

5.1 The spectral workflow: what each technique answers

Structural characterization asks four questions, and the syllabus's four techniques answer one each, which is why they are taught as a set:

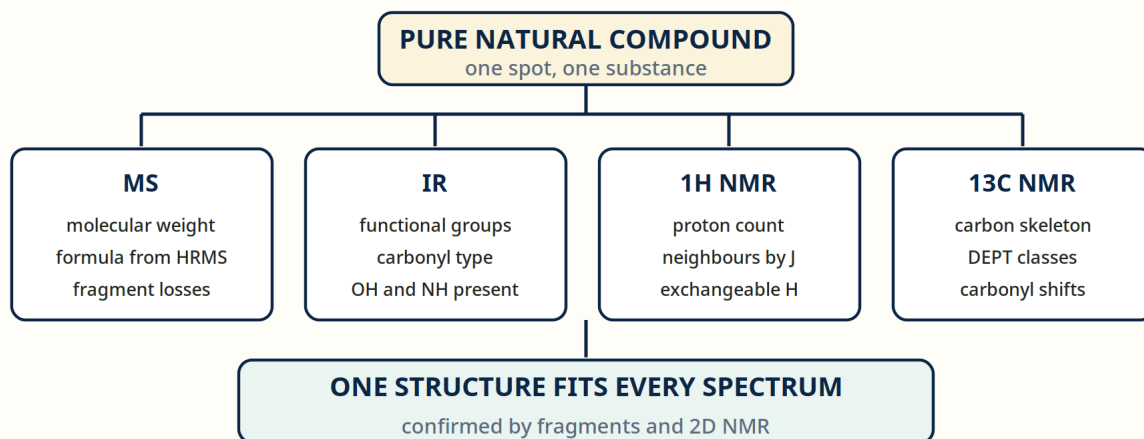
Technique	Physical basis	Question it answers first
Mass spectrometry (MS)	Ionisation and separation by mass to charge ratio	What is the molecular weight, and (high resolution) the molecular formula? What fragments does it break into?
IR spectroscopy	Absorption by vibrating bonds	Which functional groups are present: hydroxyl, carbonyl and of what kind, aromatic ring, alkene, ether?
¹ H NMR	Proton environments, integration, splitting and coupling	How many hydrogens of each kind, and what is next to what? Where are the methyls, the exchangeable protons, the olefinic and aromatic protons?
¹³ C NMR	Carbon environments (with DEPT editing)	What is the carbon skeleton: how many carbons of each type, CH ₃ , CH ₂ , CH and quaternary; carbonyls by their low field shifts

The professional workflow runs: purity first (a spectrum of a mixture characterises nothing), then HRMS for formula and unsaturation, IR for the functional group inventory, NMR for the skeleton assembled proton by proton and carbon by carbon, with 2D experiments (COSY for proton neighbours, HSQC and HMBC for carbon proton connections) as the modern assembly tools, and MS fragmentation as the cross check that the assembled structure breaks the way the real molecule breaks:

STRUCTURAL CHARACTERIZATION WORKFLOW

```
graph TD
    A["Pure compound  
(chromatography, one spot, sharp behaviour)"] --> B["HRMS: molecular weight and formula  
(formula, double bond equivalents)"]
    B --> C["IR: functional group inventory  
(OH or NH, carbonyl type, aromatic, alkene, ether, lactone)"]
    C --> D["1H NMR: proton map  
(count, environment, neighbours, J values)"]
    D --> E["13C NMR and DEPT: carbon skeleton  
(every carbon classified and placed)"]
    E --> F["Assemble draft  
(consistent with formula and all spectra)"]
    F --> G["Confirm: MS fragments, 2D NMR, chemical tests,  
comparison with authentic data, synthesis or X-ray"]
```

Four Spectra, Four Questions



© 2026 PharmacoLearn, pharmacolearn.fyi

Figure: The four syllabus techniques answer four different questions; only a structure consistent with all four spectra is accepted.

