

4.1 ADMET: why good binders fail as drugs

Ask why development candidates die and the answer, historically, was rarely potency. Compounds failed because they were not absorbed from the gut, were cleared in minutes, accumulated where they should not, or injured the liver and heart. **ADMET** names the five gates: **Absorption**, **Distribution**, **Metabolism**, **Excretion** and **Toxicity**. Predicting these properties in silico, the syllabus's first Unit 4 item, matters because of timing: a potency problem found in a screening assay costs an assay, an ADMET problem found in the clinic costs a programme. The modern discovery pipeline therefore predicts early and often, using ADMET models as filters in virtual screening and as design feedback in lead optimisation, and Unit 4 treats prediction and design together for exactly that reason.

4.2 The properties predicted, gate by gate

- **Absorption:** aqueous solubility, the gatekeeper property, and intestinal permeability, estimated from descriptors by models trained on cell assays such as Caco 2; log P and log D (Unit 1) feed both.
- **Distribution:** plasma protein binding, which sets the free fraction able to act; volume of distribution predictions; and blood brain barrier penetration models, wanted for CNS drugs and feared for others.
- **Metabolism:** prediction of metabolic stability, of the cytochrome P450 isoforms a compound inhibits or is cleared by, and of likely soft spots on the structure where enzymes attack; a predicted liability is a design instruction, block that position or soften it.
- **Excretion:** renal and biliary clearance tendencies, largely read from size, charge and lipophilicity.
- **Toxicity:** structural alerts for mutagenicity and reactivity, and models for specific mechanisms such as hERG potassium channel blockade, the classic cause of fatal cardiac arrhythmia surprises.

Every prediction carries an applicability domain, as in Unit 2: the model knows the chemistry it was trained on and guesses outside it.

4.3 Drug likeness rules: Lipinski and Veber

Two rule sets turn ADMET experience into numbers a chemist can apply at the drawing board. **Lipinski's rule of five** (1997), distilled from the property ranges of orally absorbed drugs, flags likely poor absorption or permeation when a compound has more than one of: molecular weight above 500, log P above 5, more than 5 hydrogen bond donors, more than 10 hydrogen bond acceptors. The rules are named for their multiples of five, are about oral exposure only, and carry known exceptions, notably transport substrates and most natural products; they are a warning light, not a law of nature. **Veber's rules** add shape and polarity: rotatable bonds of 10 or fewer and polar surface area of 140 square Angstrom or less predict good oral bioavailability across chemotypes. Both rule sets are routinely used as the first, cheapest filter in virtual screening (Unit 5), and both are exactly the kind of crude, useful model this subject teaches students to use and to doubt at the same time.

4.4 De novo design: growing a molecule inside the cavity

Ligand based methods remodel what exists; **de novo design** builds what does not. The programme is given the binding site, empty, and assembles a molecule inside it, atom by atom or fragment by fragment,

scoring as it grows. The design cycle runs:

```
Analyse the cavity: size, shape, interaction hot spots
|
v
Place seed fragments at the hot spots
|
v
Grow or link fragments into complete molecules
|
v
Score fit, strain and synthetic accessibility
|
v
Chemist reviews: can it be made? is it new?
|
v
Synthesise a shortlist and test
```

Early programs such as LUDI worked by placing fragments matching the site's hydrogen bond and hydrophobic hot spots and then linking them; later systems grow chains atom by atom under genetic or Monte Carlo search. The chronic weakness of the method is the same as its charm: the computer is a more imaginative chemist than the team, and many de novo structures are unsynthesisable, unstable or absurd, so scoring includes a synthetic accessibility estimate and the chemist's veto is structural, not decorative.

