Preparatory Problems for
the 58th International Chemistry Olympiad 2026
Tashkent, Uzbekistan

Theoretical Problems

Contents

Fields of advanced difficulty

Theoretical

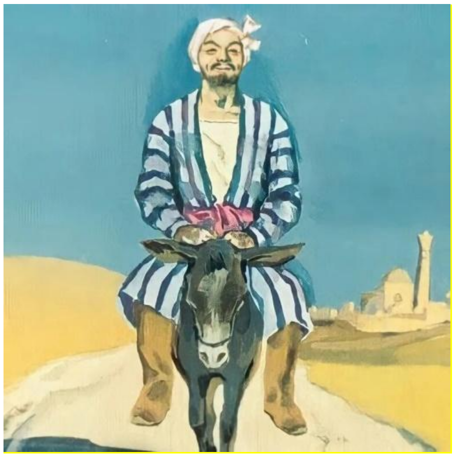
1. **Mass Spectrometry**: molecular ions, mass-to-charge ratio, isotope distribution.
2. **Kinetics**: steady-state and quasi-equilibrium approximations, enzymatic kinetics, interpretation of phase portraits.
3. **Thermodynamics**: heat capacity, temperature dependence of equilibrium constant, entropy and enthalpy change.
4. **Transition metal catalysis**: elementary steps in the catalytic cycles with transition metal complexes; single electron transfer (SET) and hydrogen atom transfer (HAT); cross-coupling reactions.
5. **Photochemistry**: Jablonski and Förster diagrams, fluorescence and phosphorescence, quenching, lifetimes, quantum yield.
6. **Carbohydrate chemistry**: represent chair conformations, carbohydrate reactions and protecting groups, repeating units.

The students are not expected to be specifically trained on the hereunder topics:

- Fragmentation in mass spectrometry.
- Advanced calculations (differentiation, iteration).
- Graph theory.
- 2nd and 3rd order kinetics.
- Cycloaddition reaction mechanisms.
- Memorising the structures of canonical amino acids.
- Archaea metabolism.
- Carothers equation.
- Memorising specific values of bond lengths.
- Memorising mechanisms of transition metal-catalysed coupling reactions.
- Details behind the techniques of STM and AFM.

Problem 1. Khoja Nasreddin is now a chemist

Khoja Nasreddin is a folklore character from the East, the hero of short humorous and satirical stories and anecdotes, and sometimes even everyday tales. In 2022, the tradition of retelling jokes about Khoja Nasreddin was inscribed on the Representative List of the Intangible Cultural Heritage of Humanity.



One day, Khoja Nasreddin decided to become a chemist. He began his journey in chemistry by studying the thermal decomposition of an unknown red salt **A**, even though such work was quite challenging for beginners.

Khoja Nasreddin carefully decomposed 1.00 g of **A** at −70 °C (**Reaction 1**) in a closed 1.00 L reactor initially filled with argon at a pressure of 100.0 kPa. After the decomposition, the pressure inside the reactor increased to 106.42 kPa, and the solid residue consisted of two compounds, one of which was identified as salt **B**.

Following the first stage of decomposition, he gradually heated the reactor to 200 °C. During heating, the solid residue first turned into a mixture of solid and liquid, then into a solely liquid phase, and finally, at 128 °C, a solid phase reappeared. At this temperature, the pressure in the reactor was 260.63 kPa.

Upon reaching 200 °C, **B** began to decompose (**Reaction 2**), and after complete decomposition, no solid residue remained in the reactor. The pressure at 200 °C was 396.82 kPa. When he introduced a glowing splint into the reactor, it ignited vigorously.

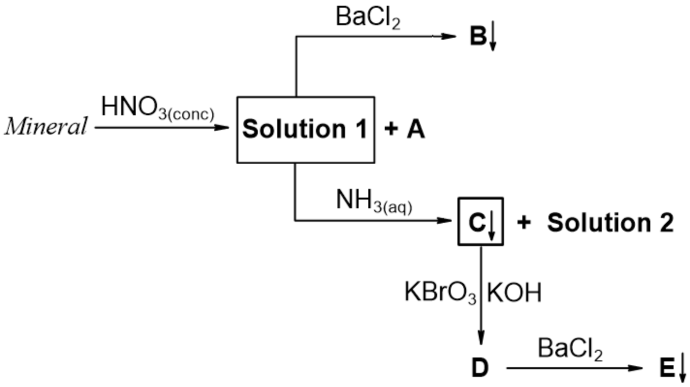
Nasreddin then attempted to synthesise salt **C**, which differs from **A** only by its cation. The mass fraction of the metal in **C** is equal to 12.63%. However, according to various sources, **C** is difficult to obtain. Instead, he successfully synthesised salt **D** (containing a complex cation, the same metal as in **C**, and 9.83% hydrogen by mass) by treating compound **E** (mass fraction of one of the elements in **E** is 4.21%) with gas **F** in liquid ammonia. He noted that at very low temperatures, **F** exists as a dark blue liquid.

To determine the amount of gas **F** required for the synthesis of 1.000 g of **D**, he passed as much **F** as necessary through 100 mL of potassium iodide solution (**Reaction 3**). For titration of a 10.00 mL aliquot of the resulting brown solution, 16.30 mL of sodium thiosulfate solution (**Reaction 4**) with a concentration of 0.0998 M was used. He also noted that when he was passing gas **F** through the iodide solution, another colourless gas was produced in the same amount as gas **F**.

1. **Identify** substances **A**–**F**.
2. **Write** the balanced equations of **Reactions 1**–**4**.

Problem 2. Avicenna

The philosopher Avicenna took a vial with a strange brownish-bronze mineral inside. Avicenna performed many reactions with this mineral and created a scheme to help to clearly identify the composition of the mineral.



Avicenna dissolved 1.00 g of the mineral in an excess of concentrated nitric acid and observed a release of a brown gas **A** with a volume of 1.65 L under standard conditions (0 °C, 1 atm). The solution formed (**Solution 1**) was divided into two equal parts for further analysis. To the first half of **Solution 1**, he added an excess of $BaCl_2$, which led to 0.929 g of white precipitate **B**. He added excess aqueous ammonia solution to the second half of **Solution 1**, causing the formation of brown precipitate **C** and deep-blue **Solution 2**. **C** was filtered off, dissolved in an alkaline solution of $KBrO_3$, giving a purple solution of compound **D**. **D** reacted with $BaCl_2$ to form 0.256 g of the red precipitate **E**.

1. **Determine** the possible elemental composition of the mineral.
2. **Determine** the molecular formula of the mineral and compounds **A**–**E**. **Show** your calculations where necessary.
3. **Write** all the reaction equations described in the scheme.

Problem 3. Unusual inorganic synthesis

Until 1968, it was believed that anion **A1** and its salts were unstable. However, in 1968, Evan H. Appelman performed an unusual synthesis of **A1**. A 90% enriched isotope of element **X** was first irradiated with thermal neutrons, then dissolved in nitric acid and oxidised by ozone in alkaline solution, giving anion **B**, which within a few minutes spontaneously turned into **A1** without absorbing or releasing any chemical substances. 1.613 g of anion **B** was obtained from 1.000 g of **X**.

After the first unusual but successful attempt, scientists managed to oxidise anion **A2** to **A1** using electrolysis. For this purpose, Pt electrodes were immersed into 3.0 mL of the solution containing **A2** (solution **S0**) and into $HClO_4$ solution. Electrolysis was carried out with a current strength of 10 A/cm^2 for 1 hour. The work surface of each Pt electrode was 0.785 mm^2, the current yield for **A1** was 11.5%.

1. **Write** the equation of the half-reaction occurring at the cathode and **calculate** the total volume (in mL at 0 °C, 1 atm) of the gas released at the cathode during electrolysis.

After electrolysis, the analyte contained a mixture of anions **A1** and **A2** (colourless solution **S1**). The Pt electrode was removed from the analyte, an excess of diluted (~1 M) acidic solution containing anion **A3** was added, and a yellow solution **S2** (containing elemental substance **A4**) was obtained. **A4** was removed by bubbling argon through **S2**, and a colourless solution **S3** was obtained. An excess of saturated (~10 M) acidic solution containing **A3** was added to solution **S3**, and a yellow solution **S4** (also containing elemental substance **A4**) was obtained. The whole volume of **S4** was diluted to 100.0 mL, and thus solution **S5** was obtained. An excess of acidified KI solution was added to a 10.0 mL aliquot of solution **S5**, and the I_2 formed was titrated with 0.100 M $Na_2S_2O_3$ solution. A titre of 1.35 mL of $Na_2S_2O_3$ solution was obtained.

Another 3.0 mL sample of solution **S0** was not subjected to electrolysis. Instead, it was diluted to 500.0 mL giving solution **S6**. An excess of acidified KI solution was added to a 10.0 mL aliquot of **S6**, and the I_2 formed was titrated with 0.100 M $Na_2S_2O_3$ solution. A titre of 10.10 mL of $Na_2S_2O_3$ solution was obtained.

2. **Write** the equation of the reaction occurring during titration between I_2 and $Na_2S_2O_3$.
3. **Determine** the chemical composition of **A1**–**A4**, **B**, and **X** (with atomic mass of enriched isotope). *Note: A1–A4 contain the same element.*
4. **Calculate** the molar concentration (in M) of **A2** in solution **S0**.

A1, formed within a few minutes from anion **B**, spontaneously decomposes within a few hours into gaseous elemental substances **C** and **D** in a molar ratio 1:2. The reaction scheme shows the formation and decay of **A1** with the corresponding rate constants

$$\mathbf{B} \xrightarrow{k_1 = 0.031\ \mathrm{min^{-1}}} \mathbf{A1} \xrightarrow{k_2 = 0.0048\ \mathrm{min^{-1}}} \mathbf{C} + \mathbf{D}$$

5. **Determine** the detailed chemical composition of **C** and **D**. Both reactions shown in the scheme above are first order with respect to the corresponding reactants.
6. **Calculate** the time (t, in min) after the start of the decomposition of **B**, which corresponds to the maximum concentration of **A1**.

It was subsequently discovered that XeF_2 also oxidises **A2** to **A1**. Interestingly, another binary compound, **E**, of xenon (which does not contain fluoride) was also obtained by exactly the same unusual method (like **B** → **A1**) from a potassium-containing salt **F** with a mass fraction of potassium of 12.62%.

7. **Determine** the detailed chemical composition of compounds **E** and **F**.

Problem 4. Nanozymes

Nanozymes are nanoparticles (NPs) that exhibit enzymatic activities under different physiological conditions, offering advantages such as high stability, low cost, and tunable activity compared to natural enzymes. Nanozyme $\mathbf{X}_nO_m$ NPs can be obtained by co-precipitation of salts **S1** and **S2**, both containing element **X**, in alkaline media. 2.7025 g of **S1** and 1.3901 g of **S2** were dissolved in an acidic solvent and treated with tetramethylammonium hydroxide, and 1.1578 g of $\mathbf{X}_nO_m$ NPs were formed as a result as the only substance containing **X**. The electrospray ionisation mass spectrometry of **S1** revealed positive ion $[\mathbf{X}\text{Hal}_2(H_2O)_4]^+$ with *m/z* 198 and negative ion $[\mathbf{X}\text{Hal}_4]^-$ with *m/z* 196 being the most intense peaks. **S2** lost 45.36% of the total of the original mass and changed colour from green to white above 300 °C. Elemental analysis revealed a mass fraction of 20.09% of element **X** in **S2**.

1. **Find** element **X**, and salts **S1** and **S2**.
2. **Determine** the values of n and m for $\mathbf{X}_nO_m$ NPs.
3. **Write** the equation of the co-precipitation reaction of $\mathbf{X}_nO_m$ NPs in ionic form.

Peroxidases are enzymes that catalyse the oxidation of substrates with H_2O_2 as an electron acceptor. Horseradish Peroxidase (HRP) catalyses a typical colourimetric reaction involving H_2O_2 and a chromogenic substrate (H_2TMB_{red}):

$$\mathrm{H_2TMB_{red} + H_2O_2 \xrightarrow{HRP} TMB_{ox} + 2H_2O}$$

$\mathbf{X}_nO_m$ NPs behave like peroxidase due to their ability to participate in the following reaction:

$$\mathbf{X}^{(n-1)+} + \mathrm{H_2O_2} \rightarrow \mathbf{X}^{n+} + \mathrm{OH}^* + \mathrm{OH}^-$$

4. The peroxidase activity of $\mathbf{X}_nO_m$ NPs is short-lived. After a certain time, $\mathbf{X}_nO_m$ NPs are converted into another oxide NPs of element **X** and lose their catalytic activity. **Determine** the chemical formula of the latter, if the mass fraction of **X** in them is smaller than that in $\mathbf{X}_nO_m$.
5. Which of the following reagents can be used to sustain the catalytic activity of $\mathbf{X}_nO_m$ NPs? **Choose** the correct answer(s).
   a) $NaIO_4$ b) $NaBH_4$ c) $Na_3C_6H_5O_7$ (sodium citrate) d) Na_3PO_4

In contrast, another $\mathbf{Y}_p$-based nanozyme can act not only as a peroxidase or catalase (converts hydrogen peroxide into water and oxygen) but also can exhibit hydrolase and phosphomonoesterase-like activities without being consumed during the reaction. $\mathbf{Y}_p$ has a spherical structure, in which the **Y** atoms are linked together by covalent bonds. The covalent spherical framework $\mathbf{Y}_p$ has a volume of 1.128×10^{-4} m^3·mol^{-1} and a surface area of 1.324×10^6 m^2·kg^{-1}.

6. **Identify** element **Y** and number p.

The catalytic activity of $\mathbf{Y}_p$-based nanozymes fits Michaelis-Menten kinetics.

The effect of functionalised $\mathbf{Y}_p$-nanozyme was tested with *para*-nitrophenyl phosphate at pH = 7.4 and $T = 37$ °C. The analysis showed $K_\mathrm{M} = 1.03$ mM and $r_\mathrm{max} = 3.73$ nM·s^{-1}.

7. The substrate's concentration should be high enough to saturate catalytic sites in nanozymes. If the recorded rate was 0.001 μM·min^{-1}, **calculate** the corresponding substrate concentration (in μM).

8. If the reaction mixture contained 0.025 mM substrate, **calculate** the time (in min) needed to produce 0.025 μM product.

Problem 5. Haber–Bosch process

Note: Necessary thermodynamic data are at the end of the problem.

In the Haber–Bosch process, ammonia (NH_3) is synthesised directly from the corresponding elements using an iron-based catalyst.

1. **Write** the equation for ammonia synthesis.
2. **Determine** the temperature, T (K), of the ammonia synthesis reaction, at which $K_p = 1$, assuming $\Delta_r H°$ and $\Delta_r S°$ do not depend on temperature.

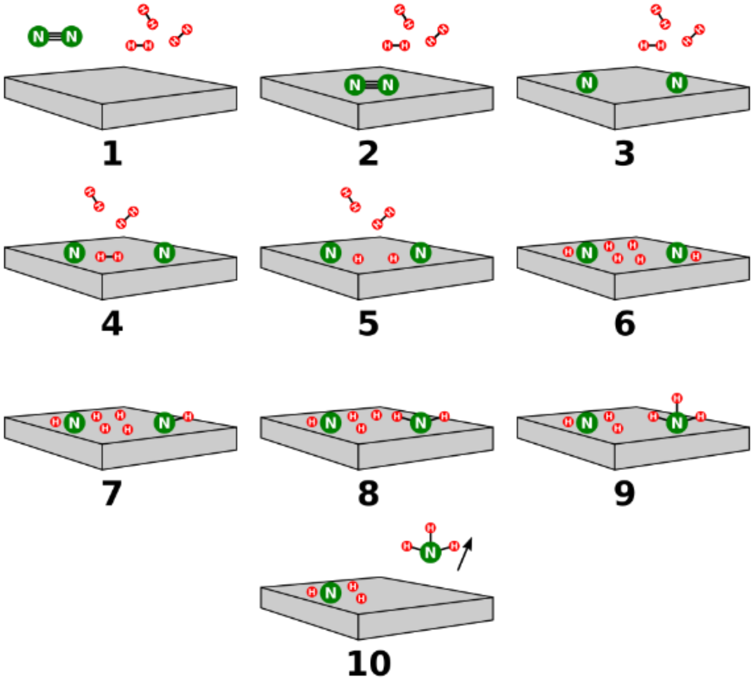
In reality, $\Delta_r H°$ and $\Delta_r S°$ depend on temperature.

3. **Calculate** the value of K_p at 300 °C, assuming that heat capacities of the gases remain constant while enthalpy and entropy depend on temperature.
4. **Determine** the minimum amount (moles) of hydrogen required to achieve 90% conversion of 1 mol of nitrogen.

Ammonia synthesis is exothermic. Thus, low temperatures favour higher equilibrium yields, but the reaction rate decreases significantly. Industrially, the process is conducted at high temperatures and pressures to achieve optimal rate and conversion.

5. a) **Calculate** the equilibrium yield, η (%), of ammonia at 300 °C and a final total pressure of 150 bar from a stoichiometric mixture of reactants.
   b) **Explain** qualitatively how the equilibrium yield would change if the total pressure were increased while keeping the temperature constant.

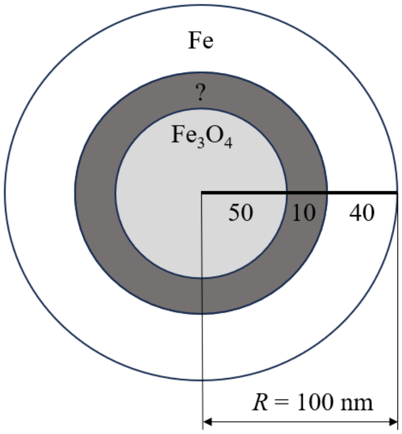
Ammonia formation proceeds via adsorption of gases on the catalyst surface, as shown in the diagram:



6. **Suggest** the possible rate-determining step of the mechanism.

The iron catalyst is produced from finely ground iron powder, which is usually obtained by reduction of high-purity magnetite (Fe_3O_4). The pulverised iron is oxidised to give magnetite particles of a specific size. The magnetite particles are then reduced. The catalyst maintains most of its bulk volume during the reduction, resulting in a highly porous high-surface-area material, which increases its catalytic efficiency.

The catalyst synthesised from 10 g of iron consists of nanoparticles with a radius, $R = 100$ nm, as shown in the diagram on the right.



7. **Choose** the composition of the "?" layer in the nanoparticle:
   a) Fe_2O_3 b) FeO c) FeOOH d) $Fe(OH)_3$
8. **Calculate** the total surface area of the catalyst, assuming all nanoparticles are spherical and uniform in size and composition.
9. **Calculate** how many times faster the ammonia synthesis reaction proceeds on the synthesised catalyst compared to a 10 g iron cube if the rate is directly proportional to the surface area.

Thermodynamic data (T = 298 K, p = 1 bar):

Substance	**$\Delta_f H°$ / kJ·mol^{-1}**	**$S°$ / J·mol^{-1}·K^{-1}**	**C_p / J·mol^{-1}·K^{-1}**
H_2(g)	0	130.7	28.8
N_2(g)	0	191.6	29.1
NH_3(g)	−46.2	192.7	35.1

Substance	Fe	FeO	Fe_3O_4	Fe_2O_3	FeOOH	$Fe(OH)_3$
ρ / g·cm^{-3}	7.9	5.7	5.0	5.3	4.3	3.9

Problem 6. Hydrogen fuel

Note: Necessary thermodynamic data are at the end of the problem.

Hydrogen is an environmentally clean fuel which is, in principle, ideal for powering the millions of vehicles on our roads.

1. **Write** the combustion reaction equations and **calculate** $\Delta_c H°$ at $T = 298$ K, $P = 1$ bar for:
   a) Hydrogen
   b) Methane
   c) Isooctane
2. **Determine** the energy density (MJ·kg^{-1}) for these fuels at $T = 298$ K and $P = 1$ bar. **Choose** the substance that is the best fuel, considering the energy-to-mass ratio?

For the rest of this problem, assume that air consists of 20% O_2 and 80% N_2 by mole fraction.

3. **Determine** the volume of air at $T = 298$ K and $P = 1$ bar required for an internal combustion engine (ICE) operating on hydrogen fuel ($\eta = 36\%$) with a power of 1 kW to work for 1 hour.

Hydrogen-fuelled cars do not release CO_2, in contrast to both gasoline and methane engines, which emit CO_2 as a direct product of fuel combustion. Uzbekistan maintains a national fleet of approximately 3.75 million cars. 65% of them operate on methane, while 35% run on gasoline (assume that isooctane is used). An average vehicle in the country travels 22 000 km per year.

Note:

- *Gasoline engines consume 9 dm^3 of fuel per 100 km ($\rho_{gasoline} = 0.72$ kg·dm^{-3}).*
- *Methane engines consume 5 m^3 of fuel per 100 km ($T = 298$ K and $P = 1$ bar).*

4. **Estimate** the amount of CO_2 emissions (in million tons) from cars that would be avoided annually, if every vehicle in Uzbekistan operated solely on hydrogen.
5. **Calculate** the amount (in moles) of nitrogen contained in an isolated system with initial temperature, $T_{\text{initial}} = 298$ K, and constant pressure ($P = 1$ bar), containing 1 mol of H_2 and excess air, if after combustion, the temperature of the system reached $T = 800$ K.

When stoichiometric mixtures undergo combustion, the system temperature increases to such a high level that the combustion reaction begins to proceed in the reverse direction. As a result, a situation arises in which both the composition of the mixture and the final temperature after combustion are unknown.