5.2 Functional group frequencies and shifts to have by heart

Observation	Interpretation
IR 3200 to 3600 cm ⁻¹ , broad	Hydrogen bonded O-H or N-H stretch
IR around 1740 to 1780 cm ⁻¹	Ester or strained lactam/lactone carbonyl; the higher the ring strain, the higher the frequency
IR around 1715 cm ⁻¹	Ordinary ketone; around 1680 cm ⁻¹ if conjugated
IR 1600, 1580, 1500 cm ⁻¹ bands	Aromatic ring vibrations
1H NMR 0.6 to 1.0 ppm singlets	Angular methyls on a steroid skeleton
1H NMR 3.2 to 4.0 ppm	Protons on oxygen bearing carbon (OCH, OCH ₃ a sharp singlet near 3.3 to 3.9)
1H NMR 4.5 to 6.5 ppm	Olefinic protons; 6.5 to 8.5 ppm aromatic protons; exchangeable OH and NH broad, disappearing on D ₂ O shake
13C NMR 170 to 220 ppm	Carbonyl carbons: ester and lactone near 170 to 180, ketone near 200 to 215
MS: loss of 18, 28, 44	Loss of water, carbon monoxide (or ethylene), carbon dioxide: first reading of a fragmentation

Deuterium exchange deserves its own sentence: shaking the NMR sample with D₂O makes OH, NH and COOH proton signals vanish, the cheapest experiment in spectroscopy and the standard way to count exchangeable protons in a natural product full of hydroxyls.

5.3 Penicillin

The problem historically: penicillin's structure was contested through the 1940s, with the fused beta-lactam thiazolidine structure competing against an oxazolone alternative; chemical degradation, IR and Dorothy Hodgkin's X-ray crystallography settled it for the strained four membered lactam fused cis to the thiazolidine ring, bearing a carboxyl group and the variable acyl side chain (benzyl in penicillin G).

Spectroscopic signature today:

- **IR:** the beta-lactam carbonyl absorbs unusually high, around 1760 to 1780 cm⁻¹, because the four membered ring prevents normal amide resonance: the carbonyl behaves almost like a ketone's. Beside

it sits the side chain amide carbonyl (around 1660 to 1680 cm^{-1}) and the carboxylic acid or carboxylate bands. No other common drug class shows a carbonyl that high, so the band is practically diagnostic.

- **^1H NMR:** the two bridgehead protons of the fused system (H-5 and H-6) appear as characteristic doublets around 5.4 to 5.6 ppm with a small cis coupling (about 4 Hz) that itself reports the cis ring fusion; the gem dimethyl group of the thiazolidine ring gives two singlets near 1.5 ppm; the side chain benzyl CH_2 and aromatic protons identify penicillin G's R group.
- **^{13}C NMR:** three carbonyl carbons in a row tell the story: lactam near 173 to 175 ppm, side chain amide near 170 ppm, carboxyl near 175 ppm, with the quaternary thiazolidine carbon near 65 ppm.
- **MS:** molecular ion (penicillin G, 334) with fragmentation dominated by cleavage around the side chain amide and the thiazolidine ring; the high resolution measurement gives $\text{C}_{16}\text{H}_{18}\text{N}_2\text{O}_4\text{S}$, the single sulfur itself confirmed by the characteristic $\text{M}+2$ isotope peak of about 4.4 percent.

5.4 Morphine

IR: a broad phenolic and alcoholic O-H envelope around 3200 to 3500 cm^{-1} , aromatic bands near 1600 and 1500 cm^{-1} , aromatic C-H out of plane bending consistent with a 1,2,4-trisubstituted ring, strong C-O stretching between 1000 and 1250 cm^{-1} for the ether bridge and the two hydroxyls, and significantly **no carbonyl band**, which at once excludes the oxidised relative morphinone.

^1H NMR: two aromatic protons as an AB pair of doublets (J about 8 Hz) near 6.5 to 6.7 ppm, the isolated olefinic protons of the ring C double bond near 5.3 to 5.7 ppm, the N-methyl a sharp singlet near 2.4 ppm, and a crowded aliphatic region whose coupling network maps the rigid bridged skeleton; the phenolic and alcoholic protons exchange with D_2O . Coupling constants around the allylic alcohol fix the configuration at C-6 relative to the ring system, spectroscopy doing the stereochemistry that once needed chemical correlation (section 2.6).

