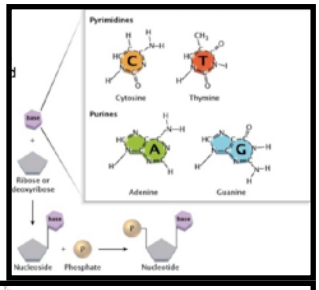
The Flow of Biological Information & Prebiotic Chemistry (W1)

The central dogma of molecular biology and DNA (L1)

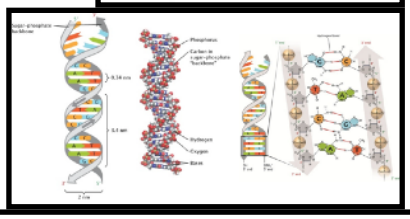
Bonding and Non-Bonding Interactions

- **The central dogma**: information flows in biology by the template-directed synthesis of:
  - DNA —> RNA —> proteins —> substrate —> products
- Each step in the central dogma results in amplification



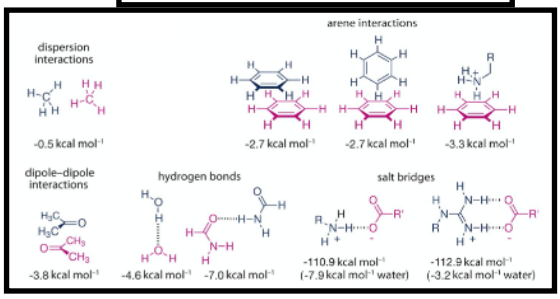
Discovery of DNA

- In 1881, Albrecht Kossel identified the chemical components of DNA
- In 1994, Oswald identified DNA as the material that contained genetic information
- In the late 1940s, Chagaff observed that amount of C-G and A-T were always the same (Foundation for base pairing)



Structure of DNA

- 1953 - Watson and Crick proposed the “double helix” structure of DNA, using models, along with data from Roslind Franklin
- 2nm diameter, 0.34nm between adjacent nucleotides, 3.4nm for one rotation



Non-bonding Interactions

- Generally weak, but **additive** (must take the sum of all energies)
- Under biological conditions, **covalent** bonds are **kinetically stable**
- Non-covalent bonds (**electrostatic** and **van der Waals**) are **kinetically unstable** and **reversible**
- e.g. Pi-stacking, ionic interactions, dipole-dipole interactions, H-bonding, dispersion

Hydrogen Bonding

- **Hydrogen bonding** is fundamental to life:
  - Hold water to the surface of the earth
  - The “glue” which unites DNA strands and holds proteins in defined conformations
- A hydrogen bond is a non-covalent bond involving three atoms
  1. A **donor atom** (electronegative, F,O,N)
  2. An **acceptor atom** (electronegative, F,O,N)
  3. A **proton** (bound to donor atom)
- Hydrogen bonds are mainly **coulombic interactions** and are strongest in low dielectric solvent, 2Å in length with 180 degrees bond angle

Hydrophobic Effect

- Water will **maximise** interactions with itself and **minimise** interactions with **hydrophobic** molecules - **hydrophobic effect**
- This **amplifies** the **energy** of **hydrophobic interactions** in water
- This is why **proteins** adopt a **globular shape** in water and why **hydrophobic ligands** bind to **hydrophobic** “patches” on proteins

Bonding and Non-bonding Interactions

1. **Bonding interactions**: lead to the formation of **covalent** bonds by mixing of **frontier molecular orbitals** (MOs)
2. **Non-bonding interactions**: involve non-covalent interactions

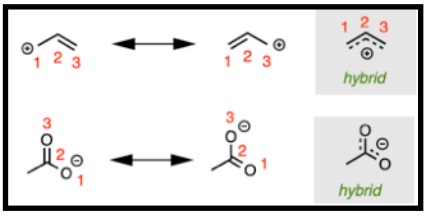
- Hydrogen bonding
- Electrostatic interactions between opposite charges (Coulombic interactions)
- Van der Waals interaction (pi-stacking/hydrophobic interactions)

Bonding Interactions - **Covalent** Bonds

- A bonding interaction between an **electrophile** and **nucleophile**
  - A **nucleophile donates** a **pair of electrons** from a **filled orbital**
  - An **electrophile accepts** them into a **vacant orbital**
  - A **new molecular orbital** is formed - a new bond.
  - Nucleophilicity: amine is more nucleophilic than oxygen because of electronegativity