6. In another isolated reactor, the process takes place under a constant pressure ($P = 1$ bar) and an initial temperature, $T_{\text{initial}} = 298$ K, with the assumption that only the combustion reaction takes place.

a) **Determine** the final temperature for combustion of a stoichiometric mixture of hydrogen and oxygen, assuming that the reaction proceeds quantitatively.

b) **Determine** the mole fractions of gases in the mixture using K_p at the temperature from question 6-a).

c) **Determine** the final temperature for the system using an energy balance, if after combustion the mixture composition corresponds to the mixture from question 6-b). **Compare** the final temperature with the temperature in 6-a).

d) **Determine** the real final temperature (T_f, with an accuracy of ~20 K) for the combustion of a stoichiometric H_2 and O_2 mixture. **Choose** the nearest temperature value for T_f:

i) 3900 K ii) 3850 K iii) 3800 K iv) 3750 K v) 3700 K vi) 3650 K

Combustion in air at high temperatures may lead to side reactions, such as NO formation.

7. **Calculate** the mole fraction of NO after the stoichiometric combustion of H_2 in air at equilibrium ($T = 3200$ K, $P_{\text{total}} = 1$ bar). Assume that only two reactions that reach equilibrium occur, and use reasonable assumptions where necessary.

8. **Choose** the option(s) which help to decrease NO formation in ICE operating on H_2 and air:

a) Lower the combustion temperature.

b) Increase the excess air ratio significantly.

c) Raise the compression ratio in the engine.

d) Increase the fuel flow speed.

For this problem, assume that C_p does not depend on temperature, while enthalpy and entropy do depend.

Thermodynamic data ($T = 298$ K, $p = 1$ bar):

Substance	$\Delta_f H°$ / kJ·mol^{-1}	$S°$ / J·mol^{-1}·K^{-1}	C_p / J·mol^{-1}·K^{-1}
$H_2(g)$	0	130.7	28.8
$O_2(g)$	0	205.2	29.4
$N_2(g)$	0	191.6	29.1
$H_2O(g)$	–241.8	188.7	37.4
$H_2O(l)$	–285.8	70.0	75.4
$CH_4(g)$	–74.8	186.3	35.3
$C_8H_{18}(l)$	–260.2	361.2	254.7
$CO_2(g)$	–393.5	213.8	37.1
NO(g)	90.3	210.8	29.9

Problem 7. 150 Years of Urease

Urease is an important enzyme because of its role in the nitrogen cycle as a key catalyst in the reaction converting urea to NH_3 and CO_2:

$$CO(NH_2)_2 + H_2O \xrightarrow{\text{Urease}} CO_2 + 2NH_3$$

Urease activity as a soluble enzyme was first identified in 1876 by Frédéric Alphonse Musculus. Urease was the first enzyme crystallised in 1926 by James Batcheller Sumner. For his discovery that enzymes could be crystallised, Sumner was awarded the Nobel Prize in Chemistry in 1946.

The active centre of Urease behaves as a diprotic acid both in the free state and the substrate-bound state, with the following equilibrium constants at 25 °C:

$$H_2U \overset{K_{a1}}{\rightleftarrows} HU^- + H^+, \qquad pK_{a1} = 5.62$$

$$HU^- \overset{K_{a2}}{\rightleftarrows} U^{2-} + H^+, \qquad pK_{a2} = 9.07$$

Only the HU^- form exhibits enzymatic activity, which obeys Michaelis–Menten kinetics as shown below:

$$\text{Urease} + CO(NH_2)_2 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} [\text{Urease}\cdot CO(NH_2)_2] \xrightarrow{k_2} \text{Urease} + CO_2 + 2NH_3$$

1. Using the steady-state approximation, **derive** the dependence of the rate of the above enzymatic reaction on the total concentrations of Urease, $CO(NH_2)_2$ and H^+. *Note:* $[CO(NH_2)_2]_{total} \gg [Urease]_{total}$.

2. **Determine** the reaction orders with respect to Urease and $CO(NH_2)_2$ in the buffer solutions when:

a) $\frac{k_{-1}+k_2}{k_1} \gg [CO(NH_2)_2]_{total}$

b) $\frac{k_{-1}+k_2}{k_1} \ll [CO(NH_2)_2]_{total}$

3. The dependence of the rate of these enzymatic reactions on pH forms bell-shaped graphs with the maximum rate when the total concentrations of Urease and $CO(NH_2)_2$ are unchanged. **Calculate** the pH value when Urease demonstrates the maximum rate.

4. In the absence of any buffer, the decomposition products of $CO(NH_2)_2$ change the pH of the solution, thereby reducing the activity of Urease. **Calculate**:

a) The pH of the solution after decomposition of 0.005 M $CO(NH_2)_2$. Neglect the contribution of Urease to the acid-base equilibrium and the dissolution of atmospheric CO_2.
$pK_b(NH_3) = 4.75$; $pK_{a1}(H_2CO_3) = 3.75$; $pK_{a2}(HCO_3^-) = 10.33$ at 25 °C.

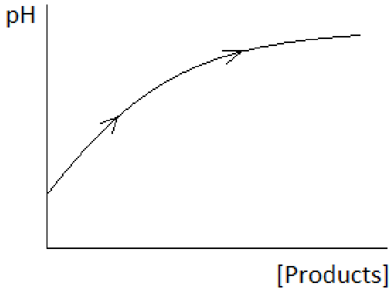
b) The factor by which the rate of the enzymatic reaction decreases when the pH of the solution changes from neutral to the value obtained in question 4-a), if the total concentrations of Urease and $CO(NH_2)_2$ are the same in both conditions.

5. **Match** the following conditions (1)–(3) with the phase portraits (a)–(c):

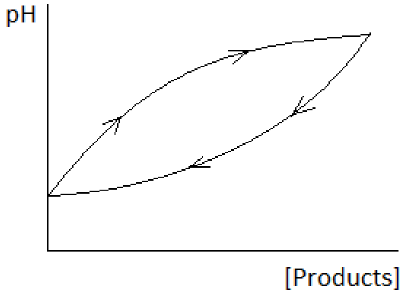
(1) Urease in buffer solution with $CO(NH_2)_2$;

(2) Urease in neutral solution with $CO(NH_2)_2$ without any buffer;

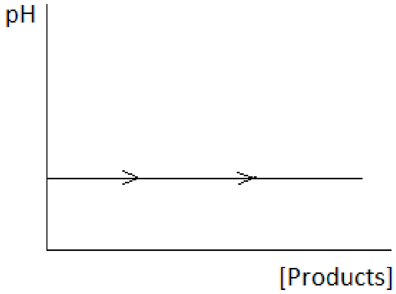
(3) Urease in liposomes with a permeable membrane with $CO(NH_2)_2$ and without any buffer inside. Assume that Urease doesn't leave the liposome, while $CO(NH_2)_2$ and products can pass through the membrane.



(a)



(b)



(c)

Biocatalytic calcification by Urease shows potential in innovative engineering applications, such as preparation of advanced carbonate materials, cleaning wastewaters/groundwaters of Ca^{2+} and $^{90}Sr^{2+}$, and most excitingly, sealing cracks and fissures in buildings and other historical stone monuments of cultural heritage during their restoration.

6. **Calculate** the minimum concentration of Ca^{2+}, at which $CaCO_3$ begins to precipitate from solution in question 4-a)? **Ignore** the dissolution of atmospheric CO_2.
$K_{sp}(CaCO_3) = 3.36 \times 10^{-9}$ at 25 °C.

In agriculture, Urease in soil plays a crucial role in nitrogen uptake by plants in the form of NH_3 produced from the hydrolysis of the fertiliser $CO(NH_2)_2$. However, if too rapid, the hydrolysis may lead to overproduction of NH_3 resulting in detrimental effects, such as losses of nitrogen by ammonia volatilisation, and ammonia- and alkaline-induced plant damage. One way to counteract this is to suppress the ureolytic activity of soil with the use of Urease inhibitors. Urease inhibition can also be exploited as an analytical technique in Urease-based biosensors for the analysis of substances that act as inhibitors of the enzyme.

Boric acid is a competitive inhibitor of Urease and their complex, [Urease·B(OH)3], also behaves as a diprotic acid with the same acidity constants:

$$[\text{Urease}\cdot B(OH)_3] \overset{K_I}{\rightleftarrows} \text{Urease} + B(OH)_3$$

7. **Rederive** the dependence of the rate of the enzymatic reaction on the total concentrations of Urease, $CO(NH_2)_2$, $B(OH)_3$, and H^+. *Note:* $[CO(NH_2)_2]_{total}$ *and* $[B(OH)_3]_{total} \gg [Urease]_{total}$. **Ignore** the acid-base properties of $B(OH)_3$.

8. According to the kinetic equation obtained in question 7, the function $\frac{1}{r} = f\left(\frac{1}{[\mathrm{CO(NH_2)_2}]_{total}}\right)$ is linear. **Calculate** K_I if the slope of this function with $[\mathrm{B(OH)_3}]_\mathrm{total}$ = 0.300 mM is 2.48 times greater than the slope with $[\mathrm{B(OH)_3}]_\mathrm{total}$ = 0.075 mM at 25 °C. The total concentration of Urease and pH values are equal in both cases.

9. **Calculate** $\Delta_\mathrm{r}H°$ and $\Delta_\mathrm{r}S°$ values for the binding of boric acid with Urease, if they are independent of temperature and $K_\mathrm{I} = 1.57 \times 10^{-4}$ at 35 °C.

Problem 8. Step-by-step

Uzbekistan's polymer industry is a growing sector, driven by domestic demand in construction, packaging, and agriculture. Polycondensation, also known as condensation polymerisation, leads to multiple important polymeric products and is known to proceed by a stepwise mechanism.

Let's consider a polymer produced by polycondensation of HO–A–OH and HOOC–B–COOH monomers. Such polymers, obtained via esterification reactions, are called polyesters. It is generally assumed that the reactivity of functional groups is independent of the length of the corresponding molecular chain.

$$n\text{HO–A–OH} + n\text{HOOC–B–COOH} \xrightarrow{\text{Cat}} \text{H}[\text{O–A–O–CO–B–CO}]_n\text{OH} + (2n-1)\text{H}_2\text{O}$$

$$-\frac{d[\text{OH}]}{dt} = -\frac{d[\text{COOH}]}{dt} = k[\text{OH}]^x[\text{COOH}]^y$$

The Initial Rate Method was used to determine the kinetic order of the reaction.

Initial Rate / M·s^{-1}	**$[\text{OH}]_0$ / M**	**$[\text{COOH}]_0$ / M**
2.70×10^{-4}	0.100	0.100
1.08×10^{-3}	0.200	0.200
1.62×10^{-3}	0.300	0.200

1. Using the information above, **calculate** the values of k, x and y.

It is reasonable to consider that the concentrations of the functional groups are equal, that is $[\text{OH}] = [\text{COOH}] = C$. Thus, the kinetic equation can be further simplified as:

$$-\frac{dC}{dt} = kC^z$$

2. **Give** the value of z in this case. **Write** the integrated kinetic equation.

The Carothers equation relates the average degree of polymerisation (DP) of polymers with the degree of conversion, p, where C is the concentration of functional groups at a particular time and C_0 is the initial concentration of those groups in monomers.

$$DP = \frac{C_0}{C}$$

3. **Derive** the Carothers equation $DP = f(p)$.
4. **Find** the dependence of DP on time, $DP = f(t)$.
5. Using the functions $DP = f(t)$ and $DP = f(p)$, **present** the relationship between p and t in a linear form.

Polycondensation reactions typically have an overall reaction order of 2. However, this can be different in particular cases. Thus, aryl halide–bisphenol A polycondensation systems are unusual because they show different kinetic behaviours depending on whether the halide is chlorine or fluorine. The following mechanism of polycondensation kinetics was found for DCDPS (dichlorodiphenylsulfone, 2-body pathway) and for DFDPS (difluorodiphenylsulfone, 3-body pathway):

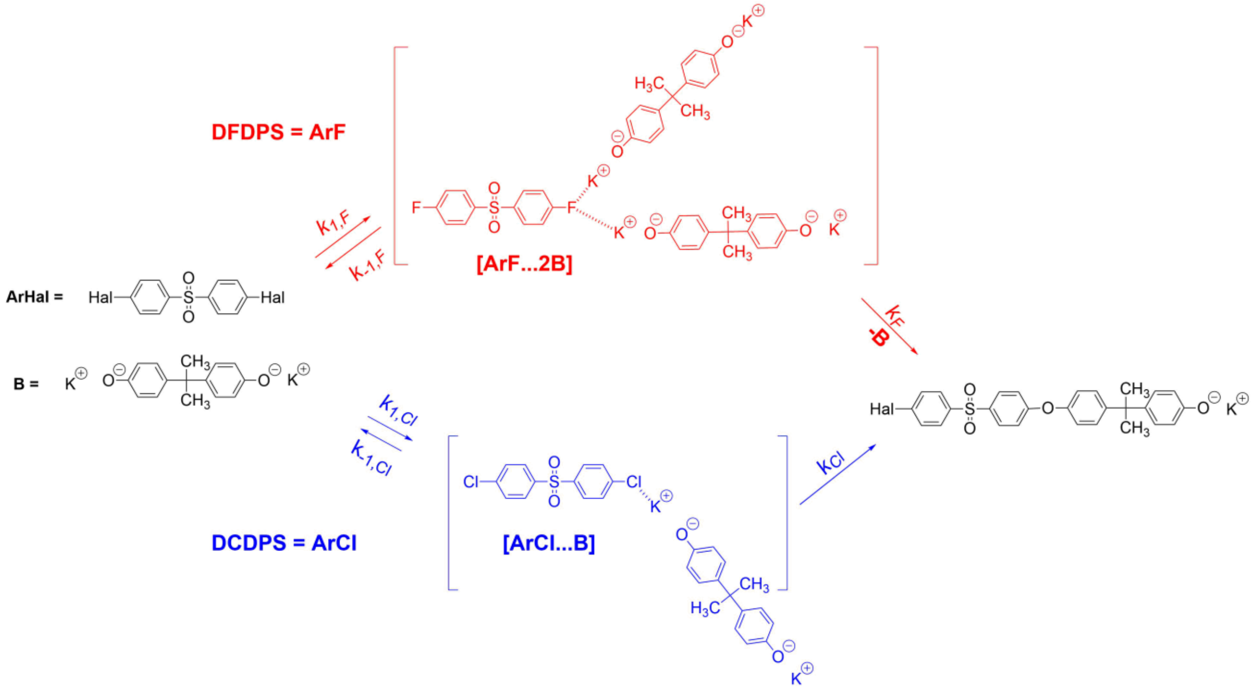


Figure 1. Substrate-dependent mechanisms of reaction for DFDPS (ArF) and DCDPS (ArCl)

6. Based on the above mechanisms, **find** the expressions for the equilibrium constants K_F and K_{Cl} of the formation of intermediates [ArF…2B] and [ArCl…B], respectively.

7. Assuming that equilibrium is established very quickly in the first step, **find** the overall reaction rate law $r = k_{obs}[\text{ArHal}]^m[\text{B}]^n$ for both halides, using K_F and K_{Cl}, respectively.

Polymerisation of DFDPS can be schematically shown as follows:

$$\mathbf{ArF + 2B \underset{k_{-1,F}}{\overset{k_{1,F}}{\rightleftharpoons}} [ArF...2B] \xrightarrow{k_F} P + B}$$

Assume that $[\text{ArF}]_0 = [\text{B}]_0$ and $[\text{ArF}] = [\text{B}]$ at any time.

8. **Deduce** the relationship between the degree of polymerisation and time for DFDPS polymerisation. **Calculate** the average degree of polymerisation after 1 hour, if the initial concentration $[\text{ArF}]_0 = [\text{B}]_0 = 0.1633$ M and $k_{obs} = 7023\ \text{L}^2\cdot\text{mol}^{-2}\cdot\text{h}^{-1}$.

So far, we've considered a model in which the rate constant is independent of the degree of polymerisation (*DP*). However, reactions are known that don't conform to this model. Consider the more complicated model where every propagation step of the step-growth reaction accelerates by a factor ξ.

9. **Sketch** qualitatively $DP = f(t)$ curves for $\xi = 1$ (common case, already done in question 4) and $\xi > 1$ in one graph.

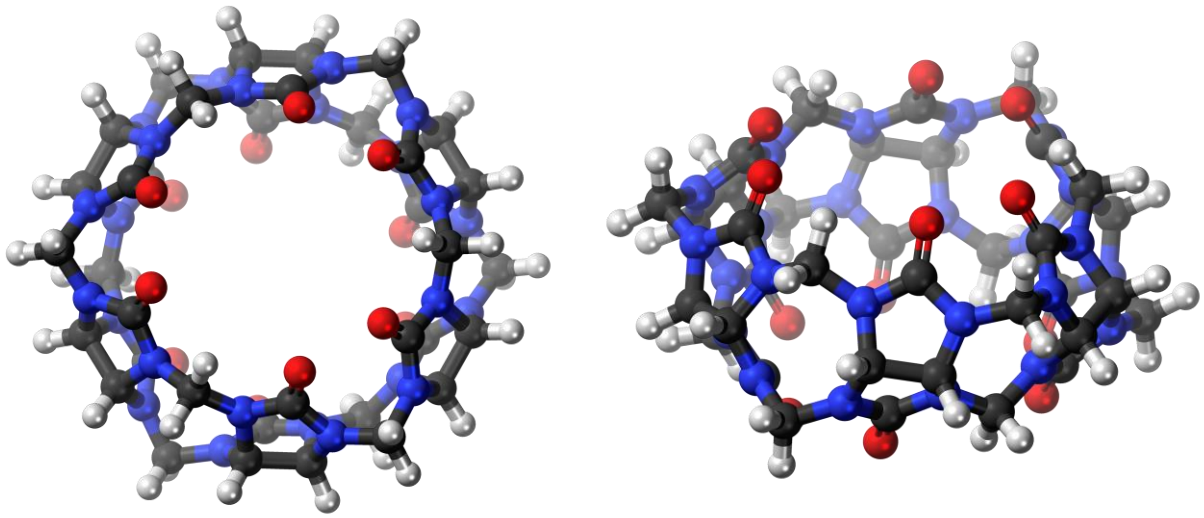
10. **Derive** the relationship $DP = f(t)$ for $\xi > 1$.

Problem 9. Host-guest chemistry of cucurbiturils as Uzbek hospitality

The cucurbiturils are a family of macrocyclic compounds, which, incidentally, resemble pumpkins. They are doughnut-shaped molecules with functional groups within their voids that can bind small guest molecules.

Fortunately, the synthesis of cucurbiturils is extremely simple. The scheme below shows the synthesis of cucurbit[6]uril (**X**), the [6] indicating a macrocycle with six repeating units.

Cucurbit[6]uril (**X**)



Top view *Side view*

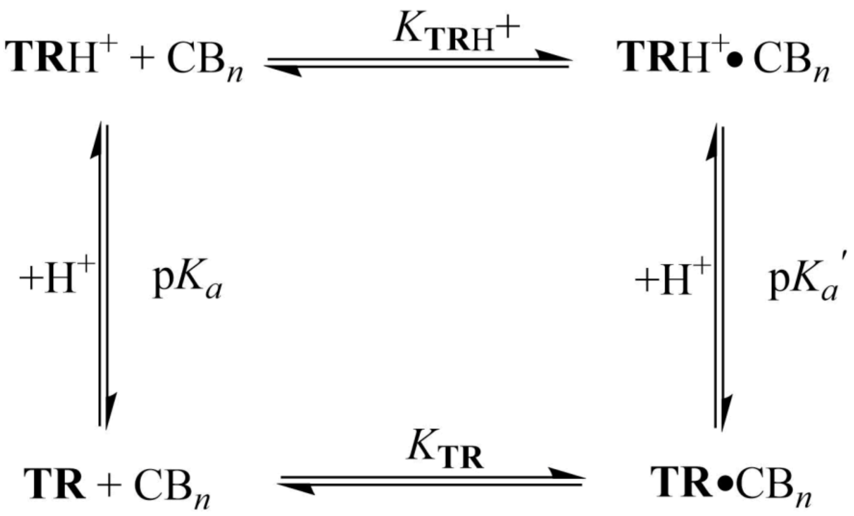
$$2\mathbf{A} + \mathbf{B} \longrightarrow \mathbf{C} \xrightarrow[110\ ^\circ\mathrm{C}]{\mathrm{CH_2O,\ H_2SO_4}} \mathbf{X}$$

Hints:

- ***A** is obtained by heating an inorganic salt of **Y**, which happens without mass loss;*
- *In the mass spectrum, **B** has peaks at m/z 58 and 29. **B** gives a positive Tollens reaction and has only one type of hydrogen atoms and one type of carbon atoms;*
- *Compound **C** is not polymeric, and **C** is the main building block for constructing cucurbiturils.*

1. **<u>Draw</u>** structures **A**–**C** and the repeating unit of **X**. **<u>Determine</u>** **Y**.

