Biochemistry

Unit 2 · Lesson 2 · Peptide bonds and protein structure

The notes for Lesson 2 of Unit 2: chapters 2.3 to 2.4 of the unit book, with their figures, worked examples, exam traps and quizzes, in one file. The full unit book, with every chapter, the exam question bank and the glossary, is linked at the end.

Lesson 2 of 4 | 2 chapters | Chapters 2.3 to 2.4

About these notes

These are the notes for **Lesson 2 · Peptide bonds and protein structure** of the University free n easy video lessons on Biochemistry Unit 2, Amino Acids and Proteins. They are chapters 2.3 to 2.4 of the unit book, unchanged, so a page here is the same page in the full book.

2.3 Peptide bond and peptides

A peptide bond is the amide bond that joins the carboxyl group of one amino acid to the amino group of the next, with the loss of one water molecule. Repeating it gives a chain called a peptide, and long chains are proteins. The peptide bond is not an ordinary single bond: it is flat and rigid, and that one fact limits how a protein chain can bend. Many small peptides are also important on their own, as hormones, antioxidants and sweeteners.

You need: amino acid structure (Chapter 2.1) and the reactions of the amino group (Chapter 2.2).

How the bond forms

The α-carboxyl of amino acid 1 reacts with the α-amino group of amino acid 2. An –OH leaves from the carboxyl and an –H from the amine; together they leave as H_2O. This is a **condensation** (dehydration) reaction, and the product –CO–NH– is an amide.



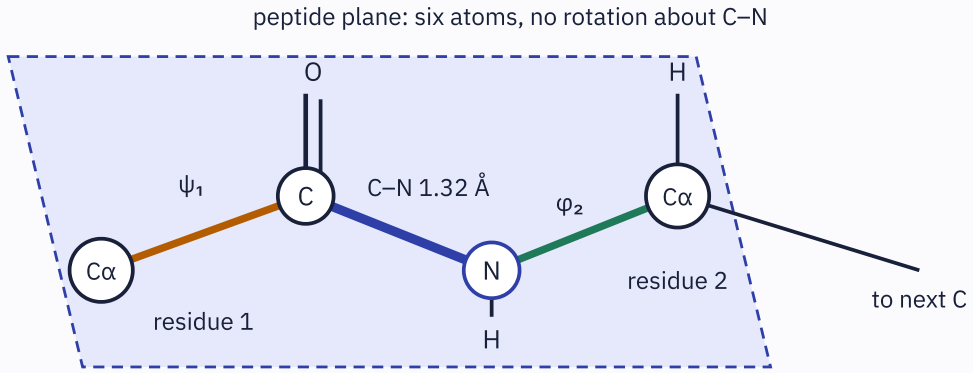
In water, the reverse reaction (hydrolysis) is favoured by free energy, but it is very slow without a catalyst, so proteins are stable for years. Cells do not make peptide bonds by simply mixing

amino acids. The ribosome joins amino acids that have first been activated by attachment to tRNA, a step that uses ATP. Breaking peptide bonds needs proteases, or boiling in strong acid (6 M HCl) for many hours, as in amino acid analysis.

The character of the peptide bond

The lone pair of electrons on the amide nitrogen is shared with the carbonyl group, so the bond is a resonance hybrid of two forms: C(=O)–N and C(–O⁻)=N⁺. This gives the C–N bond **partial double-bond character**. Its length, about 1.32 Å, lies between a C–N single bond (about 1.47 Å) and a C=N double bond (about 1.27 Å). Four results follow.

1. **Planar.** Six atoms lie in one plane: the Cα of residue 1, its C and O, the N and H of residue 2, and the Cα of residue 2. This flat unit is the **peptide plane** (amide plane).
2. **Rigid.** There is no free rotation about the C–N bond. The chain can rotate only about the two single bonds on either side of each Cα: N–Cα (angle φ, phi) and Cα–C (angle ψ, psi). These two angles are the subject of the Ramachandran plot (Chapter 2.5).
3. **Trans.** The two Cα atoms usually sit on opposite sides of the C–N bond (trans), which avoids a clash between side chains. The cis form is seen mainly in bonds to proline, whose ring makes cis and trans closer in energy.
4. **Polar but uncharged.** The C=O can accept a hydrogen bond and the N–H can donate one. These backbone hydrogen bonds build the α-helix and β-sheet (Chapter 2.4). Unlike free amino acids, the peptide bond itself does not ionise between pH 0 and 14.