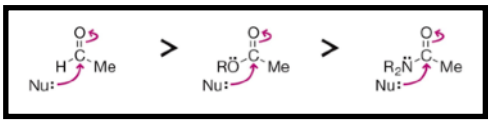
Resonance

- Two (or more) forms of the same molecule differing ONLY in the placement of electrons is called **resonance**
- The different structures are called **resonance forms**/contributors
- Resonance can be used to predict which atom in the molecule which will most likely be **protonated**



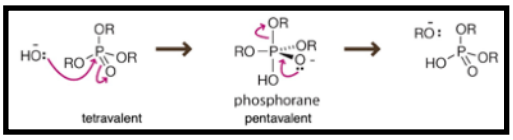
Bond Strength and Bioinformation

- Why are proteins made of **amide bonds**, not esters?
- Amides are more **kinetically stable** than esters, due to **electron delocalisation** of neighbouring **nitrogen atom** into **pi* MO of adjacent C=O**
- Carbon has an empty pi* orbital which electrons can move into
- Less likely to react with nucleophiles - it follows:
- Amides are more resistant than esters towards nucleophilic attack due to efficient delocalisation of lone e- pair on nitrogen into pi* of C=O



Phosphate Esters

- Phosphate esters **resists hydrolysis** (large activation barrier) - this is why DNA resists hydrolysis
- Via a **pentavalent phosphorane intermediate**



Phosphate esters are less reactive than carboxylic esters:

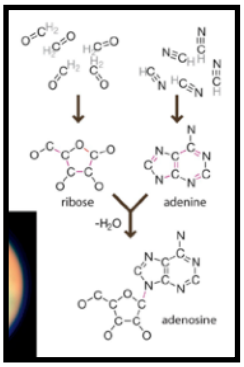
1. Large **energy barrier** for re-hybridisation for 3rd row elements
2. **Equatorial** electronegative substituents (horizontal)
3. Phosphate **oxygens repel incoming negatively charged** nucleophiles

- Reactivity: Phosphate esters < Carboxylic Esters

Modular Design

- The **stability** of a covalent bond linking biopolymers determines **information longevity**
- The **more stable** the bond, the **longer** the information storage of the biopolymer and therefore less copies need to be made
- Stability goes down as process of protein synthesis occurs
- Why **nature uses amide bonds**: **esters** are more **easily hydrolysed**. Partial double bond of amide gives rise to secondary and tertiary structures.

Functionality	Relevance	Half-life at pH 7 (years)
carboxylic ester	lipids	<1
carboxylic amide	peptides	300
ribose phosphate diester	RNA	2200
phosphate diester	DNA	220,000
β-glucofuranoside	RNA/DNA	22,000,000

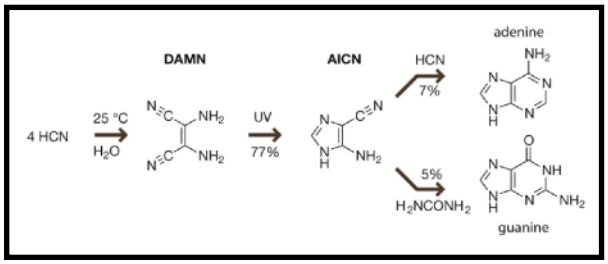


Prebiotic Chemistry (L2)

pH and pKa

- pH = log[H+]
- pKa = log[H+][A-]/[HA]
- These are related by the Henderson-Hasselbalch Equation
- pH = pKa + log[A-]/[HA]
- When pH = pKa, 50% of acid is deprotonated
- Smaller pKa means stronger acid
- H bond **accepts** - elements with free lone pair
- H-bond **donors** - X-H, where X is more eletronegative than H

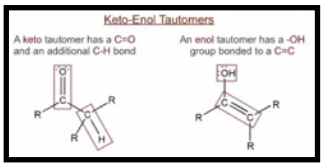
- If **pH < pKa**, then group will be **protonated**
- If **pH > pKa**, then group will be **deprotonated**