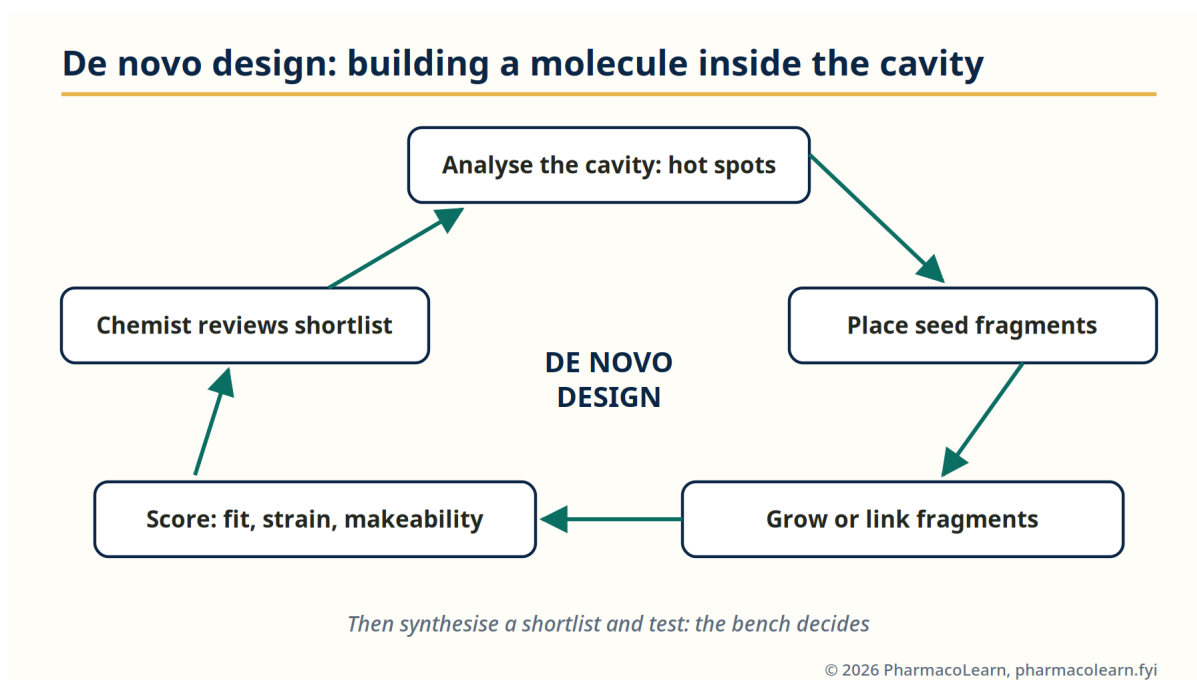


Figure: De novo design is a cycle inside the cavity, and the chemist's review step is what keeps its imagination synthesisable.

4.5 Reading a binding cavity: size and functional components

The syllabus breaks cavity analysis into three tasks, and a de novo or fragment project does them in order.

Interaction analysis maps where the site makes its demands: the hydrogen bond donors and acceptors of backbone and side chains, the charged residues, the hydrophobic walls, any metal centre or ordered water; protein ligand complexes of relatives and fragment soaking experiments reveal which of these are hot spots, positions where binding energy concentrates. **Cavity size prediction** measures the pocket's volume and shape from the protein surface, deciding whether the site can host a drug sized molecule at all, a druggability judgement, and which direction a growing molecule can extend. **Predicting the functional**

components of the cavity classifies each region by what a ligand should present to it: donor here, acceptor there, aromatic ring against that wall, positive charge nowhere near the acidic residue. The output is a blueprint, and sections 4.4 and 4.6 are two ways of building to it.

4.6 Fragment based drug design

Fragment based drug design (FBDD) starts from the observation that small fragments, molecular weight roughly 150 to 300, bind weakly but efficiently, and that nearly every pocket will bind some fragments at its hot spots. A fragment library of one or two thousand members is screened by sensitive methods, biophysical techniques such as NMR, surface plasmon resonance and X ray crystallography, rather than by potency assays, because fragment affinities are millimolar, not nanomolar. Crystallography then shows exactly where and how each fragment binds. Design grows the fragment, links two fragments binding side by side, or merges overlapping ones, keeping the efficient core and adding interactions, the path by which potency is built step by step. The **rule of three** guides library design, the fragment sized cousin of Lipinski's rules: molecular weight below 300, log P at or below 3, donors and acceptors at or below 3 each. The approach's textbook successes are real drugs: the BRAF inhibitor **vemurafenib** and the BCL 2 inhibitor **venetoclax** both began as fragment campaigns, proof that the syllabus item is a living industrial method and not a classroom curiosity.