Orange: Cα–C bond, rotates (angle ψ). Green: N–Cα bond, rotates (angle φ).
Blue: peptide C–N bond, partial double bond, fixed. Cα atoms are trans.

Figure 2.3.1 The peptide unit. The six atoms from one Cα to the next lie in a flat plane because the C–N bond is partly double. Rotation happens only at the bonds that touch each Cα, so a protein chain is like a row of flat cards hinged at the α-carbons.

Naming and writing peptides

A chain is always written from the **N-terminus** (free α-amino group) on the left to the **C-terminus** (free α-carboxyl) on the right. Each amino acid in a chain is a **residue**, because it has lost the parts of water. Name a peptide by changing the ending of every residue except the last to "-yl": Gly-Ala-Ser is glycylalanylserine. Two residues make a dipeptide, a few make an oligopeptide, and many make a polypeptide; chains of more than about 50 residues are usually called proteins, though the line is not sharp. With 20 choices at each position, there are 20^n possible sequences of length n: 400 dipeptides and 8,000 tripeptides.

A chain of n residues has n − 1 peptide bonds. Its charge comes from the two ends and from ionisable side chains only.

Biologically important peptides

PEPTIDE	SIZE AND MAKE-UP	ROLE
Glutathione	3 residues: γ-Glu-Cys-Gly; the Glu is linked through its side-chain (γ) carboxyl	Main antioxidant in cells; keeps proteins' –SH groups reduced
Oxytocin, vasopressin	9 residues each, with a disulfide ring; they differ at two positions	Hormones of the posterior pituitary: uterine contraction and milk release; water retention by the kidney
Insulin	51 residues in two chains (A 21, B 30) held by two disulfide bridges, with a third inside chain A	Lowers blood glucose; first protein to be sequenced (Sanger)
Enkephalins	5 residues (Met-enkephalin: Tyr-Gly-Gly-Phe-Met)	Natural pain-relieving peptides of the brain
Aspartame	Asp-Phe methyl ester (a dipeptide)	Artificial sweetener; labels warn people with phenylketonuria

Finding the ends and making peptides in the lab

End-group analysis. Sanger's reagent (1-fluoro-2,4-dinitrobenzene) labels the free N-terminal amino group; after acid hydrolysis, the yellow DNP-amino acid shows which residue was first. Edman degradation uses phenyl isothiocyanate to remove the N-terminal residue one at a time without breaking the rest of the chain, so repeated cycles read the sequence.

To join a chosen carboxyl to a chosen amine, every other reactive group must be blocked by a **protecting group**. The α-amino group is protected by Boc (removed by trifluoroacetic acid) or Fmoc (removed by the base piperidine). Using groups that come off under different conditions, so one can be removed while the others stay, is called **orthogonal protection**. The carboxyl is

activated by a **coupling reagent**: a carbodiimide such as DCC or DIC, often with HOBt added to reduce racemisation, or a phosphonium reagent such as PyBOP.

In **solid-phase peptide synthesis** (SPPS), introduced by R. B. Merrifield, the C-terminal residue is fixed to an insoluble resin bead. Each cycle removes the amino protection, couples the next protected amino acid and washes away excess reagent, so no purification is needed between steps. At the end, acid cleaves the finished peptide from the resin. Chemical synthesis builds the chain from the C-terminus to the N-terminus, the opposite of the ribosome.

WORKED EXAMPLE

Question (6 marks). For the tripeptide Gly-Ala-Ser: (a) name it and mark its N- and C-terminal residues; (b) how many peptide bonds does it have? (c) Calculate its molar mass, given Gly 75.07, Ala 89.09, Ser 105.09 and H_2O 18.02 g/mol. (d) What is its net charge at pH 7? (e) How many different tripeptides can be made using Gly, Ala and Ser once each?