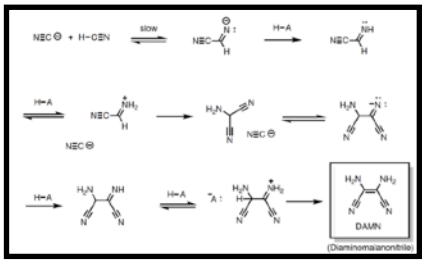


Prebiotic Chemistry
- How did life arise on primordial planet earth? Was RNA the original blueprint for life?
- The field of prebiotic chemistry is concerned with studying the biological chemistry of primordial earth before life
- Primordial earth was rich in **water**, ammonia, methane, **hydrogen cyanide** and **cyanoacetylene** (3.8-3.6 billion years ago)
- Formaldehyde —> Ribose
- **Cyanide** —> **Adenine**. **Urea** —> **Guanine**
- Ribose + Adenine —> Adenosine

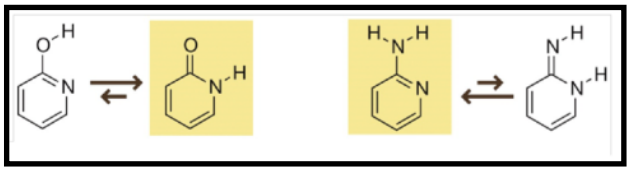
DNA vs RNA
- Note the chemical differences between DNA and RNA
- DNA is more stable, but RNA is easier to synthesise from pre-biotic chemicals
- DNA: Thymine. RNA: Uracil
- 2 hydroxyl groups on ribose sugar in RNA. 1 hydroxyl group on ribose sugar in DNA

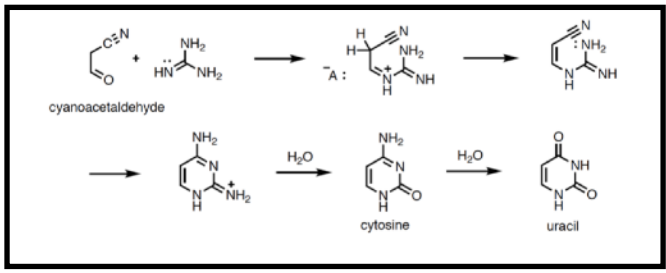
Prebiotic synthesis of purines
- **Purine nucleobases** (**adenine** and **guanine**) are derived from the oligomerisation of HCN - just add water and light
- Purine = 2 rings. Pyrimidines = 1 rings
- Cyanide exists in a 1:1 ratio of CN- and HCN at pH 9.2
- **Diaminomalanonitrile** (DAMN)





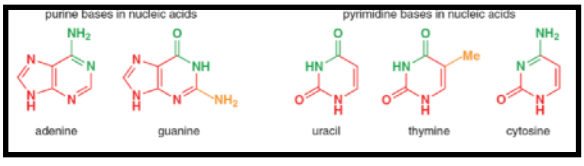
Tautomerism
- **Tautomers** are **constitutional isomers** - **migration** of **proton**
- Rapid interconversion by proton transfer: Tautomerisation (prototropic tautomerism)
- Acid or base catalysed
- **Keto-enol** tautomerism in nucleobases (G,T,C,U)
- Amides are usually favoured over the enolamine tautomer in cyclic form. Even though enolamine is aromatic, amide form is partially aromatic due to resonance

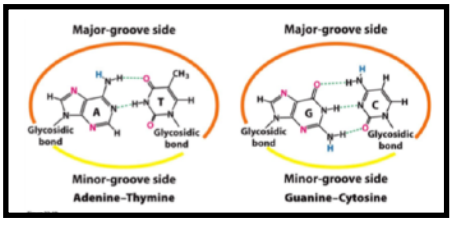




Prebiotic synthesis of pyrimidines
- **Pyrimidines** (**Cytosine** and **Uracil**) are derived from **cyanoacetylene**
- Cyanoacetaldehyde + Guanidine -> precursor of cytosine and uracil

Nucleotides
- **Nucleotides** are **phosphate monoesters** of ribose with unique heterocyclic base
- The phosphate is **deprotonated** at physiological pH
- DNA: 2'-deoxyribose and A,G,T,C
- RNA: Ribose and A,G,U,C
- **Nucleoside**: sugar + base
- **Nucleobase**: base