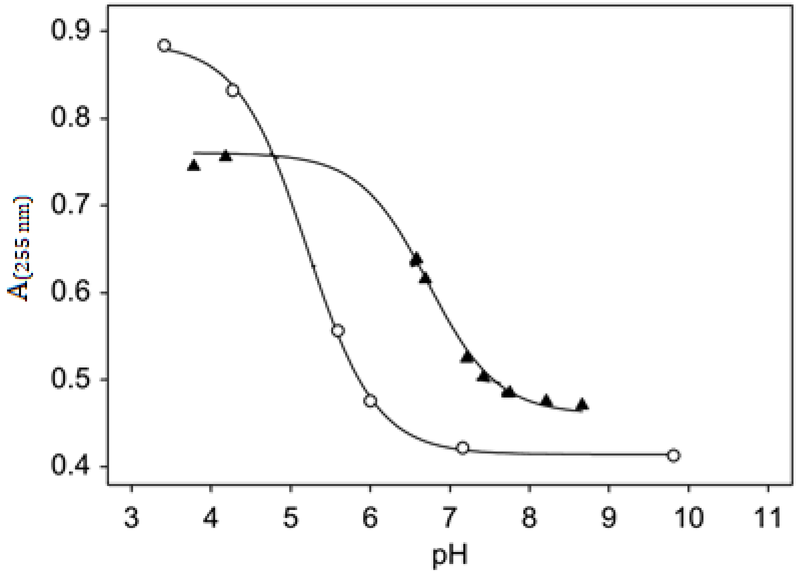
Cucurbiturils form host-guest type complexes. For example, the host cucurbit[*n*]uril (CB_n) can stabilise the ocular drug tropicamide (TR – $C_{17}H_{20}N_2O_2$). The four-state complexation model of neutral and protonated TR with CB_n is shown below:



The [$TR \cdot CB_8$] complex was studied using mass spectra. Peaks at *m*/z 285, 665, and 807 were detected in the spectrum.

2. **Determine** the species to which these peaks correspond.

For the [$TR \cdot CB_8$] complex, the formation constant, $K_{TR} = 1.4 \times 10^4$. pH titrations monitored by UV-spectrophotometry of pure TR (○), and of TR with CB_8 (▲) are shown below. The concentration of the host was sufficient to ensure virtually quantitative complexation of TR over the investigated pH range.

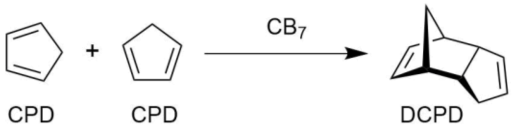


3. **Determine** the approximate values of pK_a, pK_a' and K_{TRH^+}.

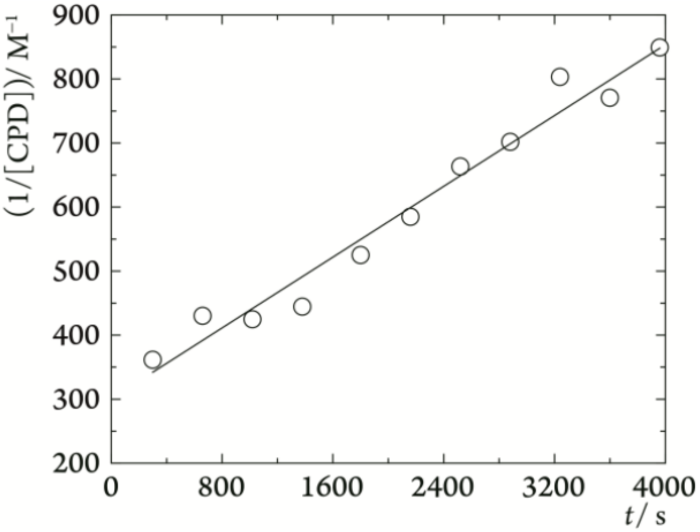
Tropicamide is an anticholinergic drug commonly used as eye drops for mydriatic effect (pupil dilation) in preparation for ophthalmological examinations and surgical procedures. The drug is available as an aqueous solution with pH = 4.0. Since this value is significantly lower than the physiological pH of the eye (pH = 7.0), topical application of the drug causes significant, albeit temporary, discomfort in patients. Using tropicamide with the CB_8 complex allows the drug to be used at physiological pH.

4. **Indicate** whether TR or TRH^+ is the biologically active form of tropicamide.
5. **Calculate** how many times the molar fraction of the biologically active form of tropicamide increases at pH = 7.0 if its complex with the CB_8 is used.

Cucurbiturils are also used as supramolecular catalysts. For example, cucurbit[7]uril (CB_7) catalyses the dimerisation reaction of cyclopentadiene (CPD) to dicyclopentadiene (DCPD):



The plot of 1/[CPD] (M^{-1}) versus reaction time (s), monitored in the presence of CB_7, is as follows:



6. **Determine** the reaction order for CPD and the observed rate constant.

The following equations describe the kinetics based on the binding equilibria:

$$CPD + CB_7 \rightleftharpoons CPD{\cdot}CB_7, \qquad K_1$$

$$CPD + CPD{\cdot}CB_7 \rightleftharpoons (CPD)_2{\cdot}CB_7, \qquad K_2$$

$$(CPD)_2{\cdot}CB_7 \rightarrow DCPD{\cdot}CB_7, \qquad k$$

7. **Derive** the dependence of the rate of the dimerisation reaction on the equilibrium concentrations of CPD and CB_7, which agrees with the experimental data from question 6.

The following table shows the observed rate constants for cucurbiturils with the same initial concentrations and different cavity volumes:

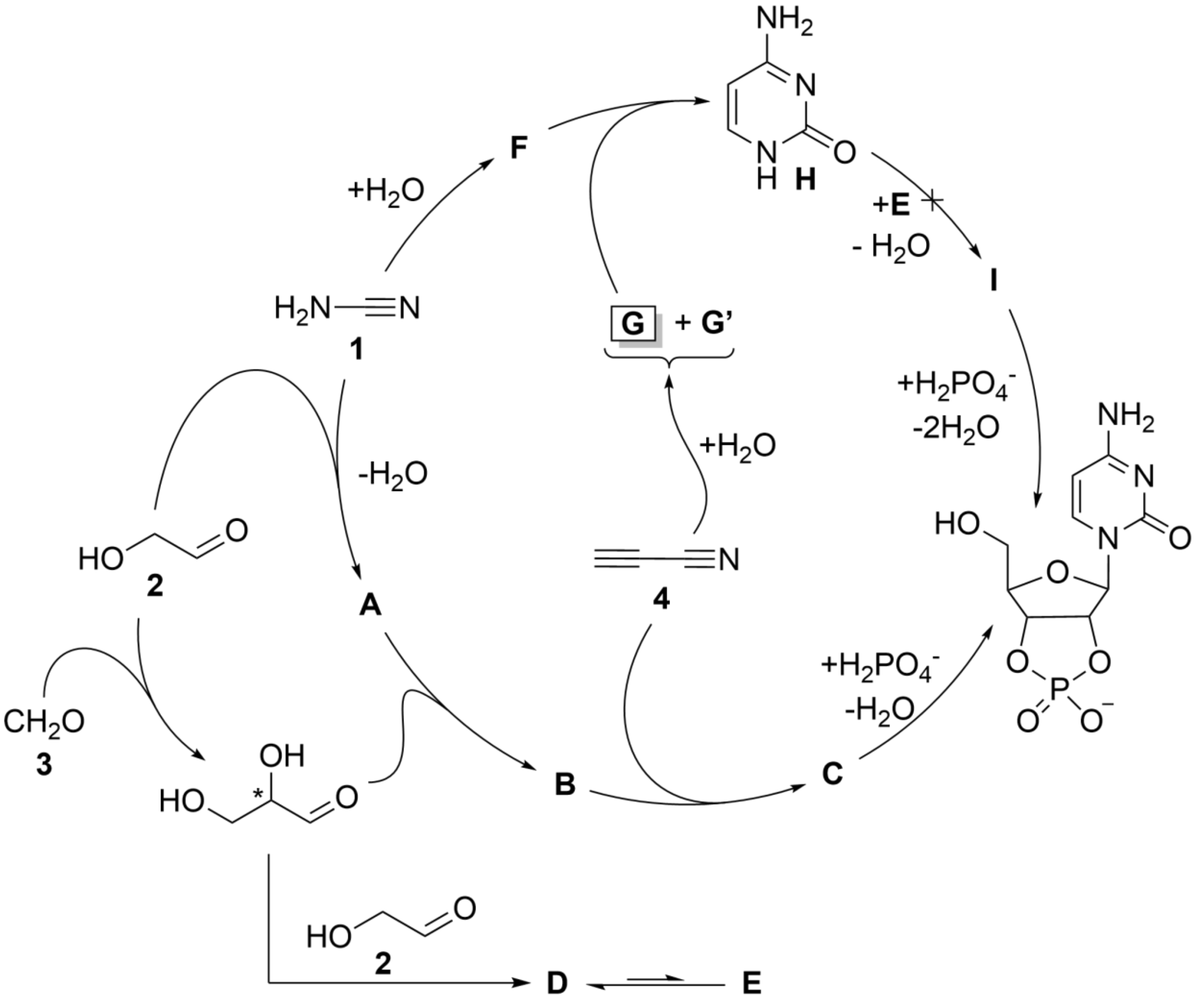
Cucurbit[*n*]uril	**CB_5**	**CB_6**	**CB_7**	**CB_8**
k_{obs} / $M^{-1}\cdot s^{-1}$	9.4×10^{-6}	5.0×10^{-6}	?	9.0×10^{-6}

8. **Explain** why CB_6 has the lowest observed rate constant.

Problem 10. Prebiotic life

What makes Earth unique is that it has life. As far as we know, it is the only place in the universe where complex forms of life exist. The impact of life has been so great that it has changed the planet's air, water, and climate. Therefore, a central goal of science is to understand what life is, how it works, and ultimately, how it emerged from non-living matter. This leads to the core mysteries of prebiotic chemistry: what conditions and processes on the early Earth transformed simple chemicals into the first self-replicating systems that kickstarted biological evolution?

One of the possible routes of how complex molecules like RNA and ribonucleotides formed from simple molecules cyanamide **1**, glycoaldehyde **2**, formaldehyde **3**, and cyanoacetylene **4** in prebiotic Earth was suggested by Powner *et al.*:



In this scheme, direct reaction between sugar **E** and cytosine **H** (which is formed via reaction of **G**) yields no product **I**, presumably because of the low nucleophilicity of **H**. Therefore, the **1** + **2** → **A** → **B** → **C** prebiotic route was suggested by researchers, where in each step the number of rings in the product increases by one.

1. **Draw** molecular structures of compounds **A**–**G**, **G'** and **I**.
2. What is the reason behind the low nucleophilicity of **H**? **Choose** the correct answer(s):
   a) electron delocalisation due to the aromaticity of pyrimidine ring;
   b) extreme ring strain in cytosine that interrupts the reaction;
   c) the positive charge due to the formation of ammonium salts that repels other molecules.

The fact that life relies almost exclusively on single molecular handedness (homochirality) is still not fully understood. It is well established that, in biological systems, most amino acids in proteins occur in the so-called L-configuration, while most sugars adopt the D-configuration.

3. In the scheme above, glycoaldehyde **2** and formaldehyde **3** react to form glyceraldehyde – the simplest sugar molecule that contains a stereocentre. **Depict** the structural formulae of L- and D-glyceraldehyde with the shown absolute configuration of the stereocentre.

There are several hypotheses about the origin of such a unique property as homochirality. A plausible early Earth scenario involves shallow lakes with magnetite deposits. UV light (200-300 nm) shining on magnetite can generate spin-polarised electrons – special electrons with aligned spins – that act as reducing agents. These electrons interact differently with left- and right-handed molecules due to the chiral-induced spin selectivity (CISS) effect, changing reaction rates selectively. This spin-based interaction could break chiral symmetry in prebiotic chemistry, favouring one enantiomer and possibly leading to life's homochirality.

In this case, it is useful to present the term of electron helicity, which is a physical property describing the direction of the electron's spin relative to its direction of motion. For example, if helicity is right-handed, then the electron's spin is parallel to its momentum, and vice versa.

Consider a process that occurs in a prebiotic lake of reactants:

$$\mathrm{LX} + \mathrm{De}^{-} \rightarrow \mathrm{LY}$$

$$\mathrm{DX} + \mathrm{De}^{-} \rightarrow \mathrm{DY}$$

The L-/D- isomer of **X** reacts with a right-handed electron donor De⁻ and produces the L-/D- isomer of compound **Y**. The spin-chirality coupling hypothesis states that the interaction of the L-isomer with the D-electron is more kinetically favoured. The amount of this preference can be expressed via the rate constant formula:

$$k_{(\mathrm{L})} = A \cdot \exp\left(-\frac{(E_a - H_{\mathrm{SO}})}{k_{\mathrm{B}}T}\right),$$

$$k_{(\mathrm{D})} = A \cdot \exp\left(-\frac{(E_a + H_{SO})}{k_{\mathrm{B}}T}\right),$$

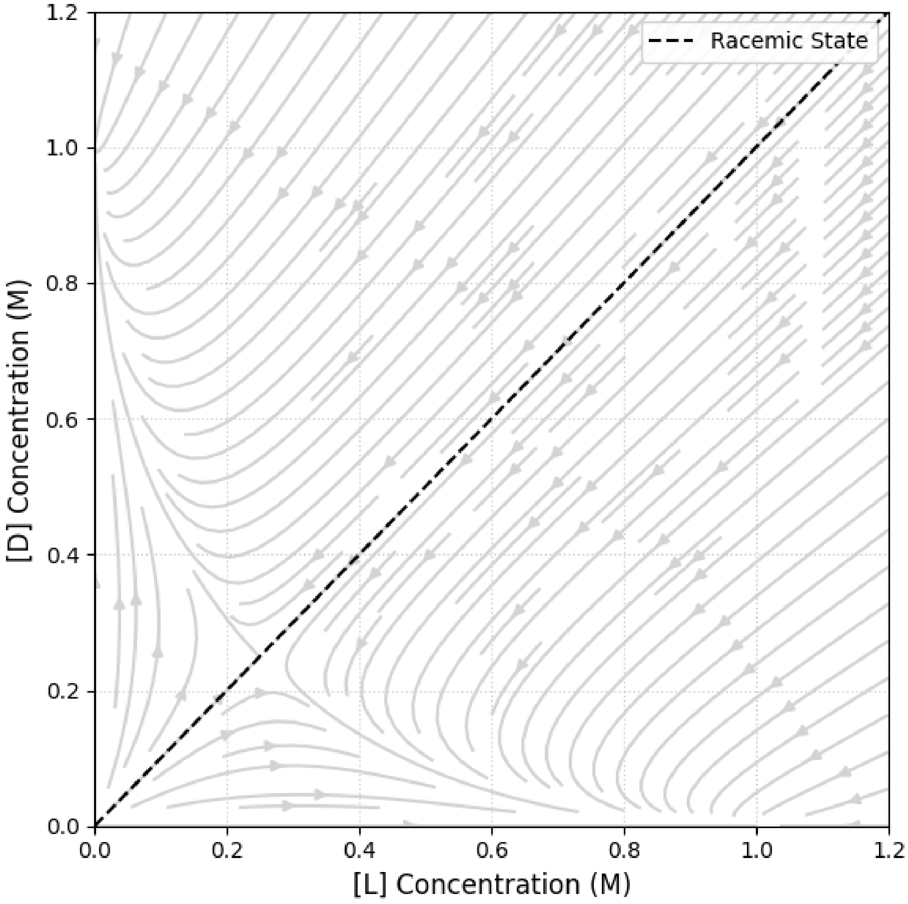
where $k_{L/D}$ is the rate constant of the reaction, E_a is the activation energy of the process, and H_{SO} is the spin-orbit energy.

The enantiomeric excess, *ee*, is a measure of the purity of a chiral substance in terms of the predominance of one enantiomer over the other in a mixture:

$$ee = \frac{[\mathrm{LY}] - [\mathrm{DY}]}{[\mathrm{LY}] + [\mathrm{DY}]}$$

4. Given $H_{SO} = 0.5$ meV, considering that the A factor is equal for L- and D-isomers and the substrate is a racemic mixture ([LX] = [DX]), **determine** the *ee* of the reaction product **Y** at $T = 313$ K (~40 °C), which is a good assumption for the average surface temperature of water bodies during the prebiotic era.

5. Consider some racemic solution of L-/D-isomers in water with a total concentration of 1 M, where symmetry breaking (i.e. conversion of one isomer into another) occurs, resulting in overall chiral amplification of the system. **Depict** possible symmetry-breaking trajectories in a phase portrait of the system.



This example of kinetic resolution has demonstrated the chiral amplification of single monomeric units. However, the origin of life requires processes for the ligation of monomers into the homochiral polymeric molecules that constitute proteins and enzymes and the genetic polymers of RNA. Can a pool of partially enantioenriched monomers successfully produce homochiral polymers?

Let's consider two amino acids in their L- and D- forms. The possible reactions between them are listed below:

$$L_1 + L_2 \rightarrow L_1L_2, \qquad r = k_{\text{homo}}[L_1][L_2]$$

$$D_1 + L_2 \rightarrow D_1L_2, \qquad r = k_{\text{hetero}}[D_1][L_2]$$

$$L_1 + D_2 \rightarrow L_1D_2, \qquad r = k_{\text{hetero}}[L_1][D_2]$$

$$D_1 + D_2 \rightarrow D_1D_2, \qquad r = k_{\text{homo}}[D_1][D_2],$$

where k_{homo} and k_{hetero} denote the homochiral and heterochiral dimer formation rate constants, respectively.

6. Considering that $k_{\text{homo}} = k_{\text{hetero}}$ and thus neglecting kinetic control, **derive** that:

$$ee_{\text{homochiral}} = \frac{[\mathrm{L_1L_2}] - [\mathrm{D_1D_2}]}{[\mathrm{L_1L_2}] + [\mathrm{D_1D_2}]} = \frac{ee_1 + ee_2}{1 + ee_1ee_2}$$

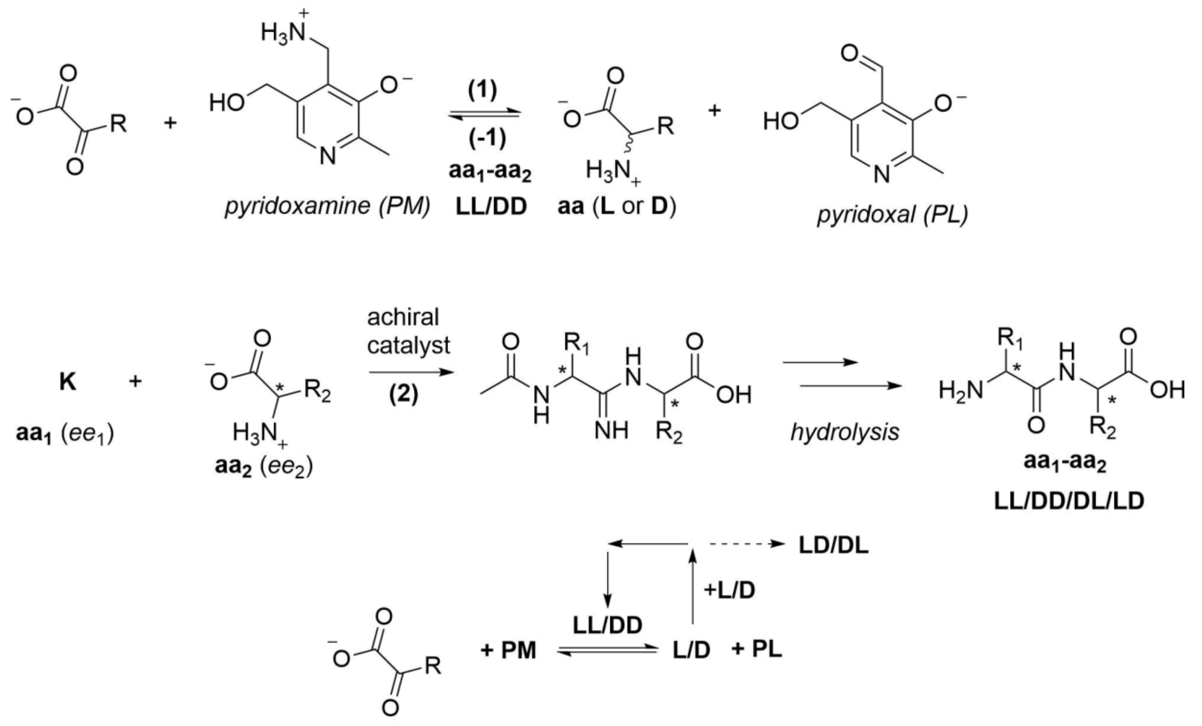
where ee_1 and ee_2 stand for:

$$ee_1 = \frac{[\mathrm{L_1}] - [\mathrm{D_1}]}{[\mathrm{L_1}] + [\mathrm{D_1}]}, \qquad ee_2 = \frac{[\mathrm{L_2}] - [\mathrm{D_2}]}{[\mathrm{L_2}] + [\mathrm{D_2}]}$$

7. **Calculate** $ee_{\text{homochiral}}$ for the prebiotic lake system from question 4 with $ee_1 = ee_2 = ee_{q4}$.
Note: If you could not solve question 4, consider $ee_{q4} = 0.05$ *(5%).*

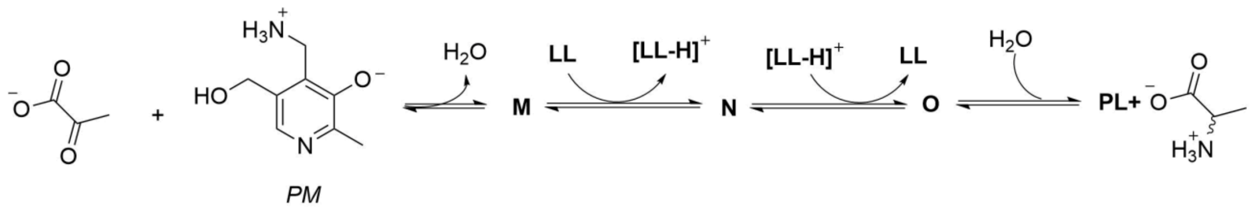
8. Now, considering that only homochiral oligomers are responsible for chain propagation, **find** the smallest number of amino acid residues needed for a resulting polymeric chain to have $ee_{\text{homochiral}} > 0.99$. For initial amino acids, **assume** each amino acid has an $ee = 0.05$.