1. Name: glycylalanylserine. N-terminal residue Gly, C-terminal residue Ser.
2. Peptide bonds = n − 1 = 3 − 1 = 2.
3. Sum of the free amino acids = 75.07 + 89.09 + 105.09 = 269.25 g/mol. Two bonds release 2 × 18.02 = 36.04 g/mol of water. Molar mass = 269.25 − 36.04 = **233.21 g/mol**.
4. At pH 7: N-terminal $-NH_3^+$ (+1), C-terminal $-COO^-$ (−1), and no ionisable side chains (Ser's –OH does not ionise). Net charge **0**.
5. Three different residues in three positions: 3! = 3 × 2 × 1 = **6** tripeptides (GAS, GSA, AGS, ASG, SGA, SAG). They have the same mass but different properties.

EXAM TRAPS

- A peptide bond is between the α-carboxyl and the α-amino group. Glutathione's Glu–Cys link uses the γ-carboxyl, so it is not a normal α-peptide bond; examiners ask about this.
- Do not say the peptide bond rotates freely. Rotation is about the N–Cα and Cα–C bonds, not the C–N bond.
- Gly-Ala and Ala-Gly are different dipeptides. The N-terminus is always written first.
- When calculating the mass of a peptide, subtract one water per peptide bond, not per residue.

CHECK YOURSELF

– REMEMBER **Which two groups react to form a peptide bond, and what is released?** 2 marks

The α-carboxyl group of one amino acid and the α-amino group of the next; one molecule of water is released.

– REMEMBER **Name two biologically active peptides and their functions.** 2 marks

Glutathione, an antioxidant; oxytocin, which causes uterine contraction and milk release. (Vasopressin, insulin or enkephalin are also correct.)

– UNDERSTAND **Why is the peptide bond planar?** 3 marks

Resonance shares the nitrogen lone pair with the carbonyl, giving the C–N bond partial double-bond character. Like a double bond, it cannot rotate, so the six atoms of the peptide unit stay in one plane.

– APPLY **How many peptide bonds and how many possible sequences are there for a pentapeptide made from the 20 standard amino acids?** 2 marks

4 peptide bonds; $20^5 = 3{,}200{,}000$ possible sequences.

– APPLY **Find the net charge of Lys-Gly-Asp at pH 7.** 3 marks

N-terminal NH_3^+ +1, Lys side chain +1, Asp side chain −1, C-terminal COO^- −1. Net 0.

– ANALYZE **Why must the side-chain amino group of lysine be protected during peptide synthesis, and why are Fmoc and a tBu-type side-chain group called orthogonal?** 4 marks

Otherwise the coupling reagent would join the next amino acid to the side-chain amine as well, giving a branched product. Fmoc comes off with base (piperidine) while acid-labile side-chain groups stay; these come off later with acid (TFA). Because each is removed under conditions that leave the other intact, they are orthogonal.

BEYOND THE SYLLABUS

Peptide drugs are a fast-growing class of medicines; semaglutide, a modified 31-residue peptide used for diabetes and weight loss, is one example. Automated SPPS machines now make peptides of 50 or more residues, and longer proteins are built by joining synthetic pieces with native chemical ligation. Mass spectrometry has largely replaced Edman degradation for sequencing. Search "native chemical ligation" and "peptide therapeutics". This is not for the exam.

2.4 Protein structure: primary, secondary, tertiary, quaternary

A protein works only when its chain is folded into the right shape, and that shape is described at four levels. Primary structure is the sequence of amino acids. Secondary structure is the local folding of the backbone into helices and sheets. Tertiary structure is the full three-dimensional fold of one chain, and quaternary structure is how several folded chains fit together. Each level is held by particular bonds, and each is decided, in the end, by the sequence. This is the one topic every university in this unit teaches.

You need: side-chain groups (Chapter 2.1) and the planar peptide bond (Chapter 2.3).