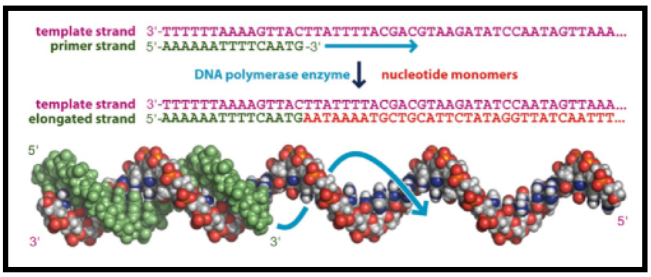




$$T_m = 2 \times (\text{no. of A-T bp}) + 4 \times (\text{no. of G-C bp})\ ^{o}C$$

Major and minor grooves
- Phosphoribose backbones are not oriented symmetrically from helical axis - leads to minor and major grooves
- G-C (3 H bonds) A-T (2 H bonds)

Replication of DNA
- **DNA Polymerase** enzyme η (DNA Pol) extends existing oligonucleotide strands
- Reactive **monomers** are **magnesium** complexes of **2'deoxynucleotidyl 5'-triphosphates**: dATP, dCTP, dGTP, dTTP
- Add at the **3' end** of **primer** strand
- DNA read in 5' —> 3' direction

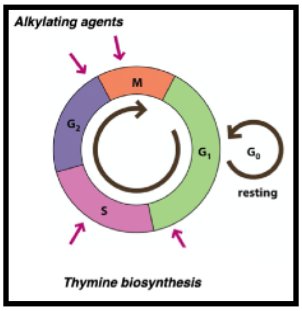


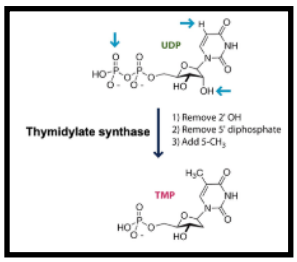
DNA sequencing
- If a polymerase incorporates a nucleotide **lacking the 3' OH group**, **chain termination** will occur - dideoxy DNA sequencing
- **Fluorescently** labeled chain terminating **dideoxynucleotides (ddN*)**
- Fluorescence resonance energy transfer (**FRET**)

Hybridisation and Melting
- The association of complimentary strands is spontaneous under physiological conditions - **hybridisation**
- The stability is measured by the melting point **(Tm)** - temperature at which **half** is double stranded and half is single stranded **(Wallace Rule)**
- Denatured has higher absorbance due to exposed nucleotides. Nucleotides are not stacked
- Absorb maximally at 260nm
- Higher G-C content —> High Tm as stronger interactions

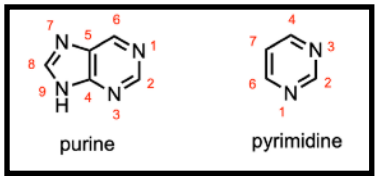
Intercalation
- **Pi-stacking** between **aromatic rings** of base pairs **stabilises** DNA double helix
- Many **cytotoxic drugs** (antitumor) and antibiotics are **DNA intercalators** (limit one per two BPs)

Cytotoxic drugs
- Tumor cells rapidly undergo cell division (mitosis)
- DNA of tumour cells can be targeted at various checkpoints to trigger cell death (apoptosis)
- Three ways of targeting DNA:
  1. **Thymine** biosynthesis
  2. DNA **replication**
  3. Inhibition of **mitosis** (alkylating agents)



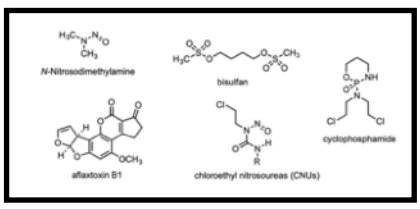


Thymine Biosynthesis
1. Inhibitors of Thymine biosynthesis
- Thymidylate synthase
- Steps 1) and 3) are targeted in chemotherapy
  - Remove 2' OH
  - Remove 5' diphosphate
  - Add 5-CH3

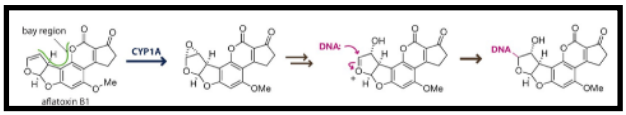


Nucleobases are nucleophiles
- Nucleobases are nucleophilic G(N7) > A(N3) >> T(N1) > C(N3)