^{13}C NMR and MS: seventeen carbons, six aromatic (two oxygen bearing near 143 to 146 ppm), two olefinic, the ether bridge carbon (C-5) characteristically downfield near 90 ppm, and the N-methyl near 43 ppm. Electron impact MS gives M^+ at 285 ($\text{C}_{17}\text{H}_{19}\text{NO}_3$ by high resolution), with the morphinan skeleton fragmenting reluctantly: a strong molecular ion is itself characteristic of these rigid polycyclic bases, in contrast to floppy molecules that shatter.

5.5 Camphor

Camphor is the cleanest teaching example of the set because it is small, rigid and symmetrical in nothing but its formula.

IR: the ketone carbonyl absorbs around 1740 cm^{-1} , high for a saturated ketone (normally near 1715), because the carbonyl sits in the strained bicyclic framework; no O-H band and no C=C absorption, consistent with a saturated ketone.

^1H NMR: three methyl singlets (the bridgehead methyl and the gem dimethyl pair, the latter two slightly different in environment) between about 0.8 and 1.0 ppm account for nine of the sixteen hydrogens without any splitting: three methyls on quaternary carbons, a fingerprint of the bornane skeleton; the remaining signals are the methylene and bridgehead protons of the bicyclic frame, the proton alpha to the carbonyl appearing downfield near 2.3 ppm.

^{13}C NMR and MS: ten carbons with the ketone carbonyl far downfield near 218 ppm and the oxygen bridgehead quaternary carbon near 57 ppm. MS gives M^+ at 152 ($\text{C}_{10}\text{H}_{16}\text{O}$); alpha cleavage beside the carbonyl and loss of carbon monoxide (152 to 124) and hydrocarbon fragments such as m/z 95 and 81 trace the ring opening of the bornane frame. The classical degradative proof (section 3.3) and this spectral

reading arrive at the same bicyclo[2.2.1] structure by independent roads.

5.6 Vitamin D

Vitamin D's spectra are dominated by one feature: the **conjugated triene** created when the B ring of the sterol opens (section 2.14).

UV and IR: the triene gives intense absorption with a maximum near 265 nm, historically the basis of vitamin D assay, and its disappearance on irradiation or oxidation monitors degradation. IR shows the hydroxyl stretch, and characteristic triene C=C and out of plane C-H bands in the 1600 to 1650 and 880 to 900 cm⁻¹ regions (the exocyclic methylene of the triene system).

¹H NMR: the olefinic region is the signature: the exocyclic methylene protons as two broad signals near 4.8 and 5.0 ppm, and the triene protons between 5.9 and 6.3 ppm with coupling patterns that distinguish the vitamin from its inactive isomers (the 5,6-trans isomer and tachysterol differ exactly here, which is why NMR polices vitamin D preparations). The angular methyl singlet near 0.55 ppm and the side chain methyl doublets place the molecule in the sterol family the triene came from.

¹³C NMR and MS: twenty seven carbons for D3 (twenty eight for D2, whose side chain carries an extra methyl and double bond); the triene carbons spread between 112 and 145 ppm. Molecular weights 384 (D3, C₂₇H₄₄O) and 396 (D2, C₂₈H₄₄O) by MS, with ready loss of water (M minus 18) from the allylic alcohol. The spectral lesson: vitamin D is identified less by any single peak than by the combination of a sterol carbon count with a triene proton and UV pattern.

5.7 Quercetin

Quercetin (section 2.11) is where the classical and modern methods are most profitably compared, because its classical UV shift reagent method remains standard practice for flavonoids.