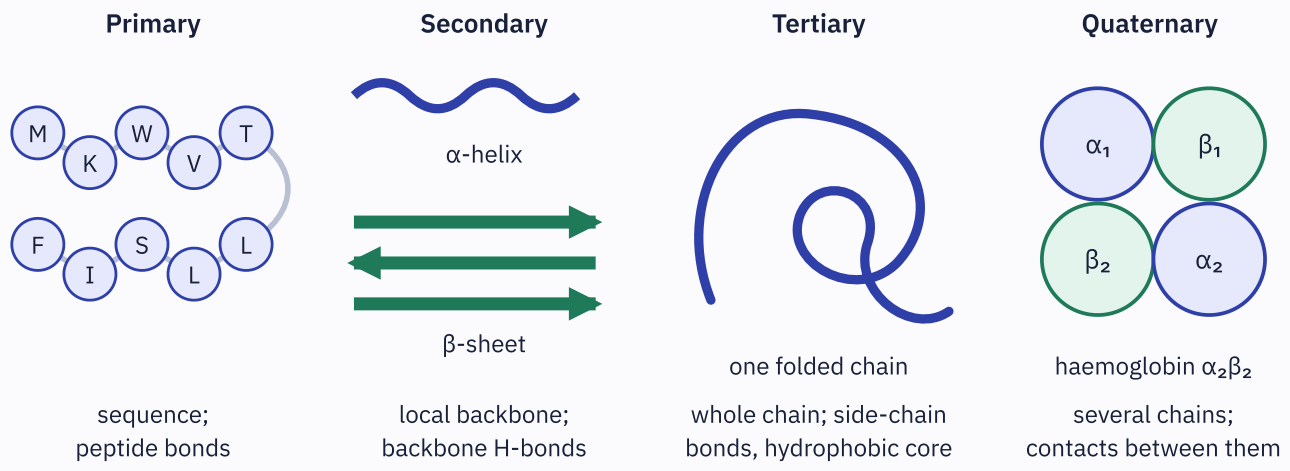


Figure 2.4.1 The four levels of protein structure and the bonds that hold each. Notice that secondary structure is held by hydrogen bonds between backbone atoms only, while tertiary and quaternary structure depend on side chains.

Primary structure and the bonds that hold proteins

The **primary structure** is the order of amino acids from N- to C-terminus, plus the positions of any disulfide bridges. It is coded by the gene, and it contains all the information needed for the higher levels (Chapter 2.6 shows the proof). One change can matter: in sickle-cell haemoglobin, glutamate at position 6 of the β chain is replaced by valine. The new hydrophobic patch makes deoxygenated haemoglobin molecules stick together into fibres that bend red cells into a sickle shape.

BOND	BETWEEN	MAIN LEVEL
Peptide bond (covalent)	α-COOH and α-NH_2 of adjacent residues	Primary
Disulfide bridge (covalent)	–SH groups of two cysteines	Tertiary, sometimes quaternary
Hydrogen bond	Backbone C=O and N–H; polar side chains	Secondary and tertiary
Ionic bond (salt bridge)	Opposite charges, such as Lys and Asp	Tertiary and quaternary
Hydrophobic interaction	Nonpolar side chains clustering away from water	Tertiary and quaternary (the main driving force of folding)
Van der Waals forces	All atoms in close contact	All levels, weak but numerous

Secondary structure

Secondary structure is a regular, repeating fold of the backbone, held by hydrogen bonds between C=O and N–H groups of the backbone. Linus Pauling and Robert Corey predicted the two main kinds in 1951 by asking which shapes keep every peptide bond planar and every C=O paired with an N–H.

α-helix. The chain winds into a right-handed spiral. There are 3.6 residues per turn, each residue rises 1.5 Å along the axis, and one turn (the pitch) is 3.6 × 1.5 = 5.4 Å. The C=O of residue n bonds to the N–H of residue n + 4, so every backbone group in the helix is hydrogen-bonded. Side chains point outward. Proline breaks a helix because its nitrogen has no H to donate and its ring cannot fit the turn; glycine is too flexible; a run of like charges (several Glu or Lys) repels itself. Keratin of hair is mainly α-helix.

β-sheet. The chain is stretched almost fully, about 3.3 to 3.5 Å per residue, into a **β-strand**. Two or more strands lie side by side and hydrogen-bond to each other, forming a pleated sheet. Strands running the same direction (N→C) form a **parallel** sheet; alternating directions form an

antiparallel sheet, which has straighter, stronger hydrogen bonds. Side chains point alternately above and below the sheet. Silk fibroin is mostly antiparallel β-sheet rich in small Gly and Ala, which lets the sheets stack closely.

β-turn. A tight turn of four residues, where the C=O of residue 1 bonds to the N–H of residue 4, reverses the chain's direction between strands. Gly and Pro are common in turns.

Collagen triple helix. Collagen, the main protein of skin, tendon and bone, has a repeating Gly-X-Y sequence, where X is often proline and Y often 4-hydroxyproline. Three left-handed chains twist round each other into a right-handed superhelix. Glycine is needed at every third position because only its tiny side chain fits in the crowded centre. Making hydroxyproline needs vitamin C, which is why its lack (scurvy) weakens gums and skin.

