 in one view

A biosynthetic pathway, once known, tells industry four practical things. It tells the grower which plant, organ and stage carries the most constituent, because accumulation follows the pathway's site and season. It tells the breeder and biotechnologist which steps limit production, and therefore which varieties to select and which genes a cell culture programme must keep switched on. It tells the manufacturer what else the extract will contain, the pathway's neighbours that standardisation must measure or control. And it tells the analyst which compounds are genuine markers, since a true biosynthetic product of the species is a far stronger identity marker than an added chemical. Unit 1 in one view: primary metabolism supplies four building block systems; the named phytopharmaceuticals are assembled from them by ordered pathways; radio tracing, with stable isotope work beside it, proved those pathways; and the general group methods of isolation, characterisation and purification turn the product into a standardised industrial constituent.

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

2.1 History of herbs as sources of drugs and drug discovery

The history this unit teaches is the record of a method being invented. Traditional medical systems, Ayurveda, traditional Chinese medicine, Unani and the European herbal tradition among them, accumulated over centuries a working knowledge of which plants act on the body, but bound inside crude drugs whose strength varied with season and source. The turning point came when chemistry learned to isolate the active principle. Morphine was isolated from opium in the first years of the nineteenth century, the first pure alkaloid and the proof that a plant's action could be carried by one definable substance. Quinine followed from cinchona bark, and a procession of isolations, atropine, caffeine, cocaine, digitoxin and many others, converted the herbal materia medica into the beginnings of the modern pharmacopoeia. Two lessons from that period still govern the field. The pure compound can be dosed, standardised and studied as the crude drug never could. And again and again the plant compound proved to be a lead rather than a finished drug, a starting structure chemistry then improved, as the salicylates led to aspirin.

The twentieth century added the other half of the method. Microbial natural products, penicillin first among them, showed that screening organisms systematically could find drugs no tradition had recorded. Synthesis and semisynthesis solved supply, plant cell culture and fermentation moved production off the field, and spectroscopic structure determination replaced degradation chemistry. Herbs remain a major source of leads because evolution has already screened their chemistry against living targets: a plant defence compound is, by design, a molecule that binds a protein in another organism. The two worked examples of this unit show the two characteristic modern stories: a traditional remedy decoded into a world medicine, artemisinin, and a traditional drug's major constituent developed as a standardised lead, the andrographolides.

2.2 The lead structure selection process

A lead compound is a compound with the desired activity in a form that can be developed: active enough to optimise, selective enough to be worth optimising, and chemically tractable enough that analogues can be made. The selection process runs in a settled order. The starting choice is the source, by one of three strategies: the ethnobotanical strategy, following documented traditional use, which begins with a human use signal attached; the ecological and chemotaxonomic strategy, choosing plants by family relationship to known producers or by predicted defensive chemistry; and the broad screening strategy, testing large extract libraries without a prior use claim. Each has found real drugs, and answers should present them as complements.

From the chosen source the process is activity led. The crude extract is tested in the bioassay; if active, it is fractionated and every fraction retested, the bioassay guided fractionation of Unit 1, until the activity is concentrated into one pure compound. Dereplication runs alongside: the extract is checked early against databases of known compounds, so a known compound is recognised at once and effort is spent only on new chemistry. The pure active is then assessed as a candidate lead: potency, selectivity against counter screens, solubility, stability, toxicity flags and structural novelty, with its structure fully proved. Selection is a comparison: several actives usually emerge, and the lead chosen is the one whose whole profile offers the best starting point for development. The flow below shows the selection funnel.

```

Source chosen: traditional use, family relationship or library screen
|
Extract prepared and tested in the bioassay
|
Active extract: dereplication against known compounds
|
Bioassay guided fractionation to the pure active compound
|
Structure proved; potency, selectivity, solubility assessed
|
Lead selected: the best whole profile, not merely the most potent

```

2.3 Structure development: selection and optimization of lead compounds

Once selected, a lead is developed by the cycle called lead optimization; the syllabus pairs selection and optimization because they are one continuous judgement. Optimization works on structure activity relationships: analogues are made, by semisynthesis from the natural product where it is plentiful, by modifying one functional group at a time, or by simplifying the skeleton to the part carrying the activity, and each is tested so the contribution of every part is mapped. The properties optimised together are potency, selectivity, solubility and permeability, metabolic stability and safety, and the standing difficulty is that improving one usually costs another. Natural product leads add two special problems: complexity, since many are large, stereochemically dense molecules, so semisynthesis is often the practical analogue route; and supply, since optimization is meaningless if the compound cannot be made at scale, so cultivation, cell culture, fermentation or total synthesis is part of structure development. The development flow below shows the cycle.