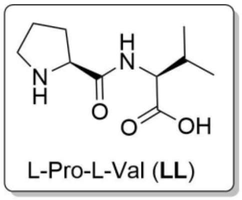
The mathematical model above was found valid for a real biochemical system, where the initial transamination reaction with pyridoxamine catalysed by a strictly homochiral dipeptide molecule **aa₁-aa₂** yields a racemic amino acid molecule **aa** and pyridoxal. The amino acid molecules formed react with another amino acid's derivative molecule **K**, which results in a dipeptide derivative. The latter yields the dipeptide **aa₁-aa₂**, thus looping the cycle:



9. **Draw** the structure of **K**.

The role of the dipeptide catalyst in reaction **(1)** is based on promoting a hydrogen atom transfer (HAT) reaction. For instance, Yu *et al.* proposed the mechanism below for the L-prolinyl-L-valine-catalysed alanine synthesis reaction from pyruvate:




L-Pro-L-Val (**LL**)

10. **Mark** the functional group(s) in the **LL** molecule involved in the hydrogen atom transfer process.

11. **Draw** the structures of compounds **M**–**O**.

Yu *et al.* also discovered that in the reverse reaction **(−1)**, which leads to pyruvate from alanine, D-alanine reacts faster than L-alanine, thus enriching the prebiotic system with the L-enantiomer.

12. Which intermolecular forces (or their absence) between alanine and **PL-LL** in the corresponding transition state could be behind the faster reaction of D-Ala compared to that of L-Ala? **Choose** the correct statement(s):
   a) Hydrogen bonds
   b) π-stacking
   c) Steric repulsion
   d) Alignment of stereocentres

Constructive and destructive mechanisms, where the corresponding amino acid was formed or destroyed, respectively, were investigated in open systems with a flow rate, φ, of corresponding reactants in each case.

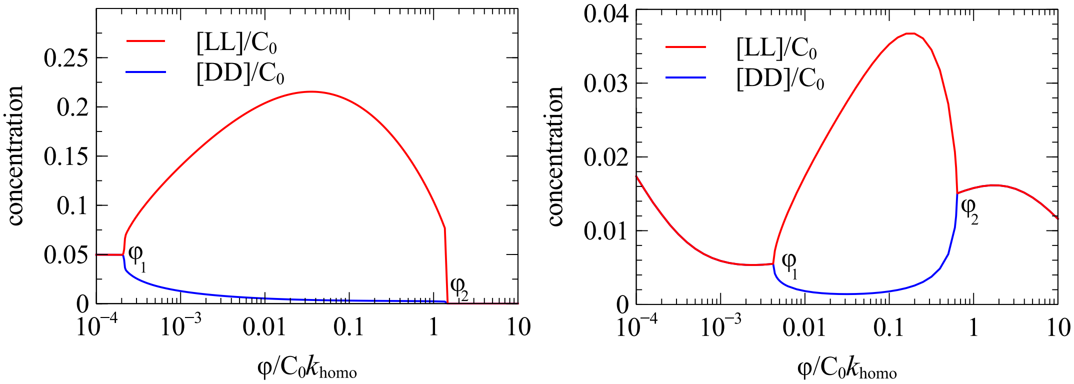


Figure 2. Concentration of homochiral dipeptides vs flow rate of a) ketoacid (left, constructive mechanism); b) amino acid (right, destructive mechanism)

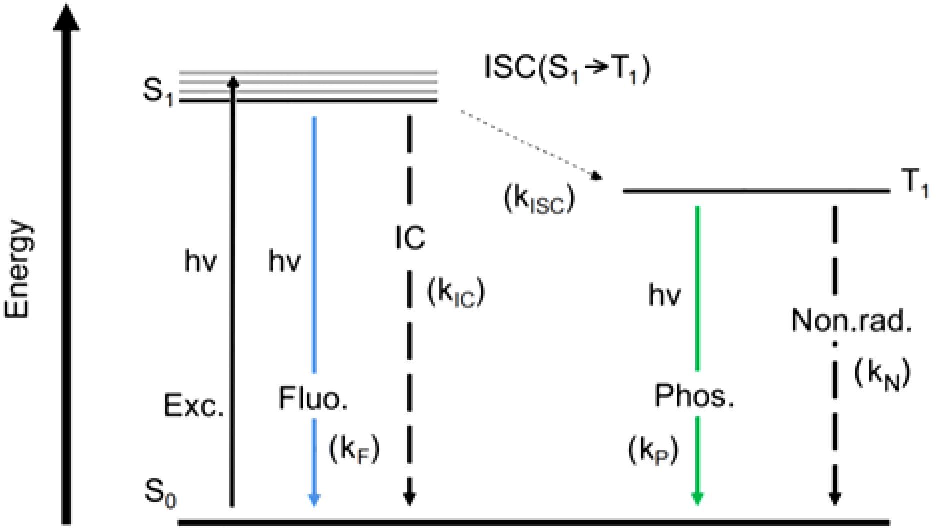
13. **Deduce** the value of φ, at which the maximum chiral asymmetry is observed.

14. **Explain** the concentration behaviour of [LL] and [DD] when:
   a) $\varphi < \varphi_1$;
   b) $\varphi_1 < \varphi < \varphi_2$;
   c) $\varphi > \varphi_2$.

Problem 11. Fluorescence, phosphorescence and quenching

Molecules can exist in different electronic spin states known as singlet and triplet states. A singlet state has all electron spins paired (total spin, $S = 0$), while a triplet state has two unpaired electrons with parallel spins ($S = 1$). Transitions between these states govern important luminescence processes such as fluorescence (singlet → singlet) and phosphorescence (triplet → singlet). Understanding the decay kinetics of these excited states is essential for interpreting emission lifetimes, nonradiative losses, and designing efficient photonic materials.

When a molecule absorbs a photon ($h\nu$), an electron is promoted to a higher energy level, called an excited state (S_1). This process is called **excitation** (Exc. with rate I_{abs}). From this excited state (S_1), the molecule can lose energy in different ways. It can undergo **internal conversion** (IC with rate constant k_{IC}), where energy is lost non-radiatively as heat through changes in vibrational states. Alternatively, it can return to the ground state (S_0) by emitting a photon ($h\nu$); this process is called **fluorescence** (Fluo. with rate constant k_F). The excited singlet state (S_1) can also convert to an excited triplet state (T_1) through **intersystem crossing** (ISC with rate constant k_{ISC}). The triplet state has a longer lifetime, and the molecule may return to the ground state (S_0) by emitting a photon ($h\nu$), known as **phosphorescence** (Phos. with rate constant k_P), or by losing energy as heat through **non-radiative decay** (Non.rad. with rate constant k_N). If a quencher is present, it can accept the excitation energy, causing the molecule to return to its ground singlet state (S_0) without emitting light.



1. **Choose** the correct option: "For electron excitation, the wavelength of the absorbed photon must be …":
   a) higher than that of the fluorescence wavelength
   b) equal to that of the fluorescence wavelength
   c) lower than that of the fluorescence wavelength

Internal Conversion (IC) is usually more common than Intersystem Crossing (ISC) because it is the spin allowed process (singlet → singlet).

2. **Choose** the correct option: "To increase k_{ISC} the energy gap between S_1 and T_1 states should be …":
   a) higher
   b) lower

To predict light emission (*e.g.*, fluorescence) quantum yield, it is necessary to know the decay mechanism of the excited singlet state:

Absorption:	$S + h\nu \rightarrow S^*$	$r_{abs} = I_{abs}$
Fluorescence:	$S^* \rightarrow S + h\nu$	$r_F = k_F[S^*]$
Internal conversion:	$S^* \rightarrow S$	$r_{IC} = k_{IC}[S^*]$
Intersystem crossing:	$S^* \rightarrow T^*$	$r_{ISC} = k_{ISC}[S^*]$

3. **Find** the order of $[S^*]$ and the observed rate constant (k_{obs}) for the rate of decay of the excited singlet state, and the emission lifetime (τ_0) after the incident light is turned off. *Note: Emission lifetime (τ_0) is the time required to reduce $[S^*]$ by e (≈ 2.72) times.*

Fluorescence quantum yield (φ_0) is the fraction of the excited singlet state molecules that **emit a photon** (fluoresce).

4. **Derive** φ_0 as a function of the rate constants (k_F, k_{IC}, and k_{ISC}).

To control fluorescence efficiency in devices (e.g., LEDs, sensors), quenching is very helpful. Fluorescence quenching is the promotion of non-radiative decay of the excited state. In a collision quenching, an excited fluorophore collides with a quencher, and due to the energy loss, there is no light emission:

$$S^* + Q \rightarrow S + Q \qquad\qquad r_Q = k_Q[S^*][Q]$$

5. **Derive** the fluorescence quantum yield (φ) as a function of the quencher concentration.
6. In the Stern–Volmer equation, the relationship between φ_0/φ and [Q] is linear. **Derive** this relationship.

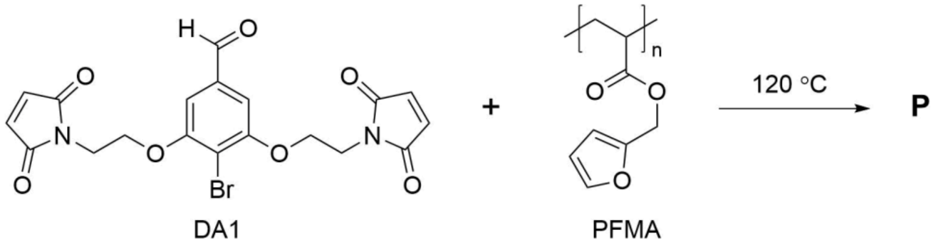
In one experiment, quenching of tryptophan fluorescence in a solution by dissolved O_2 gas was monitored by measuring emission lifetimes at 348 nm in aqueous solutions.

$[O_2]$ / M	0	2.3×10^{-2}	5.5×10^{-2}	8×10^{-2}	10.8×10^{-2}
τ_0 / s	2.6×10^{-9}	1.5×10^{-9}	0.92×10^{-9}	0.71×10^{-9}	0.57×10^{-9}

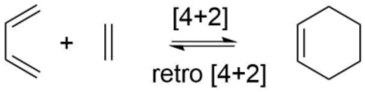
7. **Determine** the quenching rate constant (k_Q) for this process.
8. **Calculate** the concentration of O_2 at which half of the excited tryptophan molecules are quenched.

Phosphorescence can also be influenced by singlet state quenchers. There are several approaches to reducing the likelihood of collisions between quencher and phosphorescent molecules. One effective method is to increase the rigidity of the molecule. This rigidity also reduces non-radiative decay by limiting additional vibrational modes and suppressing vibrational relaxation pathways.

Phosphorescence molecule DA1 was heated at 120 °C with PFMA and the product obtained suppressed the non-radiative decay (k_N) of the triplet state.



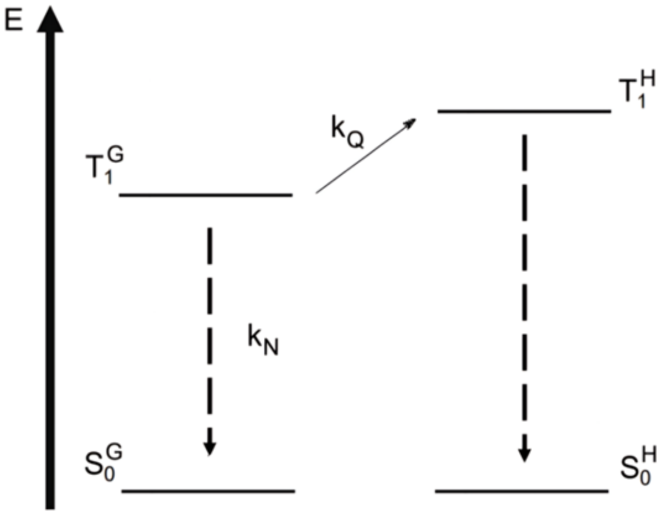
9. **Draw** the structure of **P**, which is an adduct of the following [4+2] cycloaddition reaction:



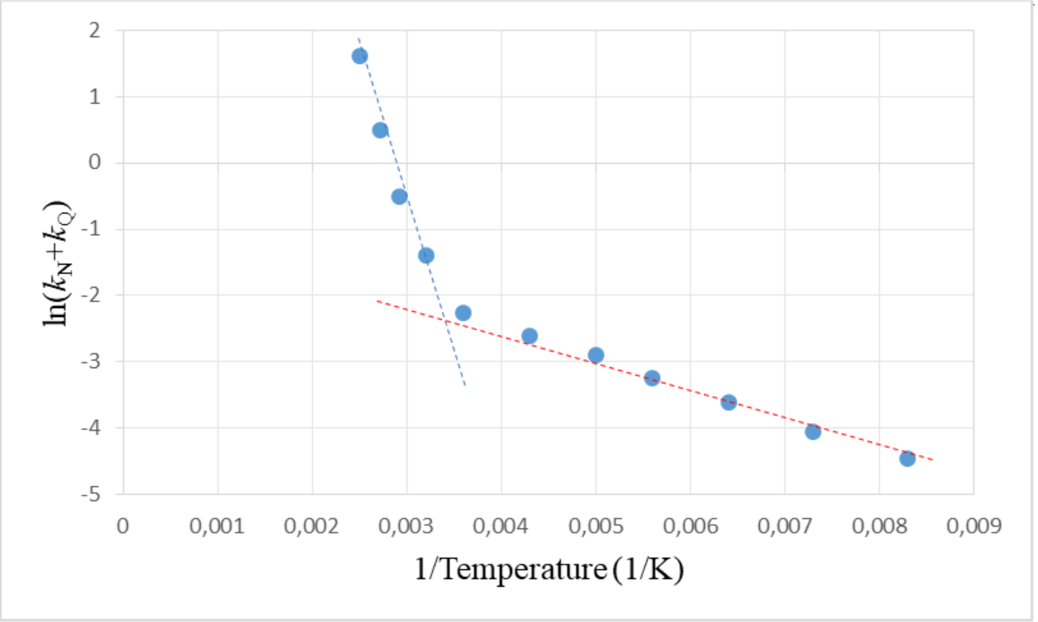
At low temperatures, the process of quenching is limited by an energy barrier:

$$\Delta E = E(T_1^{\text{Host}}) - E(T_1^{\text{Guest}})$$

Therefore, non-radiative decay can be suppressed by maintaining rigidity with a polymer matrix under this condition. At higher temperatures, vibrations increase and triplet energy transfer from the triplet state of the phosphorescent guest molecule to the polymer matrix (host) becomes possible:



ln(k_N+k_Q) as a function of (1/Temperature) is given below:



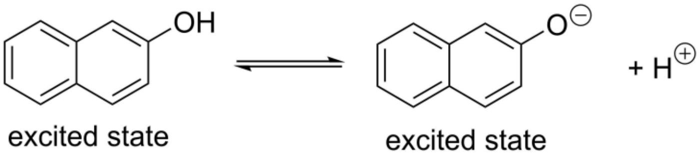
k_N – rate constant of the non-radiative decay of the triplet state.

k_Q – quenching of the triplet state by T–T energy transfer from the guest to the host molecule.

10. **Estimate** the energy barrier ΔE of T–T energy transfer.

Problem 12. Photoacids

Photoacids form a unique class of molecules that undergo a major change in their pK_a values between their ground and excited states. The loss of a proton is facilitated by light excitation. This phenomenon is also called excited state proton transfer (ESPT).

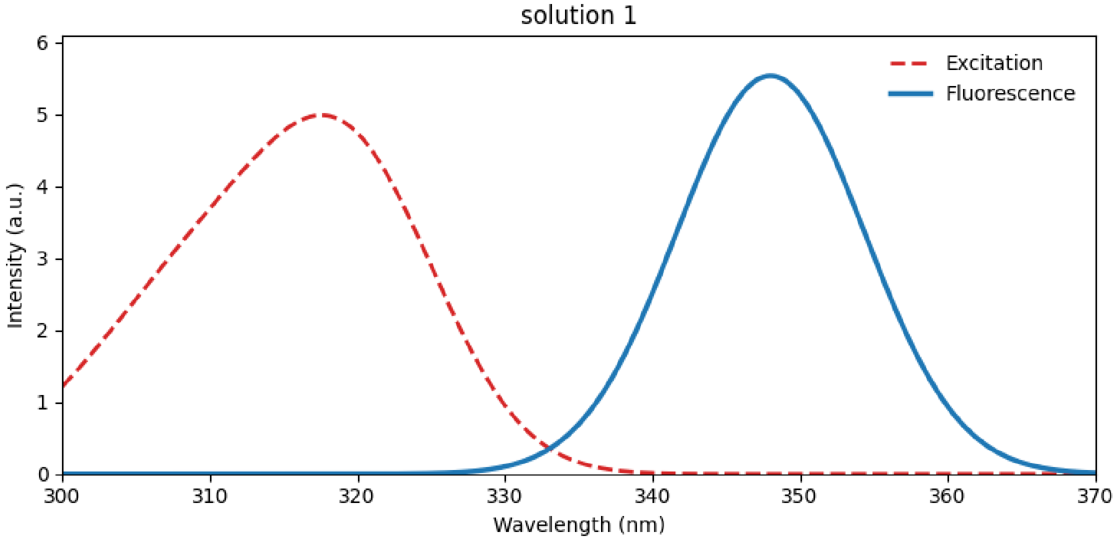


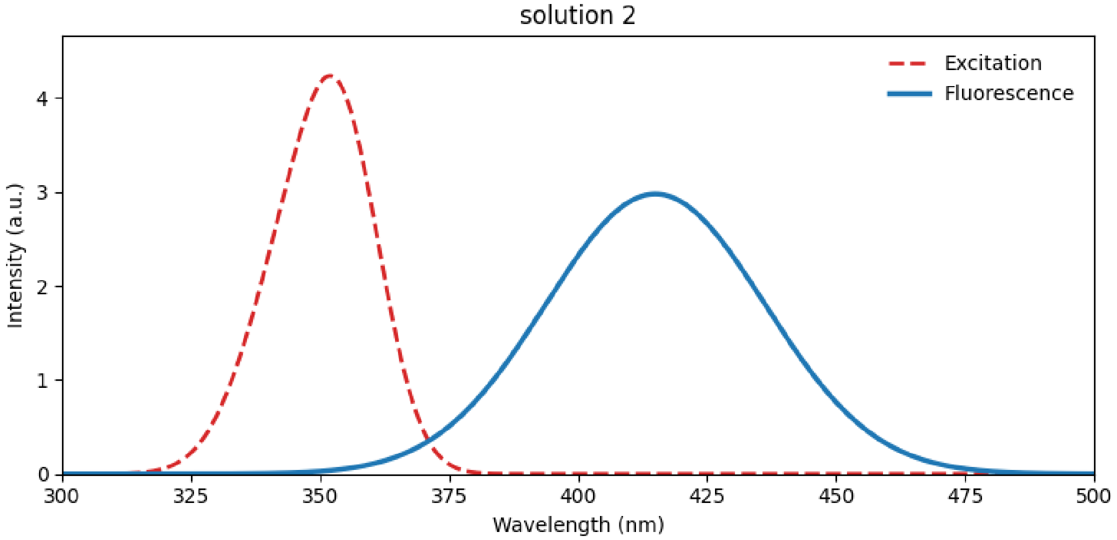
Once, Khoja Nasreddin, who was bored with his usual routine, decided to investigate 2-naphthol, a simple molecule that exhibits the ESPT phenomenon. He prepared five 2-naphthol solutions (each of 25 mL) in different pH media, using a 2.1 mM 2-naphthol solution (**solution Z**) and measured light absorbance (l = 1.0 cm) at 344 nm:

Solution	**1**	**2**	**3**	**4**	**5**
Composition	5 mL **Z** + 20 mL 0.02 M HCl	5 mL **Z** + 20 mL 0.02 M NaOH	5 mL **Z** + 20 mL NH_4Cl/NH_3 (3:1 buffer)	5 mL **Z** + 20 mL NH_4Cl/NH_3 (2:1 buffer)	5 mL **Z** + 20 mL NH_4Cl/NH_3 (1:1 buffer)
Absorbance	0.219	0.440	0.256	0.273	0.300

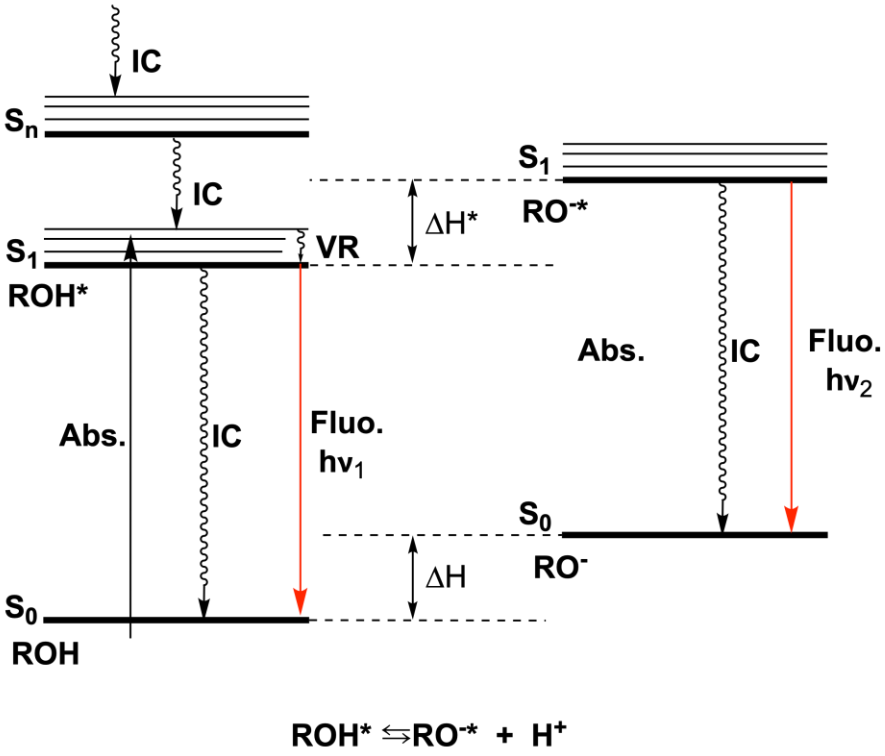
1. **Find** the final concentration (mol·L^{-1}) of 2-naphthol in solutions **1**–**5**.
2. **Calculate** the molar extinction coefficients, ε_{ArO^-} and ε_{ArOH} at this wavelength.
3. **Plot** the graph pH vs log $\frac{[ArO^-]}{[ArOH]}$. **Estimate** the pK_a of 2-naphthol if pK_b(NH_3) = 4.75.