Each amino acid has a measured tendency, or propensity, to be found in each structure. In the **Chou–Fasman method**, Glu, Met, Ala and Leu are strong helix formers; Val, Ile, Tyr, Phe and Trp favour β-strands; Pro and Gly break both and are common in turns. A stretch of a few residues rich in helix formers is predicted to start a helix, which is then extended until breakers appear. Such methods are about 50 to 60 percent accurate; modern programs that use many related sequences and machine learning do much better.

Tertiary and quaternary structure

Tertiary structure is the complete three-dimensional arrangement of one polypeptide, including how its helices and sheets pack against each other. In water-soluble proteins, nonpolar side chains are buried in a core and polar ones cover the surface. Large proteins fold into separate compact units called **domains**, each often with its own job. **Fibrous proteins** (keratin, collagen, silk) have long, simple shapes and are insoluble; **globular proteins** (enzymes, myoglobin, antibodies) are compact and soluble (Chapter 2.7). Myoglobin, the oxygen store of muscle, was the first protein whose tertiary structure was solved by X-ray crystallography (John Kendrew); it is one chain of 153 residues, mostly α-helix, wrapped round a haem group.

Quaternary structure exists only in proteins with more than one chain (subunit). Haemoglobin has four: two α and two β chains ($\alpha_2\beta_2$), each with a haem. The subunits are held by hydrophobic contacts, hydrogen bonds and salt bridges. Because they touch, oxygen binding to one subunit changes the others and makes them bind oxygen more easily; this **cooperative binding** lets haemoglobin load oxygen in the lungs and release it in tissues. Myoglobin, with one chain, cannot do this.

WORKED EXAMPLE

Question (5 marks). A 36-residue segment of a protein forms an α-helix. (a) How many turns does it make? (b) How long is the helix? (c) If the same segment were a fully extended β-strand at 3.5 Å per residue, how long would it be? (d) Which form would fit across a cell membrane about 3 nm thick?

1. Turns = residues ÷ residues per turn = 36 ÷ 3.6 = **10 turns**.
2. Length = residues × rise per residue = 36 × 1.5 Å = 54 Å = **5.4 nm**. (Check: 10 turns × 5.4 Å pitch = 54 Å.)
3. β-strand length = 36 × 3.5 Å = 126 Å = **12.6 nm**, more than twice as long.
4. A membrane about 3 nm (30 Å) thick needs 30 ÷ 1.5 = 20 residues of α-helix. So a 20-residue helix of hydrophobic residues spans it; that is why membrane-crossing segments are usually helices of about 20 residues.

EXAM TRAPS

- Secondary structure uses backbone hydrogen bonds only. Do not say that side-chain bonds make the α-helix.
- Rise per residue (1.5 Å) and pitch (5.4 Å) are different numbers; pitch = 3.6 × 1.5.
- Quaternary structure needs more than one folded chain. Myoglobin (one chain) has none; haemoglobin (four chains) has it. Do not confuse it with tertiary structure, which is the fold of one chain.
- Disulfide bridges are covalent, so they survive urea; they are broken only by reducing agents (Chapter 2.6).
- The examiner's favourite: "Describe the α-helix with a neat diagram." Give residues per turn, rise, pitch, the n to n + 4 hydrogen bond, right-handedness and the role of proline.

CHECK YOURSELF

– REMEMBER **State the residues per turn, rise per residue and pitch of the α-helix.** 3 marks

3.6 residues per turn; 1.5 Å rise per residue; pitch 5.4 Å.

– REMEMBER **What is the subunit composition of adult haemoglobin?** 1 mark

Two α and two β chains, $\alpha_2\beta_2$, each carrying one haem.

– UNDERSTAND **Why is proline called a helix breaker?** 2 marks

Its nitrogen has no hydrogen to give for the n to n + 4 hydrogen bond, and its ring fixes the φ angle and clashes with the turn of the helix.

– UNDERSTAND **Compare parallel and antiparallel β-sheets.** 4 marks

In parallel sheets all strands run N→C in the same direction and hydrogen bonds are slanted; in antiparallel sheets neighbouring strands run in opposite directions and hydrogen bonds are straight, making the sheet more stable. Both have side chains alternating above and below.