Lead development: the optimization cycle



Each turn of the cycle trades one property against another

© 2026 PharmacolLearn, pharmacollearn.fyi

Figure: The lead development cycle. Each turn of the loop trades potency, selectivity, solubility, stability and safety against each other until one compound carries the best balance.

Optimization ends at a development candidate: one compound, or a defined standardised preparation in the phytopharmaceutical case. In the single compound route the candidate is one optimised molecule. In the phytopharmaceutical route the product is a standardised extract carried by a characterised group of constituents, and development means fixing the source, the extraction and the marker based specification so the preparation tested is the preparation sold. The two examples below divide between the routes.

2.4 The product discovery process and drug registration

From candidate to marketed product the process is regulated, in the same outline for a natural product derived drug as for any drug. Preclinical development completes the pharmacology and runs the toxicology programme, together with a stable, manufacturable dosage form and the analytical methods that will control it. Clinical development (Section 2.7) then tests the product in humans under an approved protocol and trial authorisation. Registration closes development: the sponsor submits the full dossier of quality, preclinical and clinical evidence to the regulatory authority, in India the Central Drugs Standard Control Organisation under the Drugs and Cosmetics Act framework for a new drug, internationally in Common Technical Document format, and the authority grants or refuses marketing authorisation on that evidence. For herbal products the registration category must be stated accurately: a standardised herbal product may be regulated as a drug, as a traditional medicine under the AYUSH framework in India, or as a food supplement, and the evidence demanded follows the category and the claims. A product claiming to treat disease is held to drug standards whatever its origin: natural origin is not an evidence exemption.

2.5 Worked example: artemisinin

Artemisinin is the unit's complete drug discovery story. The source was traditional: qinghao, *Artemisia annua*, recorded in Chinese medical literature for fever and selected in a systematic screening programme of traditional remedies. The decisive insight was extraction method: the traditional preparation suggested heat destroyed the activity, and a low temperature ether extraction preserved it, giving a strongly antimalarial extract from which artemisinin was isolated. Its structure carried a feature no antimalarial had shown: a sesquiterpene lactone skeleton, built from farnesyl diphosphate through amorpha-4,11-diene and artemisinic acid, bearing an endoperoxide bridge, and that bridge is the pharmacophore, reacting in the parasite to generate the damaging species that kill it. Structure development followed the optimization pattern: artemisinin has poor solubility and a short duration, so semisynthetic derivatives were made, dihydroartemisinin, the oil soluble artemether and arteether, and the water soluble artesunate, each fixing a pharmaceutical defect while keeping the peroxide intact. Supply was developed in parallel, by cultivation, improved extraction and engineered yeast fermentation of artemisinic acid with chemical conversion. The artemisinin based combination therapies became the world standard for falciparum malaria, and the discovery was recognised by the Nobel Prize in 2015.

2.6 Worked example: the andrographolides

Andrographolide, from *Andrographis paniculata*, kalmegh, shows the phytopharmaceutical face of the same process. The plant is a major Ayurvedic drug for fever, liver complaints and infections, and its bitter diterpene lactones, andrographolide chief among them, are its characteristic constituents, built by the diterpene route from geranylgeranyl diphosphate. Lead selection here identified the major standardisable constituent group of an already used drug, and development proceeded on both tracks of Section 2.3. On the standardised extract track, andrographis products are defined by their andrographolide content, assayed chromatographically, with source and extraction fixed so the tested and sold preparations match. On the analogue track, andrographolide has been widely modified in research, at its hydroxyl groups and lactone system, seeking better solubility and potency. The contrast with artemisinin is the point: artemisinin became a family of single compound drugs, while andrographolide remains chiefly a standardised phytopharmaceutical lead and marker. Both are correct outcomes of the same process.