DNA alkylators

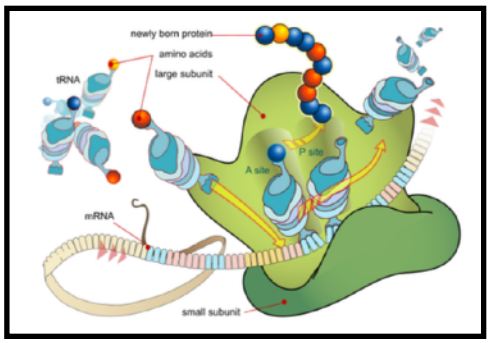


- Reacts as **electrophiles** to form stable covalent bonds (attack nucleophile)
- Frequent reactions lead to mutations then cancer
- Some are used in low concentrations to treat cancer
- **Aflatoxins** are natural products and belong to the **epoxide alkylation class**
- Are converted to a **highly electrophile epoxide** in the liver by **CYP1A**
- **Intercalates** and **reacts** with the **N7** nitrogen at **guanidylate** (1 phosphate, ribose and guanine) residues



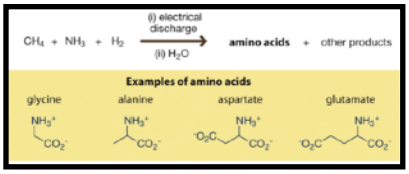
Amino Acids, Peptides and Proteins (W2)

The amide bond, peptides and proteins (L1)



Translation

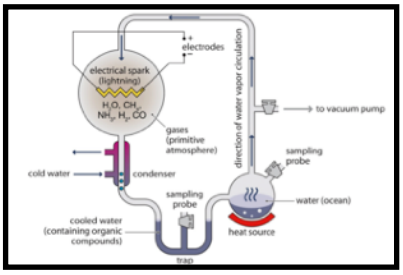
- mRNA is **translated** into protein by the **ribosome**
- The ribosome is composed of proteins and catalytic RNA (tRNAs)
- Each amnio acid is coded by a 3-base **codon**
- **tRNA** (transfer RNA) matches to the codons in **mRNA** (messenger RNA)
- E site: tRNA translocates. P site amino acid joins growing chain. A site tRNA leaves



Prebiotic synthesis of amino acids

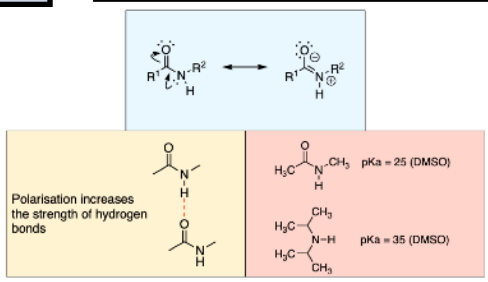
- Simple amino acids arise from **methane**, **ammonia** and **hydrogen** and an electrical discharge
- The Miller-Urey experiment

The Miller-Urey experiment:



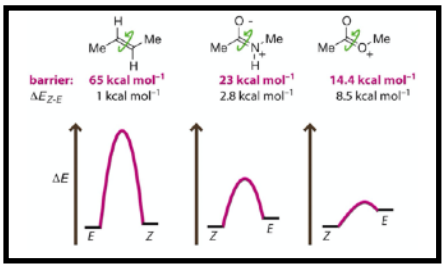
Prebiotic synthesis of amino acids

- Prebiotic amino acids biosynthesis explained by the **Strecker reaction**
- The reaction of an **aldehyde**, **HCN** and **ammonia**
  - Produces product with -CN and -NH2 on ketone C
  - Aqueous acid will be required to convert product into amino acid
- The amide bond:
  - Polarisation increases the strength of hydrogen bonds



Barriers to rotation about double bonds

- Alkene
  - Barrier: **65 kcal/mol**
  - deltaE(Z-E) = 1 kcal/mol
- Secondary amide with double bond
  - Barrier: **23 kcal/mol**
  - deltaE(Z-E) = 2.8 kcal/mol
- Ether with double bond
  - Barrier: **14.4 kcal/mol**
  - deltaE(Z-E) = 8.5 kcal/mol

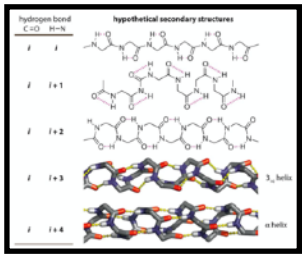


Allylic Strain (why amide bond is **stable**)