To check the ESPT phenomenon in 2-naphthol, Khoja Nasreddin recorded the fluorescence spectra of solutions **1** and **2**:





Dissociation of a photoacid in the ground and excited states can be described with a combined Förster and Jablonski diagram:



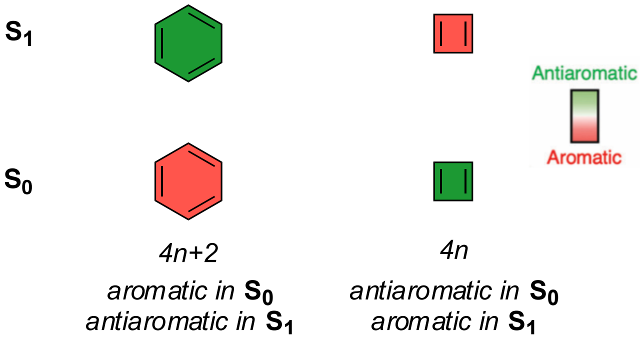
S_0, S_1, …, S_n stand for different electronic states, each of them having multiple vibrational levels. ΔH and ΔH^* indicate changes in the enthalpy of dissociation in the ground and excited states, respectively. ν_1 and ν_2 are the frequencies of S_1 to S_0 (0–0) transitions for the photoacid and its deprotonated form, respectively. Internal conversion (IC) is a spin-allowed non-radiative transition from a higher to a lower electronic state and does not show any peak in the spectra. Fluorescence, on the other hand, is the emission of a photon during a transition between the states with the same spin quantum number (for example, $S_1 \rightarrow S_0$). Vibrational relaxation (VR) is a non-radiative transition from a higher vibrational level to a lower one in the same electronic state of the molecule. VR is very fast, so it occurs right after light absorption and precedes other radiative processes like fluorescence.

In practice, transition wavelengths λ_1 and λ_2 are commonly used instead of corresponding frequencies.

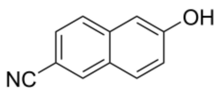
4. Assuming $\Delta H^{\circ *} - \Delta H^{\circ} = \Delta G^{\circ *} - \Delta G^{\circ}$, **derive** the relationship between ΔpK_a, λ_1, λ_2, and temperature. $\Delta pK_a = pK_a - pK_a^*$.

5. **Determine** the pK_a^* of the excited state (ES) of 2-naphthol at 298 K. **Calculate** $[ArOH^*]/[ArO^{-*}]$ at pH = 5.

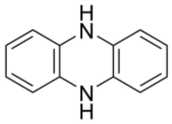
Khoja Nasreddin wondered whether every acid decreases its pK_a in its excited state. After consulting with Prof. Beruni, Khoja found out that this might relate to Baird's rules – the analogue of Hückel's aromaticity rules for excited states. According to Baird's rules, in the excited state, planar conjugated structures with $4n$ π-electrons are considered aromatic, and $(4n + 2)$ π-electrons are considered antiaromatic. In this terminology, photoacidity is the result of Baird-antiaromatic molecules "relieving stress" through deprotonation. This is made possible because of the excess charge redistribution from the anionic centre to the ring.



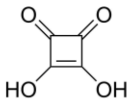
6. **Classify** the structures below as Baird-aromatic and Baird-antiaromatic. **Identify** those that could be used as potential photoacids.



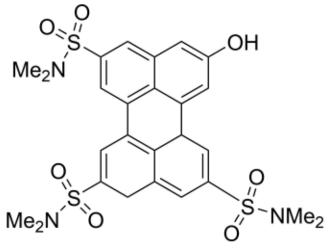
(a)



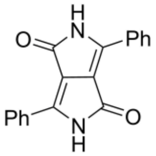
(b)



(c)



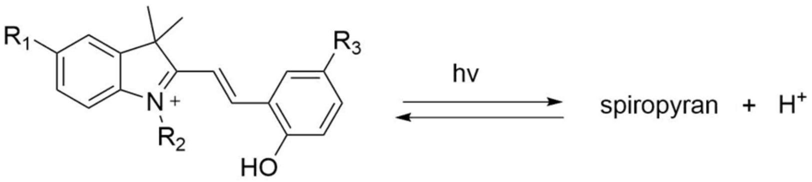
(d)



(e)

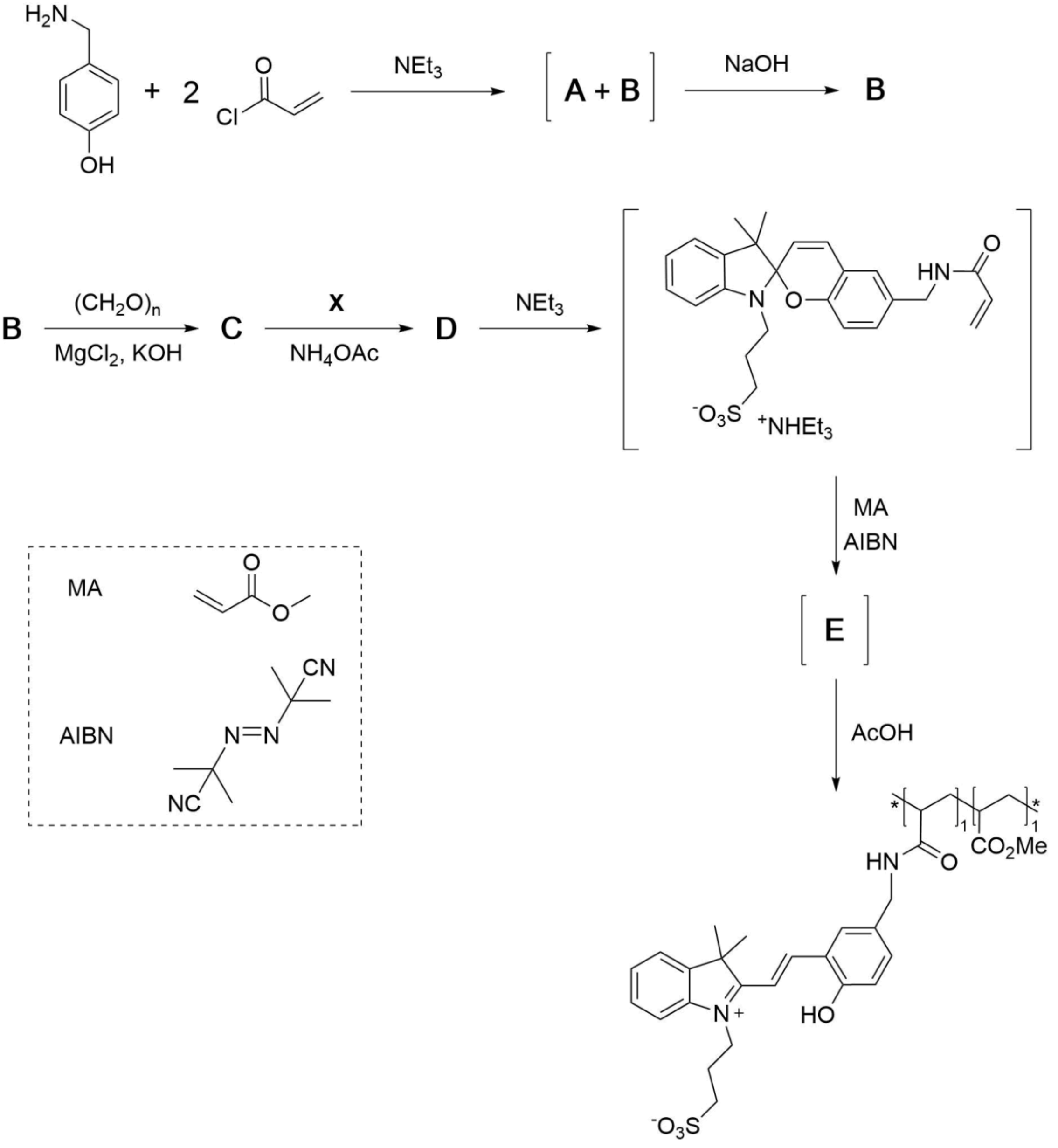
7. 2-Naphtalenecarboxylic acid behaves the opposite way when treated with light – its $\Delta pK_a = pK_a - pK_a^* < 0$. **Explain** this phenomenon.

Merocyanines are an excellent example of photoacidity, turning into spiropyrans under light irradiation.



8. **Draw** the structure of the spiropyran.

The synthesis of a merocyanine photoacid polymer is given below.



Hints:

- *Hydrolysis of compound **A** results in compound **B**.*
- ***C** can be converted into **C'** with a carboxylic acid moiety under NaClO$_2$.*

9. **<u>Draw</u>** the structures of **A**–**F** and **X**.

Problem 13. Interested in micelles?

Part 1. The forms of micelles

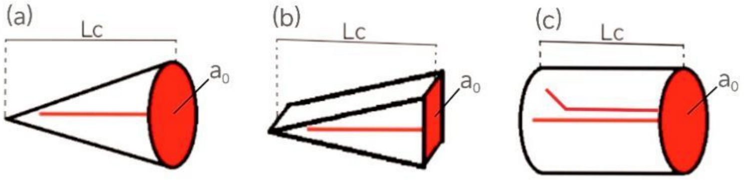
The form of a micelle can be described by the surfactant packing parameter, N_S. The packing parameter considers the volume and the length of the hydrophobic chain (V_C and L_C, respectively), and the equilibrium hydrophilic area per molecule at the aggregate interface, a_0:

$$N_S = \frac{V_C}{a_0 \times L_C}$$

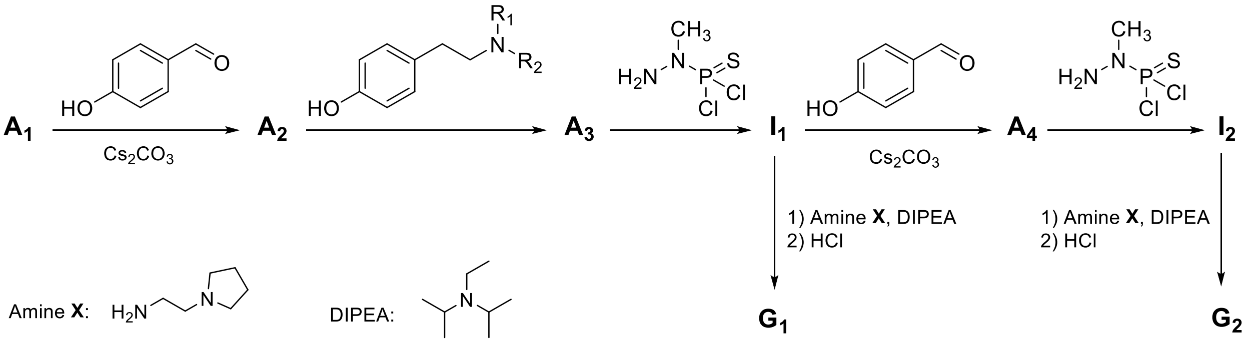
The table below shows the relationships between the surfactant packing parameter and the form of the structure:

Form of micelle	Sphere	Cylinder	Bilayer
N_S	$\leq \frac{1}{3}$	$\frac{1}{3} < \cdots \leq \frac{1}{2}$	$\frac{1}{2} < \cdots \leq 1$

1. If there are cone (a), rectangular wedge (b) and cylinder-shaped (c) surfactants, **match** them with their appropriate form of structure. **Prove** your answers by calculation.
Note: It is only an assumption. In fact, surfactants themselves are more complicated structures and can have shapes which are different from these three below.

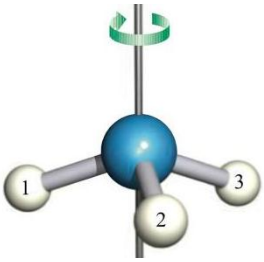


The phosphorus dendrons **G₁** and **G₂** act as charged surface active substances, although their structure differs from that of classical surfactants with a well-defined polar head and non-polar tail. They aggregate into clear structures and can be synthesised through the scheme below.



A₁ is a cyclic compound with a threefold symmetry axis, with mass fractions of P, N, and Cl of 26.73%, 12.09%, and 61.18%, respectively. With the addition of an appropriate number of equivalents of 4-hydroxybenzaldehyde to **A₁**, **A₂** is produced as the main product with only one Cl atom.

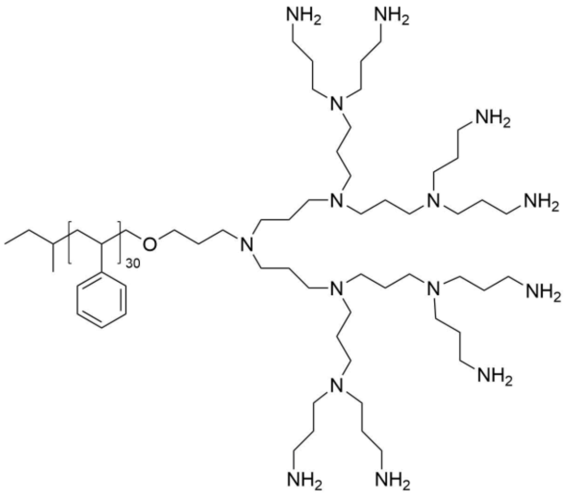
Note: A threefold symmetry axis is an imaginary line through an object around which it can be rotated by 120° and appear unchanged. For example, it is present in the molecule of ammonia.



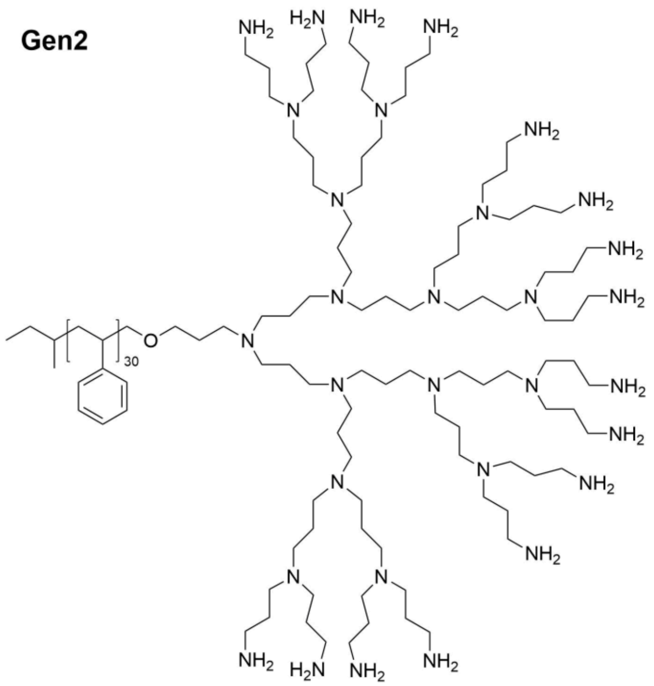
2. **Identify** structures $\mathbf{A_1}$–$\mathbf{A_4}$, $\mathbf{I_1}$–$\mathbf{I_2}$, and $\mathbf{G_1}$–$\mathbf{G_2}$.
3. Based on the synthesis pattern for $\mathbf{G_1}$ and $\mathbf{G_2}$ shown in the scheme, **calculate** how many cationic groups will be present in $\mathbf{G_{10}}$ if the same reaction stages are repeated.
4. By analysing the structure, **determine** which type of surfactants the $\mathbf{G_1}$ and $\mathbf{G_2}$ molecules resemble. **Indicate** the type of assembly they most likely show. **Specify** without calculations.
5. **Choose** the main purpose of adding HCl at the end of the synthesis:
   a) to hydrolyse unreacted P–Cl bonds for purification;
   b) to make the fluorescent label brighter;
   c) to create a water-soluble, positively charged form that will show biological activity;
   d) to rigidify the dendron structure via cyclisation.

In contrast to $\mathbf{G_1}$ and $\mathbf{G_2}$, in a series of amphiphilic polystyrene (PS)-poly(propylene imine) dendrimer diblock copolymers **Gen1–Gen3**, varying in head group size, the form of micelles within a polar medium depends on the generation number.

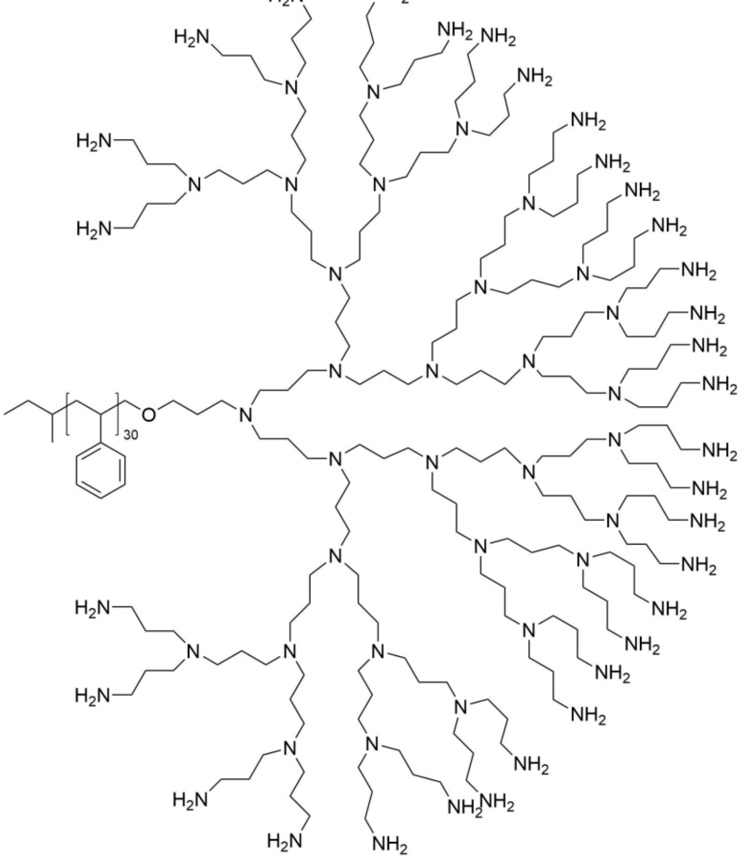
Gen1



Gen2



Gen3



6. **Propose** the forms of micelle in which **Gen1**, **Gen2** and **Gen3** are assembled in water, if each generation forms a distinct micelle type.

Part 2. Thermodynamics of micelle formation

The micelle formation process can be seen as an equilibrium between two states:

$$n[\mathrm{M}](\text{surfactant}) \rightleftharpoons [\mathrm{M}_n](\text{micelle})$$

$$K_n = \frac{[\mathrm{M}_n]}{[\mathrm{M}]^n} = e^{\frac{-\Delta_\mathrm{r}G^\circ_\mathrm{micelle}}{RT}}$$

where n is the aggregation number, i.e. the number of surfactant molecules (further referred to as monomers) in the micelle; [M], [M_n] reflect the concentrations of monomers and micelles, respectively.

7. **Derive** the equation for the fraction of monomers present in micelles as a function of [M], n and K_n.

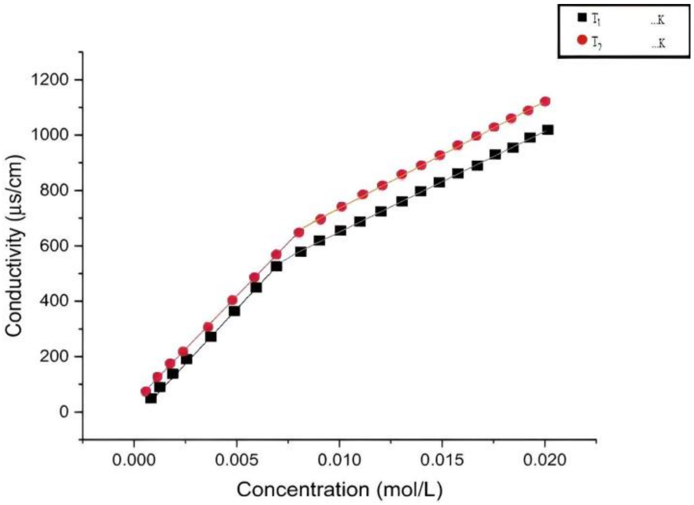
8. **Choose** which of the suggested conditions would lead to the maximum number of monomers inside the micelles. **Assume** that in all conditions the equilibrium constants are the same and the monomer concentration is lower than 1 $\mathrm{mol{\cdot}dm^{-3}}$.