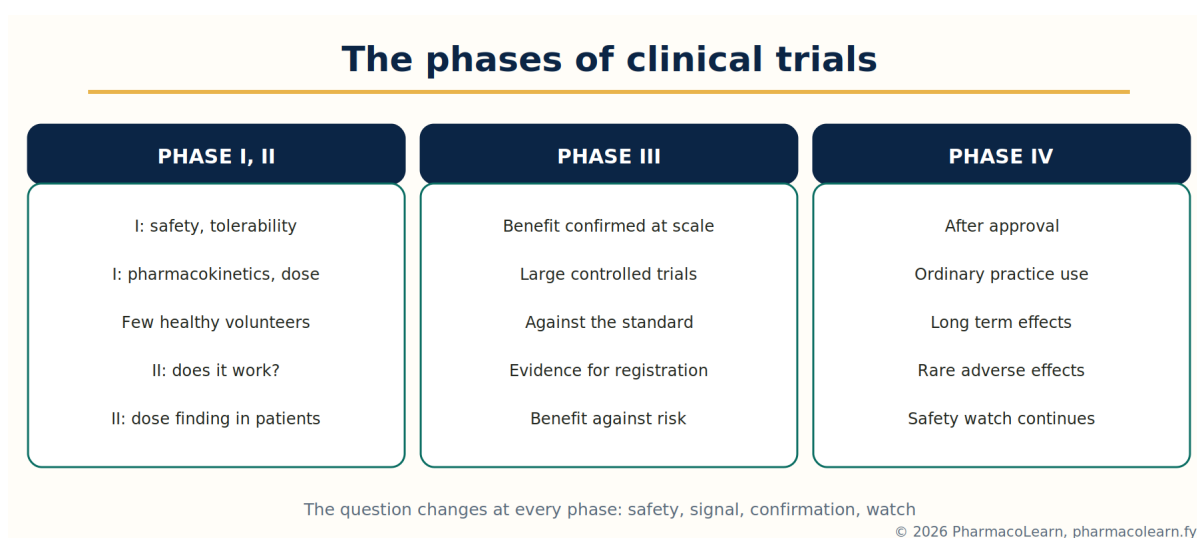
2.7 Clinical studies: the phases of clinical trials

Clinical studies are the controlled testing of a lead molecule or product in humans, and the syllabus asks for emphasis on the phases, because the phases are the structure every protocol obeys. The phases answer

different questions and must not be merged in an answer.

Phase	Main question	Who takes part	What is chiefly measured
Phase I	Is it safe and tolerable, and how does the body handle it?	Small numbers of healthy volunteers, or patients where the drug's nature requires	Safety, tolerability, pharmacokinetics, dose range
Phase II	Does it work in the target disease, and at what dose?	Patients with the condition, in controlled studies	Efficacy signals, dose response, further safety
Phase III	Is the benefit confirmed at scale against the standard?	Large patient groups, usually randomised and controlled	Confirmed efficacy, safety in broad use, benefit against risk
Phase IV	What emerges in ordinary practice after approval?	Patients in routine care after marketing	Long term and rare adverse effects, real world use

Phase I establishes the human safety and pharmacokinetic profile and the doses worth carrying forward. Phase II is the first test of the therapeutic idea in patients and the dose finding stage, where most weak leads end. Phase III provides the confirmatory evidence for registration. Phase IV continues the safety watch after marketing. For natural product leads two points are specific. The material tested must be the material defined in development, a pure compound of stated purity or an extract of stated standardisation, or the trial proves nothing transferable. And traditional use, however long, does not replace any phase for a new claim: it informs dose and safety expectations and may support a traditional use registration category, but a drug level therapeutic claim is proved in trials. The diagram below shows the phases.



© 2026 PharmacoLearn, pharmacolearn.fyi

Figure: The phases of clinical trials. Numbers grow and the question changes at every phase: safety, then dose and signal, then confirmation, then real world watch.

2.8 Protocol design for lead molecules

A clinical trial protocol is the written plan that makes a trial ethical, lawful and interpretable; protocol design for a lead molecule writes that plan around the lead's profile. Its standard elements are these. The objectives and endpoints, with primary endpoints defined before the trial starts. The design: randomised, blinded and controlled wherever the question allows, the comparator chosen and justified. The population: inclusion and exclusion criteria selecting the patients the question is about and protecting those an early lead should not yet reach. The treatment plan: dose and regimen from preclinical and Phase I evidence, route, duration, and the rules for dose change and stopping. The assessments: efficacy measures, safety monitoring and pharmacokinetic sampling. The statistics: sample size and analysis plan, fixed in advance. And the ethical frame: ethics committee approval, informed consent, trial registration and good clinical practice. For a phytochemical lead the protocol carries three extra definitions examiners look for: the exact identity and standardisation of the test material, with marker assay and batch; control of background

exposure, since patients may already take the same herb as a food or supplement; and interaction vigilance, since a lead affecting drug metabolising enzymes, as piperine does, can distort the handling of other medicines.

2.9 Unit 2 in one view

Herbs gave pharmacy its first pure drugs, and the method built on those isolations is now formal: choose the source by tradition, relationship or screen; follow activity through bioassay guided fractionation with dereplication to a pure active; select the lead on its whole profile; optimize it by analogue work; complete preclinical development; prove the product in the four phases of clinical trials under a defined protocol; and register it in the category its claims demand. Artemisinin shows the single compound outcome; the andrographolides show the standardised phytopharmaceutical outcome.