– APPLY **How many residues are needed for an α-helix 4.5 nm long, and how many turns is that?** 2 marks

45 Å ÷ 1.5 Å = 30 residues; 30 ÷ 3.6 ≈ 8.3 turns.

– APPLY **Using Chou–Fasman propensities, which segment is more likely to be a helix: E-A-L-M-K-A or V-I-Y-T-V-G?** 2 marks

E-A-L-M-K-A, which is rich in the helix formers Glu, Ala, Leu and Met. V-I-Y-T-V-G is rich in β-formers and ends with a breaker.

– ANALYZE **Explain why replacing Glu6 by Val in the β chain of haemoglobin causes disease, though it changes only one residue of 146.** 5 marks

Position 6 is on the surface. Glu there is charged and faces water; Val is hydrophobic and creates a sticky patch. In deoxygenated haemoglobin this patch fits a hydrophobic pocket on a neighbouring molecule, so molecules polymerise into long fibres that distort red cells into sickles. A change in primary structure alters the surface and how molecules associate.

BEYOND THE SYLLABUS

Predicting the full three-dimensional structure from sequence was a 50-year problem until deep-learning systems such as AlphaFold reached near-experimental accuracy for many proteins around 2020. The AlphaFold Protein Structure Database now offers predicted structures for a very large number of proteins, free to all. Experimental structures are still solved by X-ray crystallography, NMR and, more and more, cryo-electron microscopy, and are stored in the Protein Data Bank. Search "AlphaFold database" and "Protein Data Bank". This is not for the exam.

Key terms of this lesson

TERM	MEANING OR FORMULA	CHAPTER
Peptide mass	Σ(amino acid masses) − (n − 1) × 18.02 g/mol for n residues	2.3
Peptide bond	Planar, rigid, trans; C–N about 1.32 Å; partial double bond	2.3
α-Helix	3.6 residues/turn, 1.5 Å rise, 5.4 Å pitch; H-bond n → n + 4; right-handed	2.4
β-Sheet	Extended strands, about 3.3–3.5 Å per residue; parallel or antiparallel	2.4

Glossary (terms used in this lesson)

Disulfide bridge A covalent –S–S– bond between two cysteine side chains.

Domain A compact, independently folded part of a larger protein.

Fibrous protein A long, insoluble, structural protein such as keratin or collagen.

Globular protein A compact, usually soluble protein such as an enzyme or haemoglobin.

Peptide bond The amide bond between the α-carboxyl of one amino acid and the α-amino group of the next.

Residue An amino acid unit within a peptide chain.

The full unit book

Every chapter of Unit 2, the unit summary, the exam question bank with model answers and the full glossary: Biochemistry · Unit 2 · Amino Acids and Proteins (one file, works offline).

Sources and further reading

Textbooks

- David L. Nelson and Michael M. Cox, *Lehninger Principles of Biochemistry*, W. H. Freeman; chapters on amino acids, peptides and proteins, and on protein structure (from memory; check). Source of the pKa values used here.

- Jeremy M. Berg, John L. Tymoczko, Gregory J. Gatto Jr. and Lubert Stryer, *Biochemistry*, W. H. Freeman; chapters on protein composition and structure and on exploring proteins (from memory; check).
- Donald Voet and Judith G. Voet, *Biochemistry*, Wiley (from memory; check).
- U. Satyanarayana and U. Chakrapani, *Biochemistry*, Elsevier India; widely prescribed in Indian universities for the classification of proteins and colour reactions (from memory; check).

Open resources opened for this book

- G. N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, "Stereochemistry of polypeptide chain configurations", *Journal of Molecular Biology* 7 (1963) 95–99, record at the Indian Academy of Sciences repository: repository.ias.ac.in/49564.
- C. B. Anfinsen, "Principles that govern the folding of protein chains", *Science* 181 (1973) 223–229 (the Nobel lecture): PDF copy. Source of the ribonuclease details in Chapter 2.6.
- Nobel Prize page for Anfinsen's lecture, 1972: nobelprize.org.

Syllabus: the University free n easy universal syllabus, built from top universities' syllabi.

Built for the University free n easy YouTube channel from the universal syllabus of Biochemistry (Oct 2026). Written with top universities' syllabi and standard textbooks; check a formula against your prescribed book before an exam, and report mistakes on the channel.