(a) Low concentration of monomers with low aggregation number.
(b) Low concentration of monomers with high aggregation number.
(c) High concentration of monomers with low aggregation number.
(d) High concentration of monomers with high aggregation number.

At the critical point, the monomers inside the micelles start to account for the largest fraction. This is the so-called "critical micelle concentration (*CMC*)", above which the monomers will mostly be within the micelles.

9. **Describe** *CMC* as a function of aggregation number.

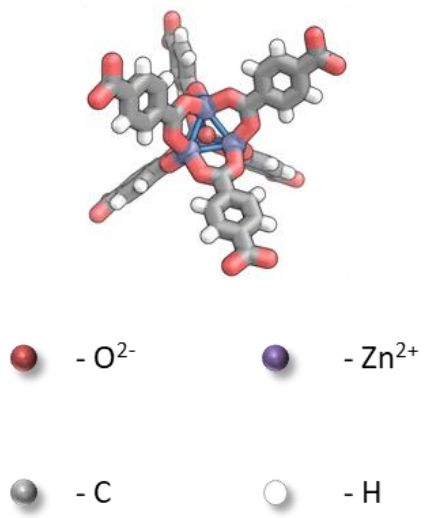
The dependence of specific conductivity on the concentration of surfactant monomers is shown in the picture below at temperatures T_1 and T_2:

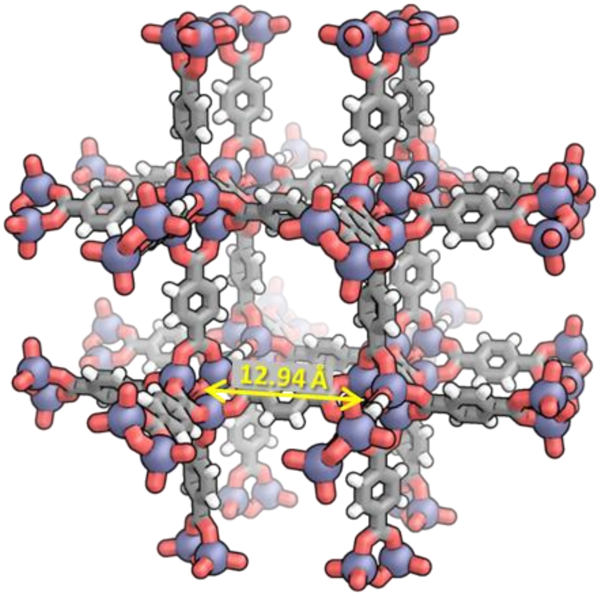


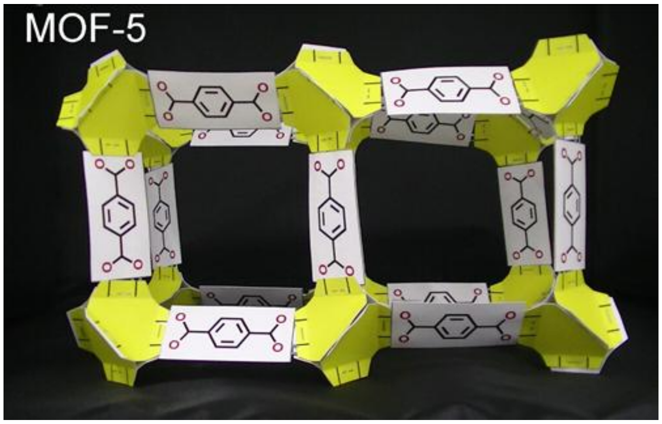
10. Assuming that the Gibbs energy of micelle formation, $\Delta_\mathrm{r}G^\circ_\mathrm{micelle}$, and the aggregation number remain constant, **compare** T_1 and T_2.

Problem 14. The basics of Metal–Organic Frameworks

The 2025 Nobel Prize in Chemistry was awarded to Susumu Kitagawa, Richard Robson, and Omar M. Yaghi for the development of Metal–Organic Frameworks (MOFs). MOFs are crystalline, nanoporous coordination polymers formed by linking metal nodes (ions or metal oxide clusters – secondary building units) with multidentate organic ligands. Their hallmark is the combination of molecular-level tunability with permanent porosity, producing materials whose structural and functional diversity exceeds that of classical zeolites or activated carbons. MOFs are 3D frameworks with a complex structure. Thus, building models is useful to understand elementary units. For that, please follow the instructions from *J. Chem. Educ.* 2025, 102, 2, 877–881 (https://doi.org/10.1021/acs.jchemed.4c01312). Normal paper, glue, and tape can also be used as alternatives. One of the famous materials in this field is MOF-5 constructed from zinc oxo-clusters (Zn_4O) and 1,4-benzenedicarboxylate (BDC) ligands.







1. From the figure above and/or constructed model, **write down** the empirical formula for the repeating unit.

The first synthesis of MOF-5 was carried out by mixing $Zn(NO_3)_2$, H_2BDC (benzene-1,4-dicarboxylic acid), and triethylamine in a *N,N*-dimethylformamide/chlorobenzene mixture. Small amounts of water were added to initiate oxo-anion formation. The crystals formed were filtered and activated (dried) in a vacuum oven.

2. **Write down** the chemical equation of MOF-5 formation.

Elemental analysis of the crystals before activation showed elemental content (by mass): C – 44.21%, H – 5.02%, N – 7.64%. In mass spectra of crystals during heating up to 350 °C, four *m/z* signals were detected: 73, 77, 112, and 114. The relative intensity of the latter two signals was 3:1.

3. **Write down** the empirical formula of the crystals before activation.

MOF-5 has a high surface area of 2900 $m^2 \cdot g^{-1}$. The distance between oxygen atoms in the cluster centres was found to be 12.94 Å.

4. **Find** the free volume (which is not occupied by framework atoms) and experimental surface area (m^2) of one cubic cell. **Account** only **for** Van der Waals radii of atoms (r), which are as follows:

Element	Zn	H	C	O
***r* / Å**	1.5	1.2	1.7	1.5

5. To be correct, there are two pores with diameters of 11.0 Å and 15.1 Å. **Find** the elementary unit cell parameter *a*.

6. MOF-5 has been tested for hydrogen uptake. It was found that 10 g of the material can take up to 1 wt% of H_2 at 20 bar, 298 K. Assuming all the material is an ideal crystal, **calculate** excessive hydrogen uptake (how much excess hydrogen can be stored compared to pure hydrogen of the same volume without MOF-5 under the given conditions).

7. **Suggest** the ligands which might change the cubic cell volume down to 69.3%. The distance between the 1- and 4-carbons of a benzene ring is 2.69 Å.

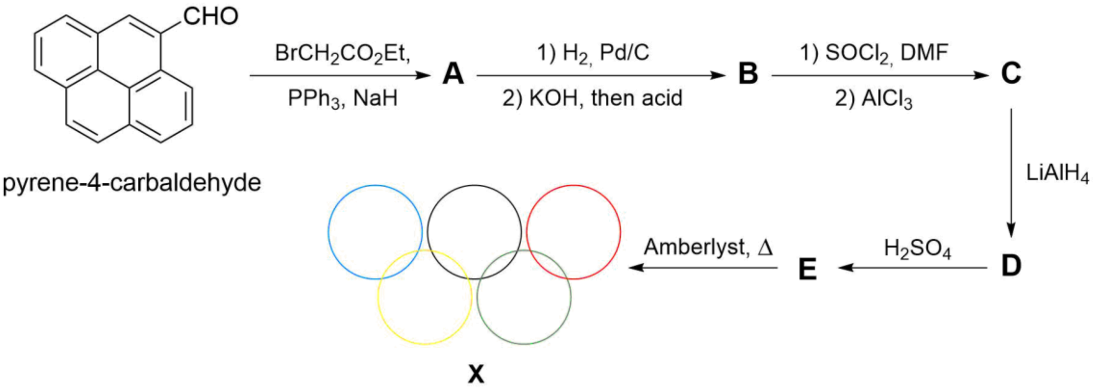
Problem 15. Microscopy in chemistry

Figure. 40 years ago, in 1986, Gerd Binnig and Heinrich Rohrer shared a Nobel Prize in Physics "for their design of the scanning tunneling microscope".

Scanning tunnelling microscopy (STM) is a microscopy method that lets us see individual atoms and molecules on the surface. It works by moving a sharp metal tip very close to the surface and applying a small voltage. In 1986, Binnig, Quate, and Gerber combined the principles of STM with a stylus profilometer to create Atomic Force Microscopy (AFM). While STM can only be used to image conductive samples, AFM does not require a current flow between the tip and the sample and can map a surface regardless of its conductivity.

In 2012, an article was published reporting that British chemists from the University of Warwick, together with a group of scientists from IBM Research, had captured an AFM image of a hydrocarbon molecule called **X** (see the image below). This image was subsequently recognised as the smallest logo for the London 2012 Summer Olympics.

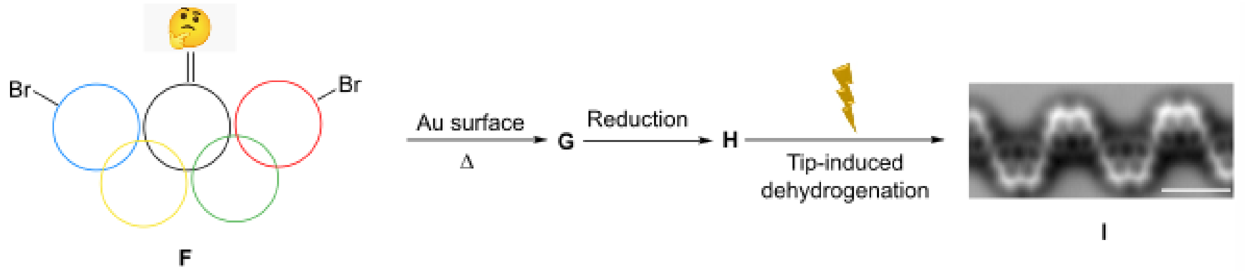
The synthesis of this hydrocarbon began with pyrene-4-carbaldehyde:



DMF – dimethylformamide, Amberlyst – protonic ion exchange resin. Initially obtained hydrocarbon molecule **E** easily isomerises to more symmetrical **X** under proton migration conditions.

1. **Draw** the structures of **A–E**, **X**. *Note: Molecule* **X** *possesses a twofold symmetry axis.*
2. Compound **X** was named as *...cene.* **Suggest** its full name.

X-based units can be used as building blocks of 1D quantum magnet chains via the formation of stable radical centres:



3. **Draw** the structures of compounds **F**–**I**, based on the following information about compound **F**:
   - A peak with *m*/*z* 411.89 was registered as the main peak in the mass spectrum;
   - Only 4 types of protons exist in **F**;
   - Compound **F** has a twofold symmetry axis.

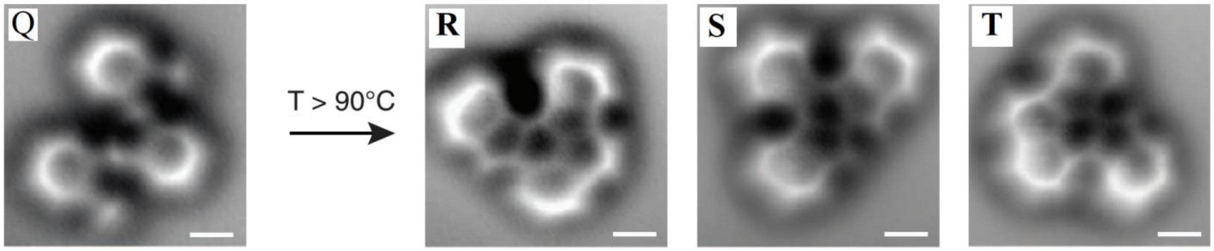
The beauty of such systems is that for finite chains like these, the lowest spin excitation energy, Δ_{LEE}, is inversely proportional to L:

$$\Delta_{\mathrm{LEE}} \leq J\frac{2\pi^2 S^2}{L}$$

where J is an exchange coupling parameter between the nearest neighbours, S is a spin number, and L is the number of repeating units in the quantum chain. Therefore, for a considerably large L, we can design spin fluids with a very low excitation gap, which can be considered gapless.

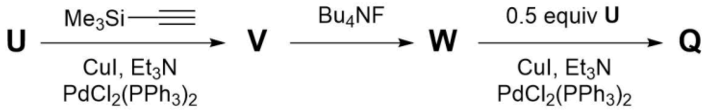
4. **Calculate** the value of L for which the quantum chain **H** becomes gapless at room temperature ($T = 298$ K), i.e. Δ_{LEE} becomes comparable with the average thermal energy $k_{\mathrm{B}}T$. Use $J = 38$ meV and $S = 1$.

AFM can be used to study the transformation of individual molecules during complex transformations. AFM was performed for the products of the thermally induced isomerisation reaction of **Q**. It is known that compound **S** has a twofold symmetry axis, **R** and **T** contain only 5- and 6-membered fused rings.



5. **Draw** the structures of **R**–**T**.

The molecule **Q** itself is obtained from **U** according to the following scheme:



Additionally:

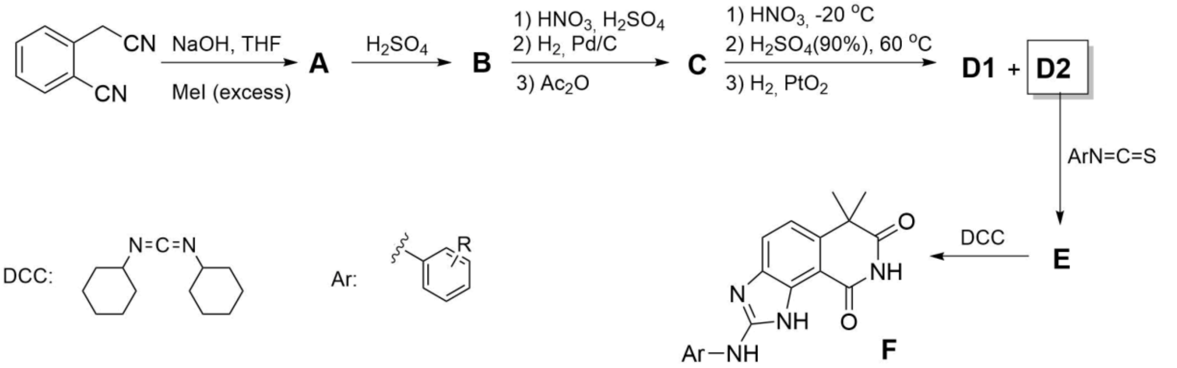
- In the mass spectrum, compound **U** has a molecular ion cluster of three peaks at *m*/*z* 234, 236, and 238 in an intensity ratio of 1:2:1;
- **U** has 2 types of protons in the aromatic region.

6. Based on the given AFM image and additional hints, **draw** the structures of **Q** and **U**–**W**.

Problem 16. Kinase inhibitors

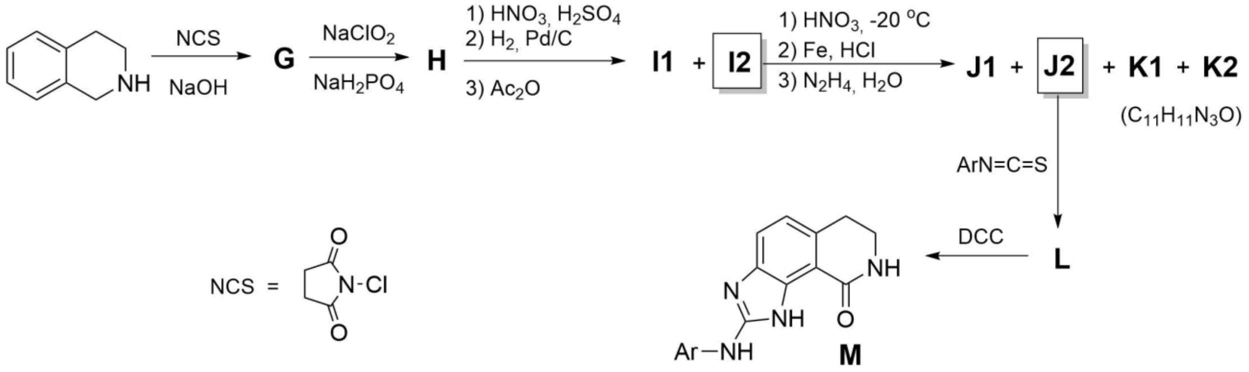
Kinases are enzymes controlling many processes in living things, including cell growth. Kinase inhibition has become an important direction for anticancer drug development. Two synthetic approaches to the class of *Lck*-kinase inhibitors are discussed.

A synthesis of promising inhibitor **F** is shown below:



1. **Draw** the structures of compounds **A**–**E**. *Note: Nitrogen content (by mass) in* ***B*** *is 7.40%.* ***D2*** *is separated from* ***D1****, and only* ***D2*** *is transformed into* ***F****.*

Another promising molecule **M** was synthesised. Compound **M** has a similar core to **F** but has a molecular mass lower than that of **F** by 42 atomic mass units. The first two steps were a benzylic oxidation by *N*-chlorosuccinimide and oxidation by NaClO$_2$. The transformation of **H** to **I1** and **I2** proceeds in the same manner as **B** to **C**, but with the formation of a mixture of isomers. The transformation of **I2** to the desired product **J2** is analogous to the conversion of **C** to **D2**. However, the order of the hydrolysis and reduction steps is switched. This change in order results in the formation of additional side products **K1** and **K2**.



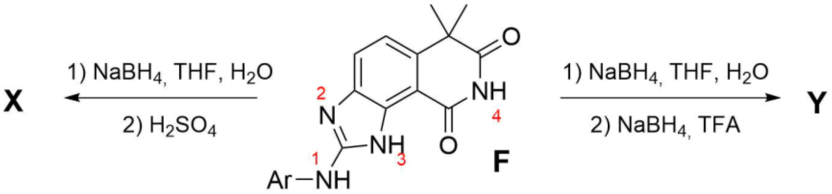
2. **Draw** the structures of **G**–**L**, including all isomers shown on the scheme.
3. **Explain** why only one product is formed during the reaction of **B**.

In the reaction of **I2** to **J2**, changing the order of the reduction and hydrolysis steps compared to the previous scheme was crucial to obtain **J**. The compound **W** could have formed following the sequence with the original order.

$$\mathbf{I2} \xrightarrow[\text{3) Fe, HCl}]{\text{1) } HNO_3, -20\ ^{\circ}C \text{; 2) } N_2H_4, H_2O} \mathbf{W}$$

4. **Draw** the structure of product **W** (21.52% nitrogen by mass).

The molecules obtained showed high potency with a specific kinase enzyme. However, two transformations of **F** were also carried out to give molecules **X** and **Y**, the molecular formulae of which differ only by two hydrogen atoms.



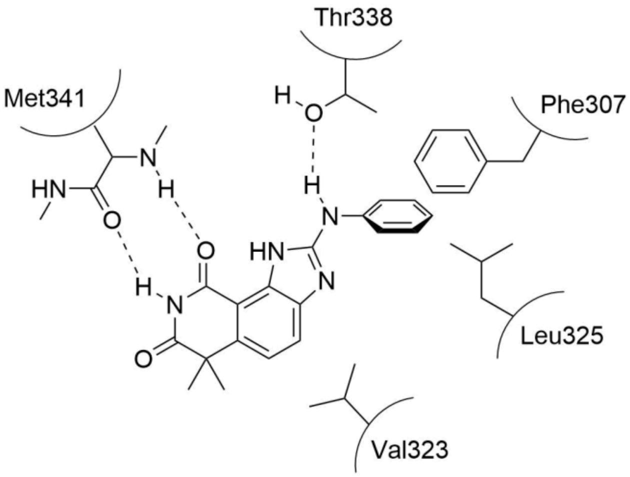
5. **Draw** the structures of **X** and **Y**. **Y** is a homologous compound to **M**. *Note: The number of* sp^2*-hybridised carbons in* ***X*** *increased by one.*

The synthesised molecules were evaluated by IC_{50} (half-maximum inhibitory concentration), which is a quantitative measure that indicates how much of a particular drug is needed to inhibit *in vitro* a given biological process or biological component by 50%. The results are shown in the table below:

Compound	R	IC_{50} / μM
F	2,6-Cl_2	0.46
F	2-Cl	>40
F	2,4-Cl_2	>40
F	4-Cl	>40
F	2,6-Me_2	0.46
F	2,6-Et_2	>40

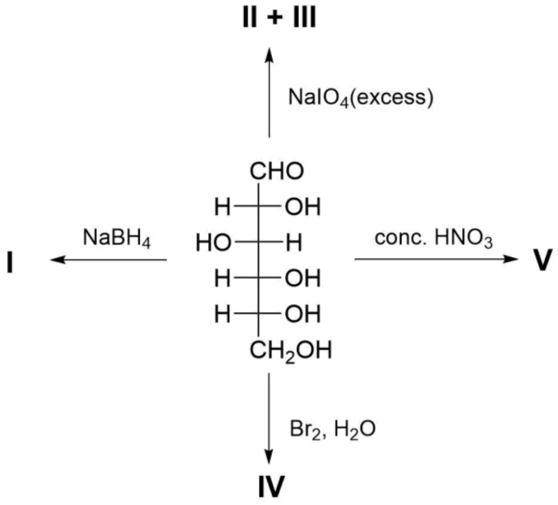
Compound	R	IC_{50} / μM
F (N3-Me)	2,6-Cl_2	1.1
F (N1-Me)	2,6-Cl_2	12.5
F (N4-Me)	2,6-Cl_2	9.6
Y	2,6-Cl_2	1.1
M	2,6-Cl_2	0.38
X	2,6-Cl_2	0.026

6. **Explain** the effect of the substituents in the aromatic ring if the coordination of **F** in a simplified enzyme pocket looks as follows:



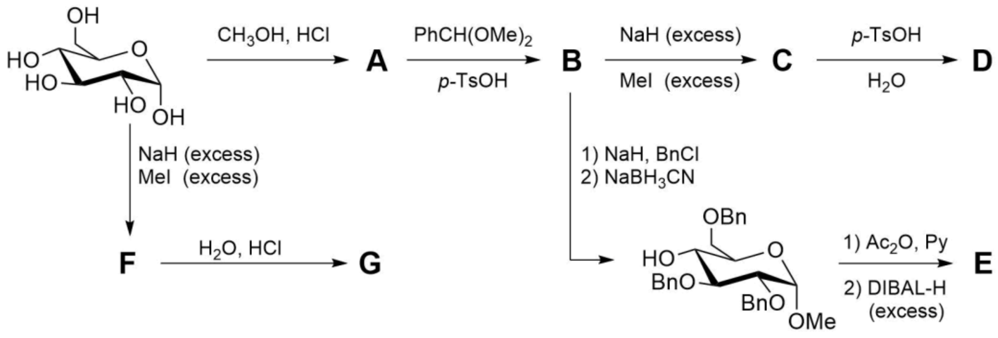
Problem 17. Carbohydrate chemistry

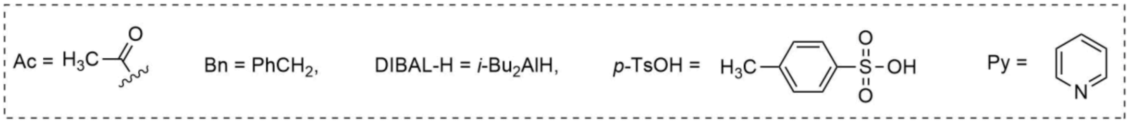
Carbohydrates are an essential class of chemicals that fuel living organisms with energy. The common saccharide glucose (Glc) exists in either a linear or cyclic form. It can undergo some common chemical transformations:



1. **Identify** products **I**–**V**.