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

3.1 Selection of method and choice of solvent

Extraction is the transfer of the wanted constituents from the plant material into a solvent, leaving as much of the unwanted matrix behind as the purpose allows. The syllabus puts selection first, of method and of solvent, because almost every later success or failure is decided there. Solvent choice follows like dissolves like, made precise by polarity: non polar solvents such as hexane and petroleum ether take fixed oils, waxes and the least polar terpenoids; medium polarity solvents such as chloroform, dichloromethane and ethyl acetate take free alkaloid bases, aglycones and many terpenoids; polar solvents, ethanol, methanol and water and their mixtures, take glycosides, saponins, tannins, alkaloid salts and sugars. Ethanol and its aqueous mixtures are the industry's workhorse solvents: wide polarity range, comparatively safe, acceptable in products for human use; methanol is the laboratory's wider net, used analytically. Four further factors complete the choice: stability (a thermolabile constituent forbids a boiling solvent; an easily hydrolysed one forbids water and acid or alkali unless they are the point); selectivity (the best solvent extracts the wanted constituent with the least companion material, not the most material overall); safety and legality (flammability, toxicity, residue limits); and purpose (assay, isolation and manufacture extract differently). Method choice follows the material: particle size and hardness, the constituent's location, the scale, and the stability limits already set.

3.2 The commonly used methods of extraction

The classical methods remain in daily use because each is the right tool for one kind of material.

Maceration soaks the powdered drug in solvent in a closed vessel for days with occasional shaking, then strains and presses the marc; it suits thermolabile constituents and small scale work, at the cost of time and incomplete extraction. **Percolation** packs the drug in a percolator and lets solvent flow slowly through the bed; it is the pharmacopoeial method for tinctures and many extracts, more complete than maceration and run cold where needed. **Infusion and decoction** are the aqueous traditional forms, a short hot steep for delicate material and boiling for hard material such as roots and barks, suited to water soluble constituents and immediate use. **Digestion** is maceration under gentle heat. **Reflux extraction** boils the drug with solvent under a condenser, fast and complete for heat stable constituents. **Soxhlet extraction** is the laboratory's exhaustive continuous method: solvent vapour condenses onto the drug in a thimble, the chamber siphons back, and the drug is washed repeatedly with effectively fresh hot solvent while the extract concentrates in the flask; its standing limitation is that the extract sits at boiling point for hours, excluding thermolabile constituents. **Distillation methods** serve the volatile oils: hydrodistillation in the Clevenger type apparatus for assay, steam distillation at production scale. The figure below compares the working set.

Extraction methods compared

Maceration, percolation	Cold	Thermolabile constituents
Reflux extraction	Boiling, short	Heat stable constituents
Soxhlet extraction	Boiling, prolonged	Exhaustive laboratory work
Microwave assisted	Minutes, controlled	Rapid, low solvent extraction
Supercritical fluid	Low, pressure tuned	Non polar, solvent free extracts

Choose by the heat the constituent can bear, then by the completeness required

© 2026 PharmacoLearn, pharmacolearn.fyi

Figure: The extraction methods compared. Heat and completeness rise together up the classical ladder; the modern methods of Section 3.4 break that trade by adding energy or pressure instead of time.

3.3 Successive and exhaustive extraction

Two syllabus terms organise extraction strategy and are often confused. **Exhaustive extraction** is extraction continued until the drug gives up effectively all of the extractable constituent; the last portion of solvent shows no further constituent by the chosen check. Soxhlet extraction is exhaustive by design, and percolation is run to exhaustion for a complete extract. **Successive extraction** is a strategy, not a degree of completeness: the same marc is extracted in turn with solvents of increasing polarity, classically petroleum ether, then chloroform or ether, then alcohol, then water, each extract collected separately. It is therefore the first fractionation as well as an extraction: the non polar extract carries oils, waxes and free terpenoids, the medium extracts aglycones and free bases, and the polar extracts glycosides, saponins and tannins, so the scheme recovers the constituents already sorted by class for phytochemical screening. Its discipline is order and completeness: each solvent is run to exhaustion before the next begins.

3.4 Recent advances in extraction: microwave assisted extraction and the modern methods

Recent advances share one aim: to get the constituent out faster, with less solvent and less damage, by putting energy or pressure where the plant structure resists. **Microwave assisted extraction (MAE)** is the method the syllabus names. Microwaves heat the water inside the plant cells directly, because the field acts on polar molecules; the heated water builds pressure, the cell walls rupture, and the constituents are released into the surrounding solvent, which may stay comparatively cool. Extraction times fall from hours to minutes and solvent volumes fall, with the cautions that the solvent must interact with the field or the material must carry moisture, and that temperature must be controlled for thermolabile compounds. Beside MAE the modern set should be known in outline. **Ultrasound assisted extraction** uses cavitation, the collapse of microscopic bubbles, to batter cell walls; it is simple and cold running. **Enzyme assisted extraction** uses cellulase and pectinase type preparations to digest the wall itself. **Pressurised liquid extraction** keeps a solvent liquid above its boiling point under pressure, gaining hot solvent speed in a fraction of Soxhlet's exposure. **Supercritical fluid extraction**, treated in Section 3.6, replaces the liquid solvent with a dense gas whose power is tuned by pressure. The comparison table fixes the set.