UV with shift reagents: in methanol, quercetin shows Band II near 255 nm and Band I near 370 nm, the long wavelength band reflecting the flavonol 3-hydroxyl conjugation. Sodium methoxide pushes Band I strongly to longer wavelength (a free 4 prime hydroxyl); sodium acetate shifts Band II (a free 7 hydroxyl); aluminium chloride shifts Band I and the shift largely survives added acid, reporting the 5-hydroxy 4-keto chelation system and the ortho dihydroxyl B ring. In twenty minutes this battery locates free hydroxyls that methylation analysis once took weeks to place.

IR and NMR: IR shows the hydrogen bonded carbonyl of the flavonol near 1660 cm⁻¹ (lowered by conjugation and chelation to the 5-hydroxyl), aromatic bands and a broad phenolic envelope. In ¹H NMR (DMSO-d₆) the chelated 5-OH appears far downfield near 12.5 ppm, a signal so characteristic it is practically a name tag; the A ring protons appear as two meta coupled doublets near 6.2 and 6.4 ppm; the B ring shows the 3 prime, 4 prime dihydroxy pattern (a doublet near 6.9, a doublet near 7.5 and a double doublet near 7.6 ppm). ¹³C NMR gives fifteen carbons with the carbonyl (C-4) near 176 ppm and the oxygenated C-2 near 147 ppm. MS gives M⁺ at 302 (C₁₅H₁₀O₇), with retro Diels-Alder cleavage of the C ring producing the diagnostic A ring fragment at m/z 153 that locates the A ring dihydroxylation.

5.8 Digitalis glycosides

The digitalis glycosides (digoxin, digitoxin; section 2.12) are the hardest of the six because they are three part molecules: a steroidal genin, an unsaturated lactone and a chain of unusual sugars (digitoxose, a 2,6-dideoxy sugar).

IR: the butenolide lactone carbonyl absorbs near 1740 to 1780 cm⁻¹ with its C=C band near 1620 cm⁻¹; hydroxyls give the broad high region band; and the absence of aromatic bands separates these steroids

from the oestrane family at a glance.

1H NMR: recognition signals: the lactone vinyl proton near 5.9 ppm; the 21 methylene of the lactone as a pair of doublets of doublets near 4.6 to 4.9 ppm; two angular methyl singlets (C-18 near 0.85, C-19 near 0.95 ppm); anomeric protons of the sugar chain between 4.8 and 5.2 ppm; and a signature that belongs to the digitoxose sugars alone, their 6-methyl **doublets** near 1.2 ppm, one per sugar unit, so counting them counts the sugars.

13C NMR and MS: the lactone carbonyl appears near 174 ppm and the olefinic lactone carbons near 117 and 174 ppm; the anomeric carbons sit between 95 and 100 ppm. Digoxin (C₄₁H₆₄O₁₄, mass 780) and digitoxin (C₄₁H₆₄O₁₃, mass 764) differ by one oxygen, the 12 beta hydroxyl of digoxin, which MS states in the molecular weight alone. Soft ionisation MS (ESI, giving [M+Na]⁺ and [M+NH₄]⁺ rather than a naked molecular ion) shows stepwise losses of 130 mass units, one digitoxose at a time, reading the sugar sequence off the spectrum; electron impact, by contrast, destroys these molecules, which is itself a teaching point: the ionisation method must be chosen for the fragility of the natural product, and glycosides demand soft ionisation.

Compound	One line spectral identity
Penicillin	IR carbonyl near 1770 cm ⁻¹ (strained lactam) plus sulfur isotope pattern in MS
Morphine	No carbonyl; phenolic OH; strong M ⁺ at 285; ether carbon near 90 ppm in 13C
Camphor	Ketone at 1740 cm ⁻¹ and near 218 ppm; three methyl singlets; M ⁺ 152
Vitamin D	Triene UV at 265 nm and olefinic 1H pattern; sterol methyls below 1 ppm
Quercetin	UV shift reagent battery; chelated 5-OH at 12.5 ppm; RDA fragment at m/z 153
Digitalis glycosides	Lactone vinyl proton near 5.9 ppm; anomeric signals; MS sugar losses of 130

Frequently Asked Questions

Q: What is a natural product lead, and why is the natural compound itself often not the drug? A lead is a natural compound whose activity shows where a drug can be built. The compound itself is usually flawed for medicine: paclitaxel was scarce and insoluble, thienamycin was unstable, podophyllotoxin was too toxic. Chemical modification, semi-synthesis and analog design repair those faults while keeping the activity the natural product revealed.