In its cyclic form, glucose contains five hydroxy groups. Regioselective transformation of each position represents a challenge for chemists. Some synthetic transformations of α-glucopyranose are shown below:

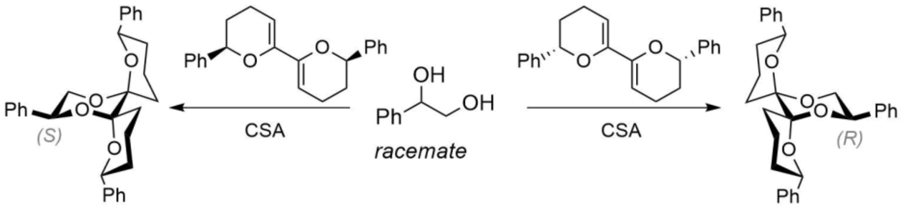


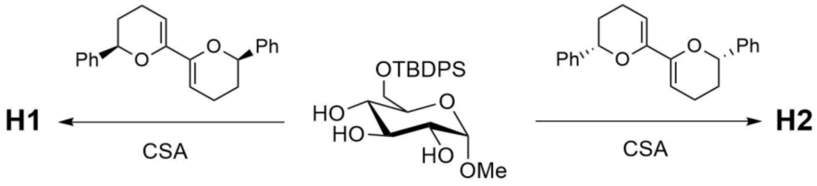


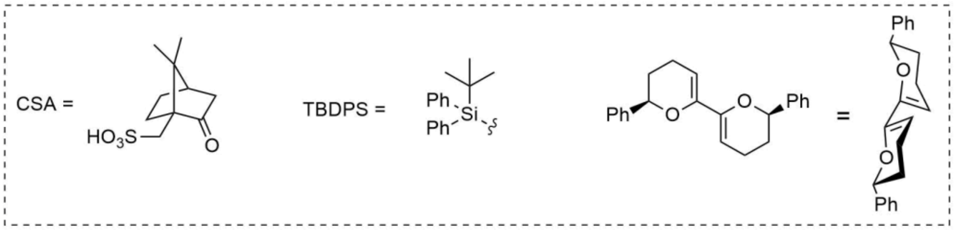
It is known that **A**, **G**, and **D** contain the same functional groups; **G** and **D** differ in formulae by CH_2; **B** and **C** contain three six-membered rings in their structures; there are two different CH_3-groups in **E**.

2. **Draw** the structures of **A**–**G**.

However, these transformations cannot differentiate neighbouring 2,3- and 3,4-diols, since in both cases the hydroxy groups are in a *trans*-configuration and occupy equatorial positions. A brilliant solution to this challenge was found by using the common tetrahydropyranyl protecting group in chiral form:

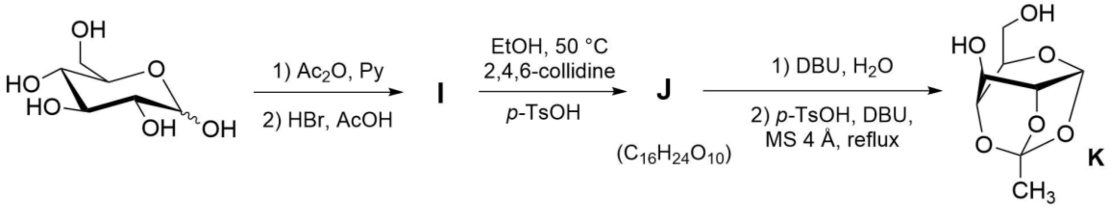


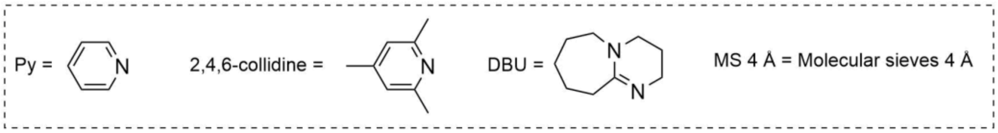




3. Based on the given example and considering that sterically bulky groups tend to occupy equatorial positions in chair conformations, **draw** the structures of **H1** and **H2**. For drawing **H1** you need to redraw the glucopyranose ring so that the 2,3-hydroxy groups are on the right side of the drawing. **Use** a chair conformation of the protecting group.

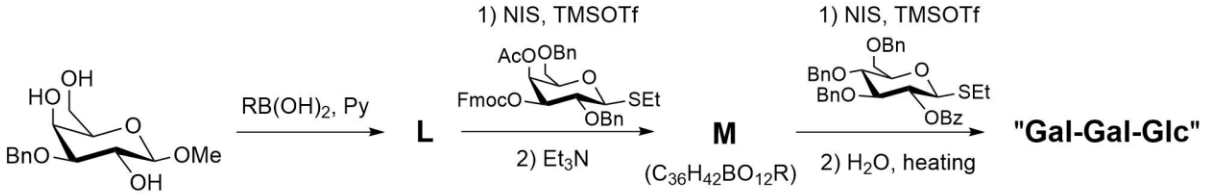
A very interesting example of a glucopyranose transformation led to adamantane-type cyclic compound **K** stepwise blocking three hydroxyl groups with only one protecting group:



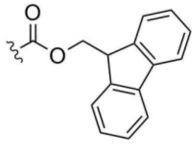


4. **Draw** the structures of **I** and **J**.

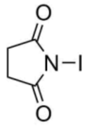
Another class of reagents that can form cyclic acetal-type structures is boronic acids. Their chemistry was successfully used to synthesise the trisaccharide “**Gal-Gal-Glc**” derivative:



Fmoc =



NIS =



TMSOTf = $CF_3SO_3SiMe_3$

5. **Draw** the structures of **L**, **M**, and trisaccharide “**Gal-Gal-Glc**”. Do not specify the configuration of the newly formed anomeric centres.

Problem 18. Inositol

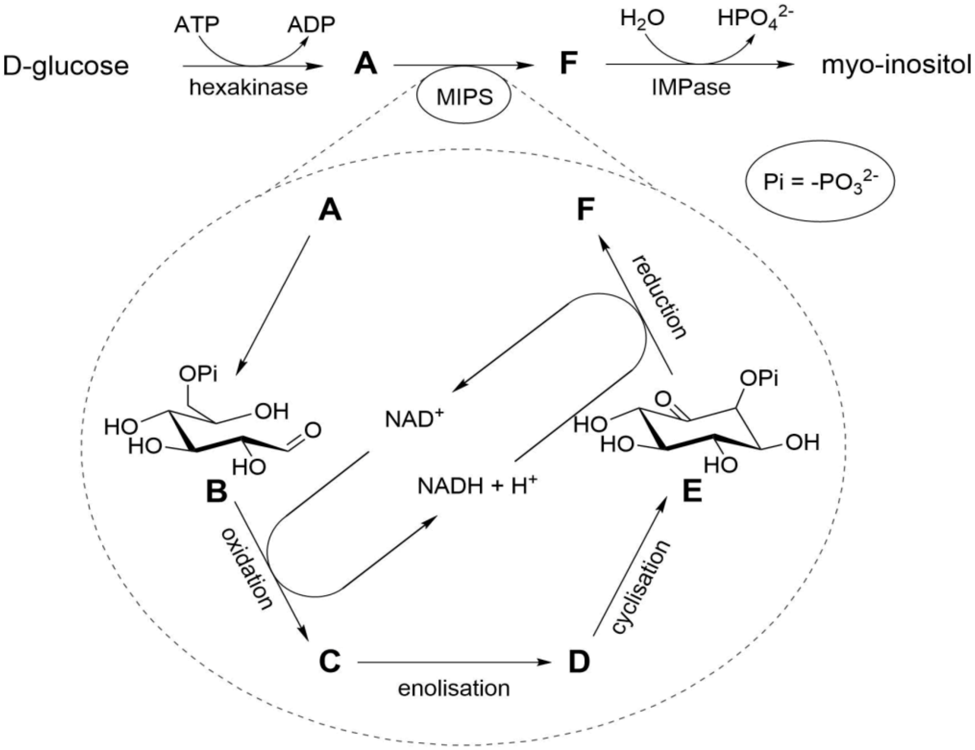
In Uzbekistan, people greatly appreciate foods that are both rich in flavour and highly nutritious. Rice is one of the main ingredients in Uzbek cuisine and forms the basis of many traditional dishes, such as *pilaf* and *mastava*.

Cooked rice consists of approximately 29% carbohydrates by mass, primarily in the form of α-amylose, which is a polysaccharide composed of α-D-glucose units linked through α(1→4) glycosidic bonds. Carbohydrates are the primary source of energy for humans and occur in nature in a wide variety of forms.

1. a) **Draw** the most stable chair conformation of D-glucose.
   b) **Draw** the structure of the repeating unit of α-amylose in Haworth projection.
2. **Calculate** the number of repeating units that one amylose chain with a molar mass of 243 kDa contains.

Glucose, which is a well-known representative of carbohydrates in living organisms, can be converted into many different biomolecules.

One such compound is **myo-inositol** (one of the isomers of inositol), whose biosynthesis pathway is shown below:

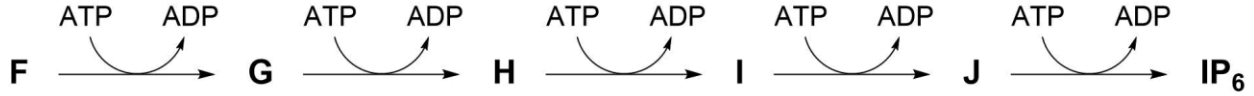


ATP = Adenosine Triphosphate, ADP = Adenosine Diphosphate

3. **Draw** one of the chair conformations for each of **A**–**F** and **myo-inositol**.
   *Note: Molecules of **F** and **myo-inositol** contain a plane of symmetry.*
4. **Draw** the structures of all inositol stereoisomers and **indicate** which of them are chiral.

Many plants synthesise myo-inositol hexaphosphate (**IP6**) as a phosphorus storage compound. Studies of various plant species have revealed that the biosynthesis of **IP6** proceeds through stepwise phosphorylation of compound **F**.

Below is the biosynthesis pathway of **IP6** that occurs in *Dictyostelium*:



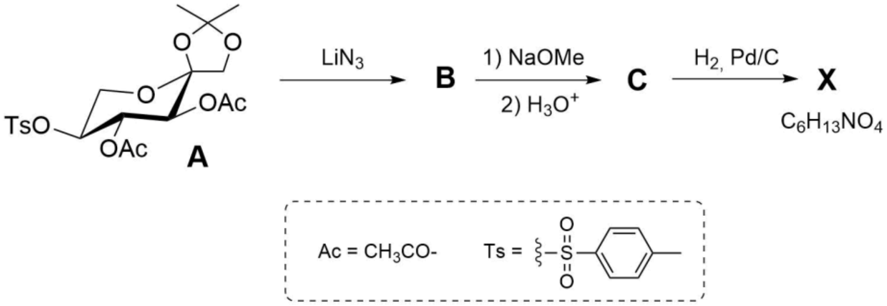
5. **Determine** the maximum number of **myo-inositol diphosphate** (**G**) isomers that can be formed in this pathway. **Draw** their structures and **indicate** which of them are chiral/achiral.
6. **Determine** the maximum theoretically possible number of isomers of **myo-inositol diphosphate** (that do not necessarily originate from **F**). **Draw** their structures and **indicate** which of them are chiral/achiral.

Phosphorylation of **myo-inositol** occurs according to the following rule: *the most distant accessible hydroxyl group relative to the previously phosphorylated hydroxyl group is phosphorylated.*

7. **Draw** the structures of **G**–**J** and **IP6**. **Indicate** the structures that contain a plane of symmetry.

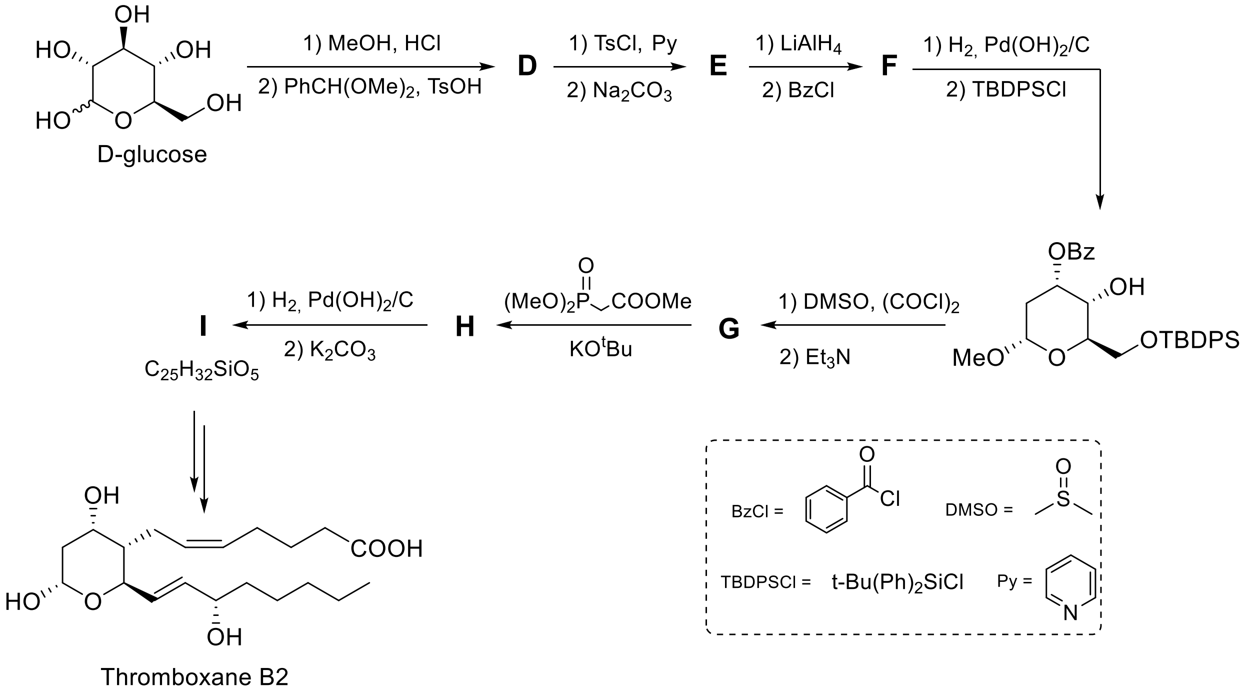
Problem 19. Carbohydrates as chiral pools

Asymmetric synthesis has attracted significant attention in recent years thanks to advances in asymmetric catalysis. However, using chiral auxiliary groups in the substrate structure remains an efficient approach for chiral total synthesis, especially when chiral substrates are readily available. Different monosaccharides are handy starting materials in natural product synthesis. For example, in 1985, the first synthesis of compound **X** from L-*sorbose* derivative **A** was published.



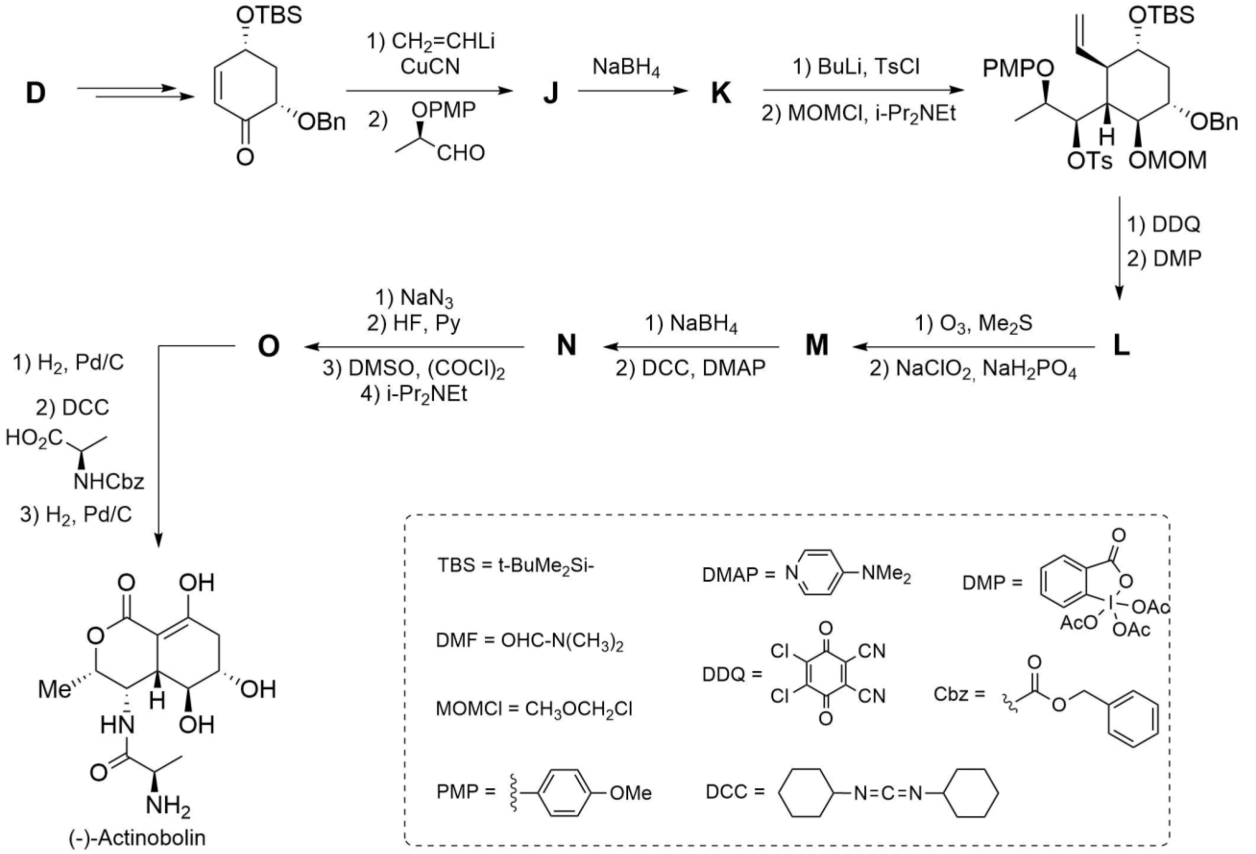
1. **<u>Draw</u>** the structures of compounds **B**, **C**, and the most stable diastereomer of **X,** including stereochemistry. *Note:* **X** *contains a five-membered heterocycle, and the formation of* **X** *proceeds through an imine intermediate.*

Thromboxane B2 (TxB2) is a stable, inactive breakdown product of *Thromboxane* A2 (TxA2), a potent molecule that causes blood platelets to clump (aggregate) and blood vessels to constrict. It was synthesised from D-*glucose*. The following scheme shows the synthesis of key intermediate **I**.