Method	Driving force	Heat exposure	Best suited to
Maceration	Diffusion into standing solvent	None (cold)	Thermolabile constituents, small scale
Percolation	Solvent flow through a packed bed	None to low	Tinctures and pharmacopoeial extracts
Soxhlet	Repeated washing with condensing solvent	High and prolonged	Exhaustive laboratory extraction of stable constituents
Microwave assisted	Intracellular water heated, cells rupture	Short, controlled	Rapid marker extraction, reduced solvent work
Ultrasound assisted	Cavitation disrupts walls	Low	Fast laboratory and pilot extraction
Supercritical fluid	Tunable dense gas solvent	Low	Non polar constituents, clean extracts, volatile oils and decaffeination type work

3.5 Methods of fractionation, and the isolation scheme for a crude extract

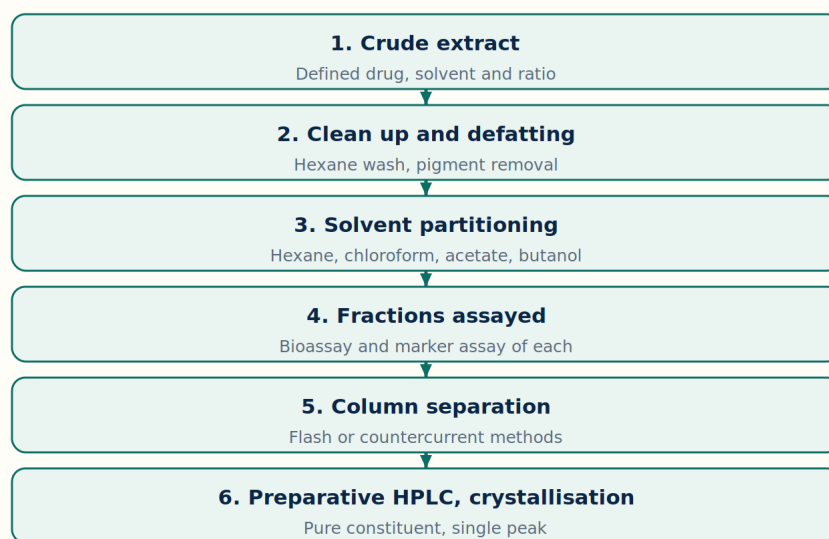
Extraction gives a mixture; fractionation begins the sorting. The standard methods, in their usual order, are these. **Solvent partitioning** distributes the crude extract between immiscible solvents, classically from water or aqueous alcohol into hexane, chloroform or dichloromethane, ethyl acetate and n-butanol in turn: each layer is a fraction sorted by polarity, the saponins preferring butanol and the aglycones ethyl acetate. **Acid base fractionation** washes an alkaloid bearing extract between acid and alkali, as in Unit 1, separating bases from neutrals and acids. **Precipitation and crystallisation** drop a constituent or an impurity out of a concentrate by solvent change, cooling or reagent addition, lead acetate precipitation of tannins being the classical clean up. **Adsorption clean up** strips chlorophyll, tannins or pigments on a short adsorbent or polyamide bed. **Size based fractionation**, dialysis, membrane filtration and gel filtration, sorts by molecular size. The full scheme for working a crude extract to a pure constituent is the daily algorithm of the phytochemistry laboratory.

```

Crude extract (defined solvent, defined drug to solvent ratio)
|
Defat and clean up where needed (hexane wash, pigment removal)
|
Solvent partitioning: hexane, chloroform, ethyl acetate, butanol, water
|
Bioassay and marker assay of every fraction
|
Active or marker rich fraction taken forward
|
Column chromatography, flash or countercurrent separation
|
Preparative HPLC for final separation of close compounds
|
Crystallisation; purity by single peak and constant constants
|
Pure phytoconstituent, ready for structure elucidation

```

Isolating a phytoconstituent



No fraction is discarded untested: the assay directs the scheme

© 2026 PharmacoLearn, pharmacolearn.fyi

Figure: The isolation scheme for a crude extract. Every fraction is assayed before any is discarded; the scheme is Unit 3 drawn as one page.

3.6 Separation by CCCET and SCFE techniques

The syllabus sentence for this section reads: separation of phytoconstituents by latest CCCET and SCFE techniques including preparative HPLC and flash column chromatography. The abbreviation CCCET is printed without expansion, and these notes do not invent one for it, as the verification record flags. What can be taught honestly is the technique family its context points to, the countercurrent family, in which two immiscible liquids do the separating and there is no solid stationary phase, together with SCFE, read in context as supercritical fluid extraction. Both are taught on that openly stated basis.