Q: How did teprotide lead to captopril if teprotide was never a drug? Teprotide, a peptide from pit viper venom, proved that inhibiting the angiotensin converting enzyme lowers blood pressure and showed how a peptide carboxylate and proline engage the enzyme's zinc site. Captopril was designed as a small, orally active molecule presenting the same binding features, so the natural product contributed the mechanism and the binding map rather than the structure.

Q: Why are alkaloids extracted by switching between acid and base? Because the two forms have opposite solubility. As salts in dilute acid, alkaloids dissolve in water and leave fat and resin behind; basified, the free bases move into organic solvents and leave salts and tannins behind. Switching twice purifies the bases from most of the plant matrix before chromatography or crystallisation begins.

Q: Morphine and ephedrine both come from plants. Why is only morphine called a true alkaloid in the strict sense? Morphine's nitrogen sits inside a ring system biosynthesised into a complex polycyclic alkaloid. Ephedrine's nitrogen is in an open side chain as a methylamino group; it belongs to the protoalkaloids, plant amines of amino acid origin whose nitrogen is not part of a heterocyclic ring. The

pharmacology overlaps, the structural class does not.

Q: What does the isoprene rule actually claim, and how did squalene confirm it? It claims terpenoid skeletons are assembled from C5 isoprene units through a few acyclic precursors, so structures must be derivable from the right precursor by cyclisation and rearrangement. Squalene, the C30 acyclic precursor, cyclises in the laboratory of the cell to lanosterol and then cholesterol exactly along the folding the rule predicted, confirming the scheme for the whole triterpene and steroid family.

Q: Quercetin has five hydroxyl groups. How do spectroscopists show where they are? Three tools combine: the UV shift reagent battery locates free hydroxyls at 7, 4 prime and the 5 position by the direction of band shifts; ¹H NMR shows the chelated 5-OH far downfield near 12.5 ppm and the aromatic proton patterns of both rings; and mass spectrometry's retro Diels-Alder fragment at m/z 153 reports the A ring hydroxylation. Together they fix 3,5,7,3 prime,4 prime.

Q: Why does vitamin D appear in both the steroid unit and the vitamins unit? Because it is both. Biosynthetically and structurally it is a secosteroid, a sterol whose B ring has opened, which places it with the steroids in Unit 2 of the syllabus; nutritionally and physiologically it is a vitamin whose hydroxylated metabolite is a calcium regulating hormone, so it is examined alongside the vitamins as well.

Q: What is the difference between a recombinant protein drug and a monoclonal antibody from hybridoma technology? A recombinant protein drug such as insulin is made by inserting the human gene into a production organism and purifying the protein it expresses. A monoclonal antibody comes from a hybridoma, a fused cell line that secretes one antibody of one specificity indefinitely; modern antibody drugs are then engineered and produced recombinantly as well, but their specificity was selected by the hybridoma route.

Q: Why does the beta-lactam carbonyl of penicillin absorb at such a high IR frequency? In an ordinary amide, resonance with nitrogen flattens and weakens the carbonyl, lowering its stretching frequency. The four membered beta-lactam ring twists the amide out of planarity and prevents that resonance, so the carbonyl behaves more like a ketone's and absorbs near 1760 to 1780 cm⁻¹. The same strain is what makes the ring reactive towards penicillin binding proteins.

Q: The indigenous diabetic drugs contain real active molecules. Why are they still described as adjuncts? Because constituent chemistry is only the first test. Gymnemic acids, salacinol and 4-hydroxyisoleucine have credible mechanisms, but human trial evidence for the crude drugs is variable and usually modest, preparations are not standardised like medicines, and diabetes carries risks that demand proven therapy. The professional position is study and adjunctive use with disclosure, never replacement of prescribed treatment.

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