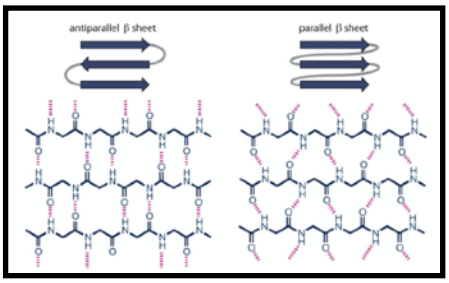
- Peptide bonds have **40% pi-bond character** due to **resonance -** partial bond character prevents rotation of bonds
- **Allylic strain** explains why proteins mainly adopt either a **beta-sheet** or **alpha-helix** secondary structure - partial bond allows this to happen
- Allylic strain limits backbone conformation

H-bonding and secondary (2 structure)

- **Alpha-helices** are formed by intramolecular H-bonding between **adjacent residues** (i, i+4)
- C=O groups in alpha helix pointing in same direction - protein is polarised

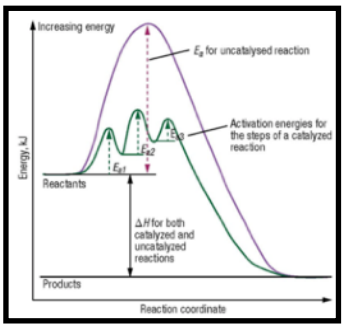


leading to macro dipole (+ large at N-terminus)
- Used by some proteins to bind negatively charged ions (eg. Sulphate binding protein PBD 1SBP)



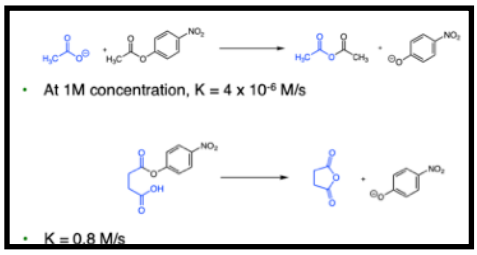
H-bonding and secondary (2 structure)
- **Beta-sheets** are formed by H-bonding between **backbone amide NHs** and **different peptide strands** - parallel and antiparallel
- **Antiparallel** are **most stable** - linear H-bonds with valine (V), isoleucine (I) and threonine (T) (most common)
- Examples: beta-sheet containing domains - beta-barrels, aggregation leading to Alzheimers

The amide bond, peptides and proteins (L2)
Roles of protein (an incomplete list)
- **Structural**
  - Give structure and rigidity to cells
  - Actin, Tubulin, keratin
- **Motor**
  - Allow cells to move
  - Actin/myosin, flagellin
- **Transport**
  - Move molecules through blood, across membranes etc.
  - Haemoglobin, albumin
- **Hormones**
  - Signalling
  - Insulin, thyroxine
- **Enzymes**
  - Catalyse chemical reactions
  - Lipase, trypsin

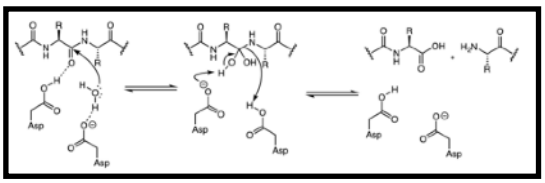


Enzymes - natures catalysts
- Enzymes lower that activation energy (Ea)
- This changes the rate of the reaction
- Enzymes do not change the position of an equilibrium

Enzymes - modes of catalysis
- Enzymes can catalyse chemical reactions in a range of ways. These can include (an incomplete list)
  - Position and orientation of molecules
  - Proton donors and acceptors
  - Metal ion catalysis
  - Covalent catalysis



Position and orientation of molecules
- Proximity effects
- First reaction: at 1M concentration, K= 4 x 10^-6 M/s
- Second reaction: K = 0.8 M/s
  - Tethering the reactants increases the rate by 200,000 fold



Proton donors and acceptors
- Amide hydrolysis is essential for numerous cellular processes
- Amide hydrolysis can be catalysed by acid or base
- Amide —> Carboxylic acid and Amine
- Aspartate proteases catalyse the hydrolysis of amide bonds

Metal ion catalysis (later)