Countercurrent separation works by partition alone. One liquid phase is held stationary and the other pumped through it; each constituent distributes itself between the phases by its partition coefficient and travels at its own rate, liquid liquid extraction repeated hundreds of times in one instrument. Its instrument form is countercurrent chromatography, including its high speed planetary centrifuge forms, in which the stationary liquid is retained in a coiled column by centrifugal force. Its virtues are characteristic: no solid adsorbent means no irreversible adsorption, so delicate polar compounds, saponins, glycosides and tannins, are recovered fully; capacity is large; and the solvent system is designed from test tube partition experiments by the logic of Section 3.5. Craig countercurrent distribution is the historical form of the same idea.

Supercritical fluid extraction and separation (SCFE) uses a fluid above its critical temperature and pressure, in practice almost always carbon dioxide, whose solvent strength is tuned by changing the density: low density carbon dioxide takes the most non polar constituents, higher density takes more, and a polar co solvent such as ethanol extends the range to medium polarity. On releasing the pressure the carbon dioxide leaves as a gas, so the extract is solvent free by construction and the working temperature is low, suiting thermolabile and volatile constituents. Decaffeination, hops and spice extraction and volatile oil work are the standard industrial examples. Its limitation mirrors its virtue: very polar constituents such as most glycosides dissolve poorly even with co solvents, so SCFE complements rather than replaces solvent extraction.

3.7 Preparative HPLC and flash column chromatography

The two named chromatographic techniques divide the final separation work by scale and resolving power. **Flash column chromatography** is rapid medium pressure column chromatography on silica gel of defined particle size, the solvent driven by air or nitrogen pressure rather than gravity. It is the workhorse purification step: gram scale loads, runs in tens of minutes, fractions monitored by thin layer chromatography, and resolving power well beyond a gravity column because the finer, uniform packing sharpens the bands. Its method is chosen from TLC first: the solvent system placing the target band at a suitable retention on a plate is transferred to the column, commonly as a gradient, and that plate to column discipline is the practical skill taught. **Preparative HPLC** is high performance liquid chromatography run for collection: larger columns and loads and a fraction collector in place of a purely recording detector. It is the final tool, used where flash chromatography cannot resolve the mixture, for close alkaloid pairs, taxane relatives, the dimeric vinca alkaloids and the last purification before structure proof. Its cost in solvent and time is why it stands last in the scheme of Section 3.5, applied only to fractions already cleaned and enriched. Separation is staged: cheap bulk methods first, resolving power spent only where it is needed.

3.8 Unit 3 in one view

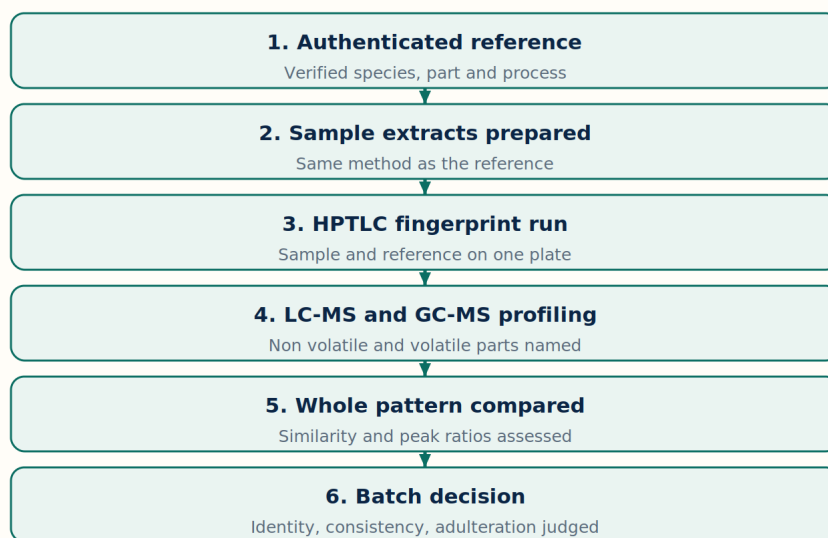
Choose the solvent by the constituent's polarity and stability and the method by the heat it can bear; extract successively to sort by class or exhaustively to recover in full; and use the modern methods, microwave assisted extraction above all, where speed and solvent saving matter. Fractionate by partitioning, acid base behaviour, precipitation and size, following activity and markers at every step. Separate finally by the staged ladder: countercurrent methods where a solid phase would lose the compound, supercritical fluid extraction for a clean non polar extract, flash chromatography for working purification and preparative HPLC for the separations nothing else can finish.