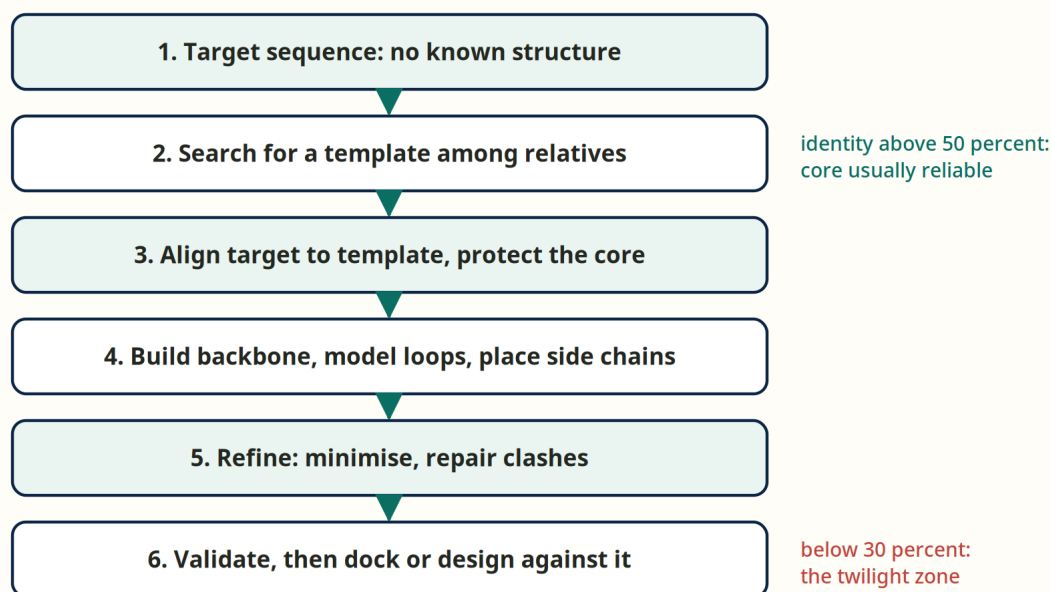
4.7 Homology modeling: a structure for a protein that has none

Docking and de novo design need a three dimensional target, and for many targets no experimental structure exists. **Homology modeling** builds one from a relative: if the target's amino acid sequence resembles that of a protein whose structure is known, the **template**, the target almost certainly folds the same way, because structure is more conserved than sequence. The workflow:

```
Target sequence
|
v
Template search (sequence alignment against known structures)
|
v
Align target to template, carefully, around conserved regions
|
v
Build the backbone on the template, model the variable loops
|
v
Place side chains, minimise, refine
|
v
Validate the model, then use it for docking or design
```

Reliability tracks sequence identity: above about 50 percent identity, models are generally trustworthy in the core; between 30 and 50 percent, usable with care; below about 30 percent, the twilight zone, where alignment errors dominate and models mislead as often as they guide. The conserved core inherits the template's backbone; insertions and deletions fall in surface loops, which are therefore the least reliable parts, and, inconveniently, often line binding sites.

Homology modeling: borrowing a relative's structure



© 2026 PharmacoLearn, pharmacolearn.fyi

Figure: Homology modeling borrows a relative's fold; reliability is set at the template step and proved at the validation step.

4.8 Validating the model before trusting it

A homology model is a hypothesis and is examined like one. The **Ramachandran plot** checks that backbone torsion angles fall in the sterically allowed regions where real proteins live; outliers mark strained, probably wrong, segments. Knowledge based energy and statistical profiles compare the model's residue environments with those of experimental structures, flagging buried charges and exposed hydrophobics that nature avoids. **Active site sanity** is the pragmatic test: residues known from mutagenesis or from family conservation to matter must land where the model needs them, and known ligands should dock into the model the way they dock into relatives. The deliverable of Unit 4, taken whole, is a judgement call this course trains deliberately: which computational evidence, ADMET prediction, de novo sketch, fragment hit or homology model, is strong enough to spend synthesis on.

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