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

4.1 The fingerprint idea: characterising a whole extract

Units 1 to 3 deal with single constituents. This unit changes the object: the herbal extract itself, a mixture no single marker assay can describe. A phytochemical fingerprint is the characteristic pattern an extract produces under a defined method: the set of peaks on a chromatogram, with their positions and relative sizes, generated under fixed conditions. Its logic is that of a human fingerprint: the pattern as a whole identifies the material, though no single peak is the identity. A genuine extract of the right species, part and process gives its pattern; an adulterated, substituted or exhausted extract gives a different one; two batches of the same product give patterns alike within defined limits. Fingerprinting therefore answers the three questions a marker assay cannot: is this the right plant and part (identity), is this batch like the approved batch (consistency), and has anything foreign or missing changed the whole composition (adulteration and deterioration). Its working rules follow: the method is fixed in every detail, reference material of verified identity is run beside the sample, and the whole pattern is compared, by profile similarity and the ratios of characteristic peaks, with marker peaks identified inside the pattern where standards exist. The workflow below shows a characterisation built this way.

Fingerprinting a herbal extract



The pattern is the identity: no single peak decides the batch

© 2026 PharmacoLearn, pharmacolearn.fyi

Figure: The fingerprinting workflow. The reference extract and the sample travel through identical conditions, because the comparison is only valid inside one fixed method.

4.2 HPTLC fingerprinting

High performance thin layer chromatography is the fingerprinting method the syllabus names first, and the identity workhorse of herbal quality control. HPTLC uses plates with a fine, uniform layer, most often silica gel, samples applied as narrow bands, development in a controlled chamber, and instrumental scanning of the finished plate, densitometry, under ultraviolet light and after derivatisation where colour forming reagents reveal further zones. One plate carries the sample and the reference standards and reference

extract side by side in the same run under identical conditions, so the comparison the fingerprint idea demands is built into the format. The fingerprint is the sequence of zones at their retention factors, with their colours under ultraviolet light and after spraying: a senna fingerprint shows its sennoside zones against sennoside standards, a liquorice fingerprint its glycyrrhizin zone, and the accompanying zones distinguish the genuine drug from a substitute even where the marker is present. HPTLC also quantifies, the scanner measuring zone intensity against standards on the same plate, so one run delivers identity, fingerprint comparison and marker assay together. Its limits are moderate resolution against HPLC and sensitivity to chamber conditions, which is why saturated chambers, fixed development distance and system suitability checks are part of the method.

4.3 LC-MS in the characterisation of herbal extracts

Liquid chromatography with mass spectrometry joins HPLC separation to mass spectrometric identification, and in extract characterisation it does three jobs no optical detector can do. First, identification by mass: each peak carries its molecular mass and, with tandem mass spectrometry, a fragmentation pattern giving the class and often the exact identity of the constituent, matched against standards and libraries; a peak co eluting with a marker under ultraviolet detection is confirmed or exposed by its mass. Second, profiling of the unknown: an extract's constituents are listed by mass and fragmentation in one run, which is how the composition behind a fingerprint is mapped and how adulterants, synthetic drug spikes and degradation products are found. Third, sensitive quantification by selected ion monitoring, where a marker is present at very low level or in a crowded matrix. The cautions belong with the power: mass alone does not distinguish isomers; ionisation varies between compounds, so peak size is not a measure of amount without calibration; and cost and expertise place LC-MS as the investigation tool behind the routine fingerprint, the instrument that explains the pattern HPTLC or HPLC records.

4.4 GC-MS in the characterisation of herbal extracts

Gas chromatography with mass spectrometry is the corresponding tool for the volatile fraction: volatile oils, monoterpenes and sesquiterpenes, and any constituent that can be made volatile, by derivatisation where needed. The gas chromatograph separates the vapourised mixture on a capillary column with a resolving power liquid methods cannot match for volatiles, and the mass spectrometer identifies each peak by its electron impact fragmentation pattern, matched against standard libraries, with the retention index as the second identifying coordinate. A volatile oil fingerprint by GC-MS is therefore a named list, not merely a pattern: a mint oil is characterised by its menthol and menthone peaks, a citrus oil by its limonene led profile, and adulteration shows as peaks and ratios the genuine profile does not contain. The syllabus pairing of LC-MS and GC-MS is a coverage statement: LC-MS characterises the non volatile majority, glycosides, saponins, alkaloids and phenolics, GC-MS the volatile minority, and a full characterisation uses both.

4.5 Structure elucidation of phytoconstituents: the general method

This unit ends where Unit 5 begins, with structure elucidation stated as a general method. To elucidate a structure is to prove the complete constitution of a pure constituent: its molecular formula, its skeleton, the position of every functional group and substituent, and where the method allows, its stereochemistry. The evidence is gathered in a fixed order, each step narrowing what the molecule can be. The molecular formula comes from mass spectrometry, checked against the carbon count of the ^{13}C NMR spectrum. The degrees of unsaturation from the formula state how many rings and multiple bonds the structure must contain. The ultraviolet spectrum reports the conjugated system, if any, and shift reagents in flavonoid work report the hydroxylation pattern. The infrared spectrum reports the functional groups, each in its standard region. The ^1H NMR spectrum reports the hydrogen environments: how many kinds, in what numbers by

integration, in what surroundings by chemical shift, and next to what by splitting. The ^{13}C NMR spectrum reports every carbon environment, with editing experiments distinguishing CH_3 , CH_2 , CH and quaternary carbons. Two dimensional NMR then assembles the fragments into one skeleton, and the assembled structure is checked against every spectrum, since a correct structure must explain all of the evidence, not most of it. Unit 5 works this method on the nine named compounds.

4.6 Unit 4 in one view

A herbal extract is characterised as a whole by its fingerprint, compared entire against an authenticated reference for identity, consistency and adulteration. HPTLC delivers the routine fingerprint and marker assay on one plate. LC-MS identifies and profiles the non volatile constituents by mass and fragmentation. GC-MS names the volatile constituents peak by peak. And structure elucidation proves a pure constituent's structure by the ordered agreement of mass, ultraviolet, infrared and NMR evidence, the method Unit 5 now applies.

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

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 ^{13}C), 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 ^{13}C NMR; and the assembly, checked against all four. The workflow figure shows that division of labour.

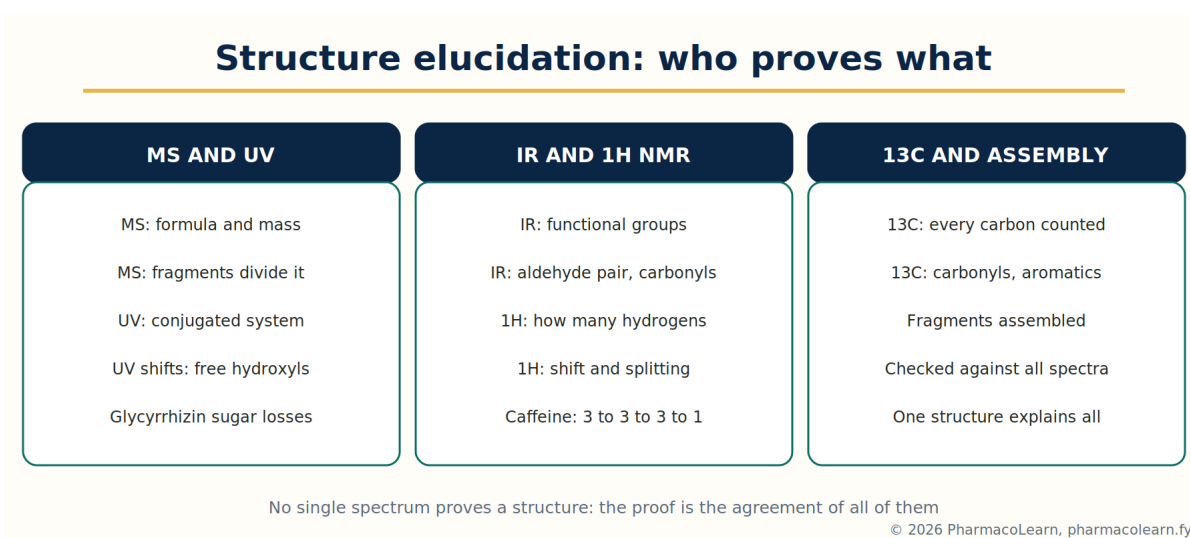


Figure: The structure elucidation workflow of this unit. No single spectrum proves a structure; the proof is the agreement of all five readings.

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, $\text{C}_{10}\text{H}_{14}\text{O}$, 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 enone 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 $=\text{CH}_2$ 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 ^{13}C 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, $C_{10}H_{16}O$, 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 ^{13}C 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 ^{13}C , carry the structure.

5.4 Menthol

Menthol is the monoterpene alcohol of peppermint oil, $C_{10}H_{20}O$, 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 ^{13}C 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, $C_{15}H_{10}O_6$, 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 ^{13}C 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, $\text{C}_{15}\text{H}_{10}\text{O}_6$, 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 ^{13}C 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, $\text{C}_{10}\text{H}_{14}\text{N}_2$, 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 ^{13}C 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, $\text{C}_8\text{H}_{10}\text{N}_4\text{O}_2$, 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 ^{13}C 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.

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 glycyrrhetic 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 ^{13}C 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 $=\text{CH}_2$
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 ^{13}C 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 glycyrrhetic 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*