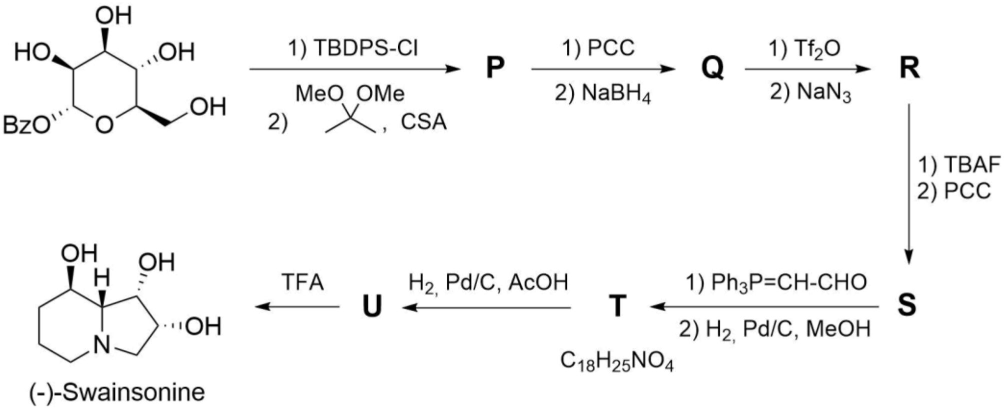
2. **Draw** the structures of **D**–**I** with stereochemistry. *Note: Acetal/ketal formation involves a primary alcohol when it is available.*
3. **Determine** in what order the following combinations of reagents (Z_1–Z_4) must be used to obtain *Thromboxane* B_2 from **I**:
   Z_1 – 1) PCC ($PyH^+ \cdot CrO_3Cl^-$), 2) $Ph_3P^+CH^-C(O)(CH_2)_4CH_3$;
   Z_2 – 1) BzCl, 2) $BuN_4^+F^-$;
   Z_3 – 1) $NaBH_4$, 2) K_2CO_3, 3) HCl;
   Z_4 – 1) $(i\text{-}Bu)_2AlH$, 2) $Ph_3P^+CH^-(CH_2)_3COOH$, 3) CH_2N_2.

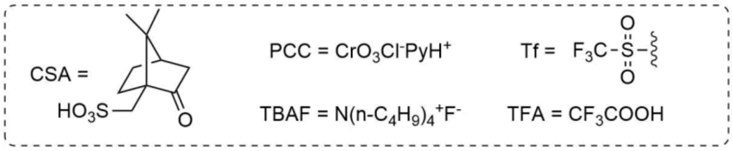
Using compound **D**, it is possible to synthesise other chiral molecules. (+)-*Actinobolin*, isolated from the culture fluid of *Streptomycin*, has a broad antibacterial spectrum and moderate antitumor activity. In the synthesis of (+)-*Actinobolin*, overall, five new stereocentres were formed using the chiral auxiliary.



4. **Draw** the structures of **J**–**O**, including stereochemistry.

Swainsonine, an indolizidine alkaloid, was synthesised using benzyl-α-D-mannopyranoside. It is a potent inhibitor of Golgi apparatus alpha-mannosidase II, an immunomodulator, and a potential chemotherapeutic agent. The synthesis of (–)-*Swainsonine* is given below.

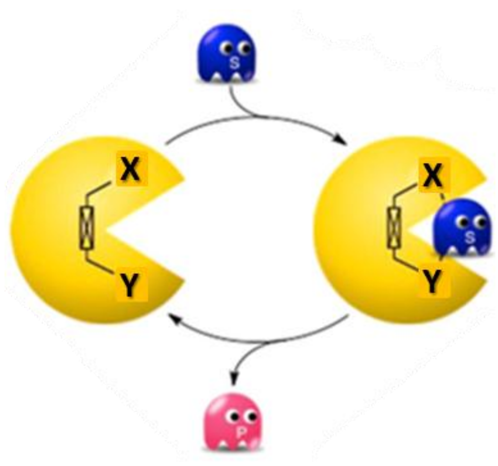




Hint: ***P*** *and* ***Q*** *are epimers.*

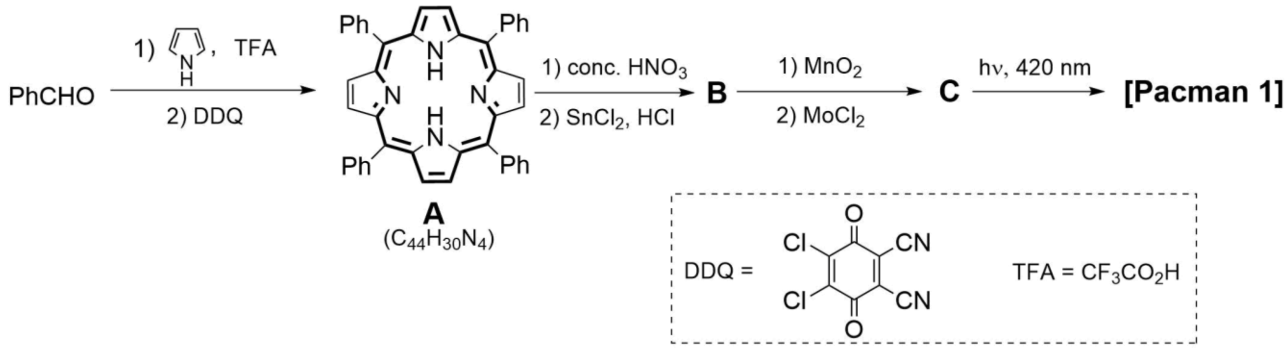
5. **Determine** the configurations of all stereocentres in (–)-*Swainsonine* and the starting material.
6. **Draw** the structures of **P**–**U**, including stereochemistry.
7. **Indicate** which stereocentres are preserved and which centres are newly formed.

Problem 20. Pacman, Hangman, Cageman



Scientists often look for analogies in culture to explain their discoveries. One example of such a comparison was the arcade video game *Pacman*, popular in the 1980s and 1990s, where the main task of the yellow ball (Pacman) is to eat all the dots on the map.

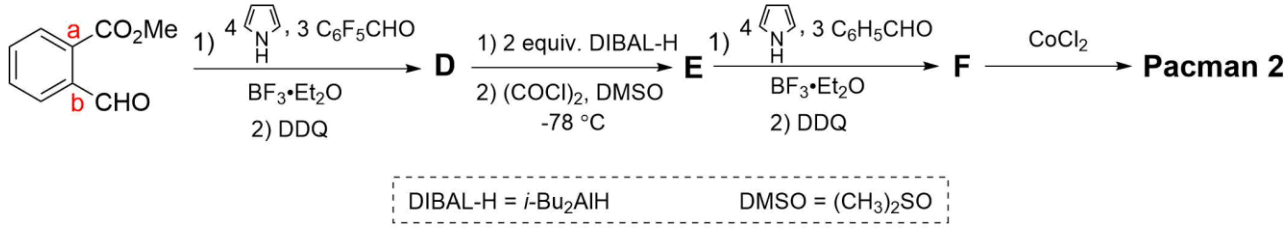
The same task must be performed by the molecules **Pacman 1** and **Pacman 2**, which must "eat" small molecules such as CO_2, N_2, or O_2, catalysing their transformations. Pacman molecules are two catalytically active structural fragments connected by a linker at a certain angle. Inspiration for such molecules comes from the active site of natural cytochrome P450, which contains a porphyrin macrocycle. Porphyrin is a macrocycle with an 18 π-electron Hückel aromatic system, which can coordinate metal ions in the middle. The synthesis of **Pacman 1** is shown below:



Note: C–H bonds of porphyrin rings are not involved in chemical transformations in this problem.

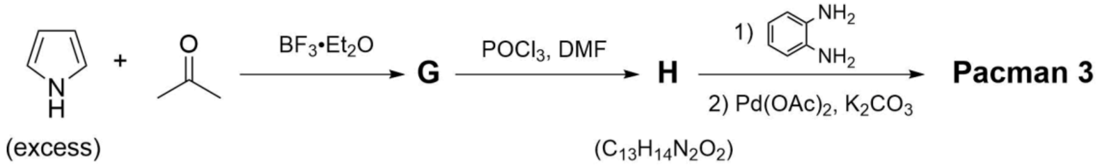
1. **Draw** the structures of compounds **B** and **C**, if it is known that neither of these compounds, nor **Pacman 1–2** contain oxygen. The nitrogen content (by mass) in compounds **B** and **C** is 11.12% and 9.70%, respectively.

Unfortunately, it was impossible to isolate **Pacman 1**. As a result, another approach to the stable **Pacman 2** with two different porphyrin rings was suggested. In the mass spectrum of this compound, the molecular ion peak was at *m*/*z* 1535.04.



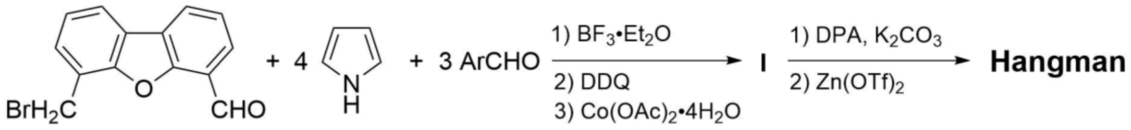
2. **Draw** the structures of compounds **D–F**, and **Pacman 1–2**. All molecules can be drawn flat. **Pacman 2** contains 7.678% of the metal by weight.
3. **Calculate** the distance between the two metals in **Pacman 2**, if the distance from carbon atoms a and b to the corresponding metals is the same and equals 4.7 Å. **Assume** that porphyrin and benzene rings are coplanar, and bond lengths and angles on benzene rings are ideal. *Note: The carbon–carbon bond distance in the benzene ring is 1.39 Å.*
4. **Draw** a 3D schematic structure of **Pacman 2** binding an oxygen molecule. *Note: You can simplify the macrocycle as a quadrilateral, only drawing out the linker and metals.*

Another version of a binuclear complex was synthesised as follows to yield **Pacman 3**:



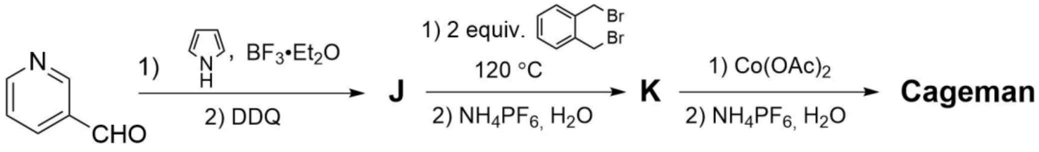
5. **Draw** the structures of compounds **G–H** and **Pacman 3**.

In the catalytic centre of cytochrome *c* oxidase, heme (a porphyrin-based iron-containing compound which forms the non-protein part of haemoglobin) and Cu atoms are placed in close proximity to ease the reduction of oxygen to water. Scientists attempted to mimic this pocket by synthesising a salt called **Hangman,** bearing two different metal ions:



6. **Draw** the structures of **I** and **Hangman**.

Another type of binuclear porphyrin complexes can be called **Cageman**. One of the examples of a **Cageman** synthesis is shown below. In the mass spectrum of the compound **K**, a series of signals at *m/z* 2667.0, 1261.0, 792.3, 558.0, and 417.4 were detected.

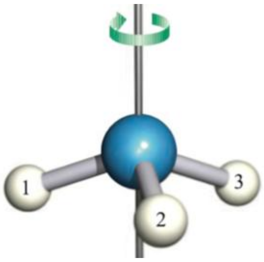


7. **Calculate** the molecular weight of **Cageman** (use integer values of atomic mass).

Problem 21. Catalyst C

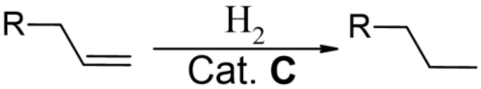
Compound **A** is a trihydrate which contains a metal of the 5th period, **M**, with a mass fraction of 39.09%. A 2.00 g sample of **A** was dissolved in ethanol. This ethanolic solution of **A** was added to an ethanolic solution of 7.96 g **B**. Compound **B** is a well-known ligand, which is composed of three elements: C, H, and P. It has 11.81% of P by mass and a threefold symmetry axis. The resulting mixture was refluxed for about two hours, yielding 6.18 g of a dark red compound **C** with 88% yield.

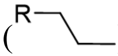
Note: A threefold symmetry axis is an imaginary line through an object around which it can be rotated by 120° and appear unchanged. For example, a threefold symmetry axis is present in the molecule of ammonia:



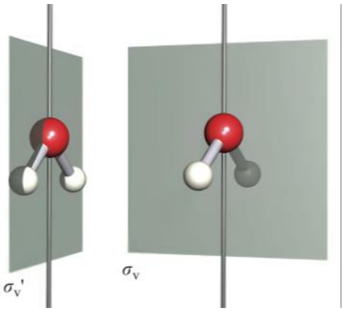
1. **Identify** the formulae of **A**–**C** and **M**, if each coefficient in the synthesis reaction is lower than 5.
2. **Write** the equation for the synthesis of compound **C**.

Compound **C** is a well-known catalyst in the hydrogenation of alkenes:

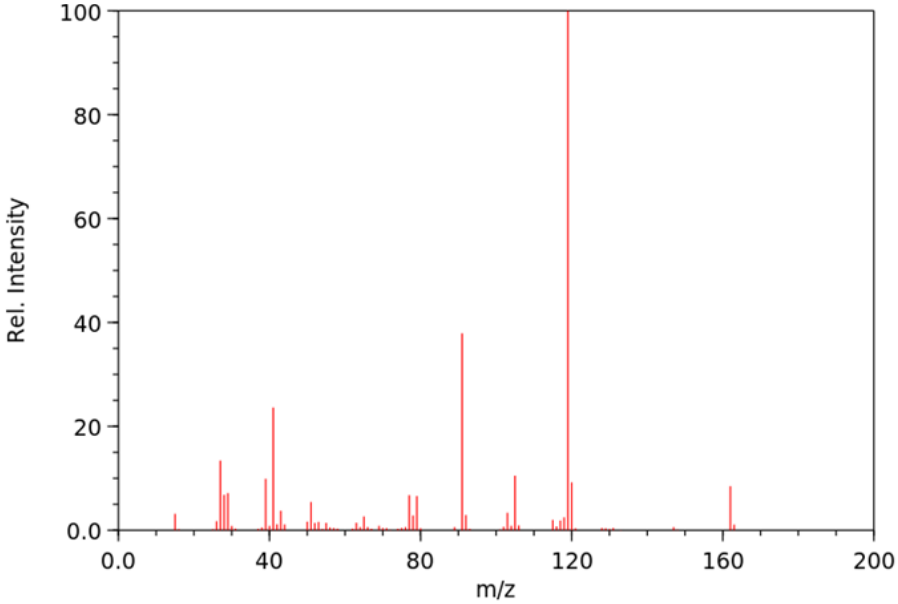


One example of the **product** () has a maximum of one plane of symmetry.

Note: A plane of symmetry is an imaginary plane that divides an object into two mirror-image halves. For example, the molecule of water has two planes of symmetry:

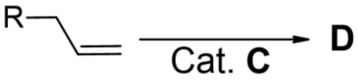


The mass spectrum of this **product** is shown below. The three peaks at *m*/*z* 41, 91, and 119 are noticeably more intense than the others, and two of them correspond to fragments containing different aromatic rings.



3. **Draw** the structures of the cations that correspond to those 3 peaks.
4. **Draw** the structure of the **product**.
5. **Find** the intensity ratio $[M]^+:[M+1]^+$ in the mass spectrum if the relative atomic mass of carbon is 12.011 and it consists of two stable isotopes.

In the catalytic system, a minor product **D** can be formed according to the reaction:

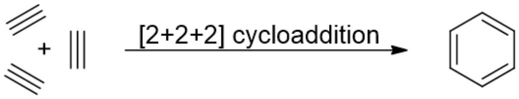


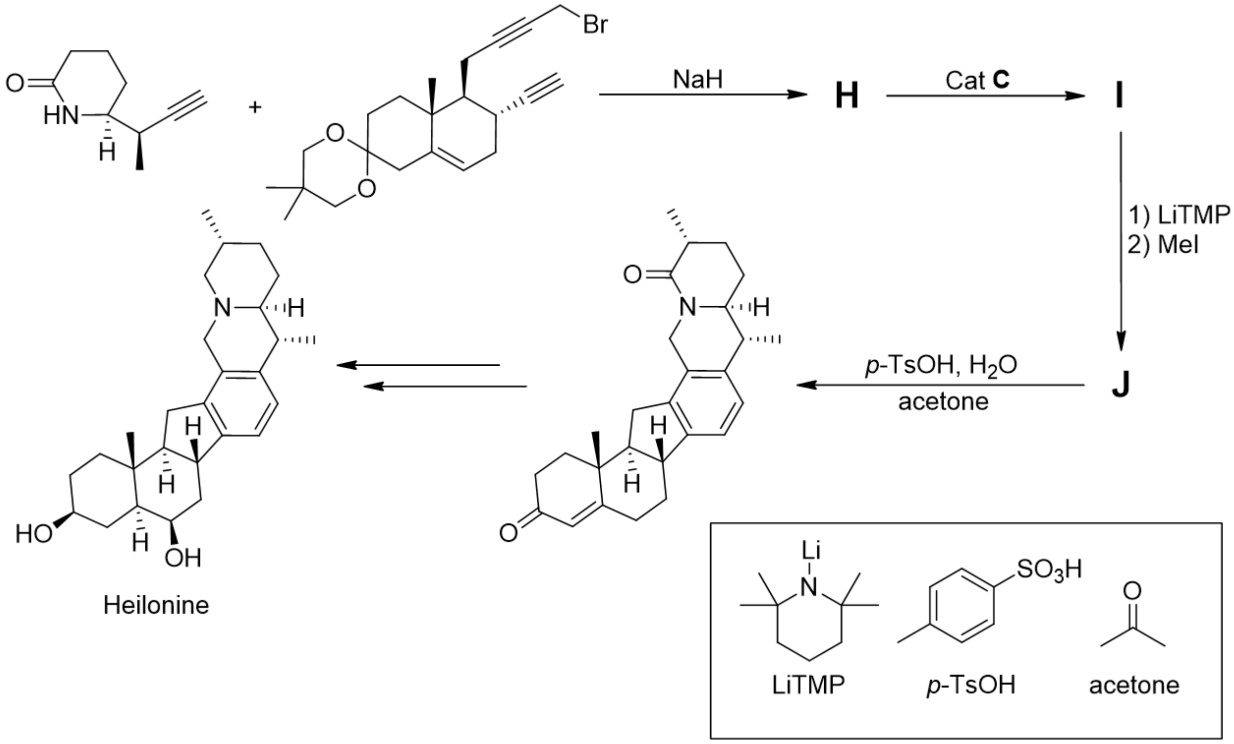
6. **Draw** the structure of **D** (you do not have to show stereochemistry).

After several catalytic cycles, the catalyst loses its activity due to the decomposition of **C**, forming products **F** and **M**. **F** is produced by the oxidation of **B** with atmospheric O_2.

7. **Identify** the formulae of **F** and **write** the equation for the oxidation reaction of **B**.

Catalyst **C** can catalyse other types of reactions. For example, **C** was used to catalyse a [2+2+2] cycloaddition reaction in the synthesis of Heilonine:





8. **<u>Draw</u>** the structures of **H**–**J** (you do not have to show stereochemistry).

Problem 22. Transition metals in organometallic catalysis

Organometallic chemistry studies various aspects of compounds containing a metal–carbon bond. Many transition metal (TM) complexes are of fundamental importance due to their role in industrial catalytic processes.

There are many parameters that should be considered when choosing a TM complex for a particular process, and four significant ones are:

- oxidation state (OS) of the metal in the complex,
- coordination number (CN),
- total number of valence electrons (VE), which is the sum of the number of the metal *d*-electrons and electrons provided by the ligands coordinated to the metal,
- overall net charge of the species that contains the transition metal.

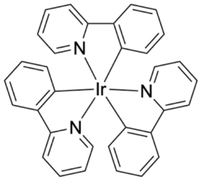
1. For complexes **1–4**, **calculate** OS, CN, and VE.

1
$Pd(PPh_3)_2Cl_2$

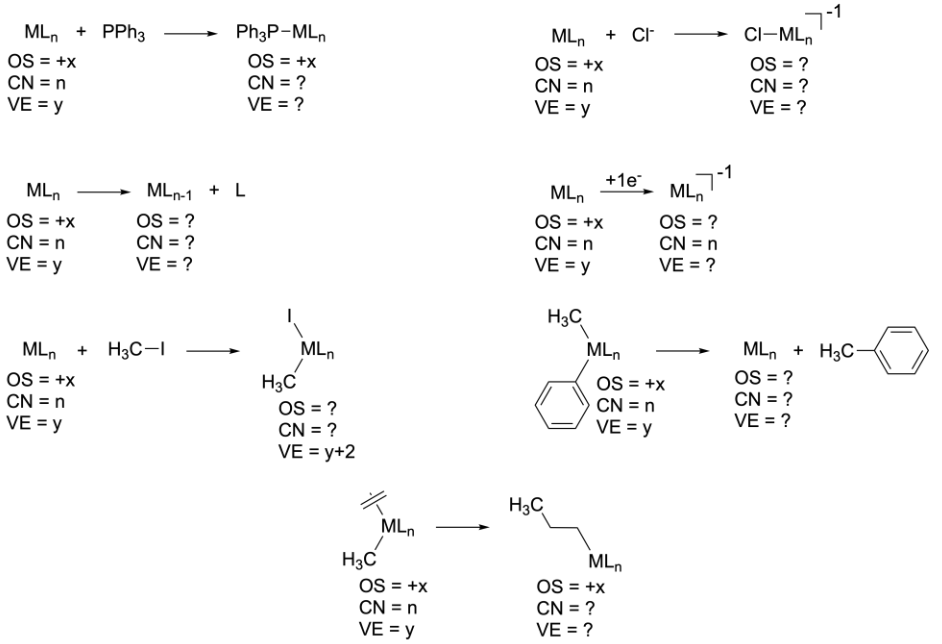
2
$Pt(PPh_3)_4$

3
$Rh(PPh_3)_3Cl$

4

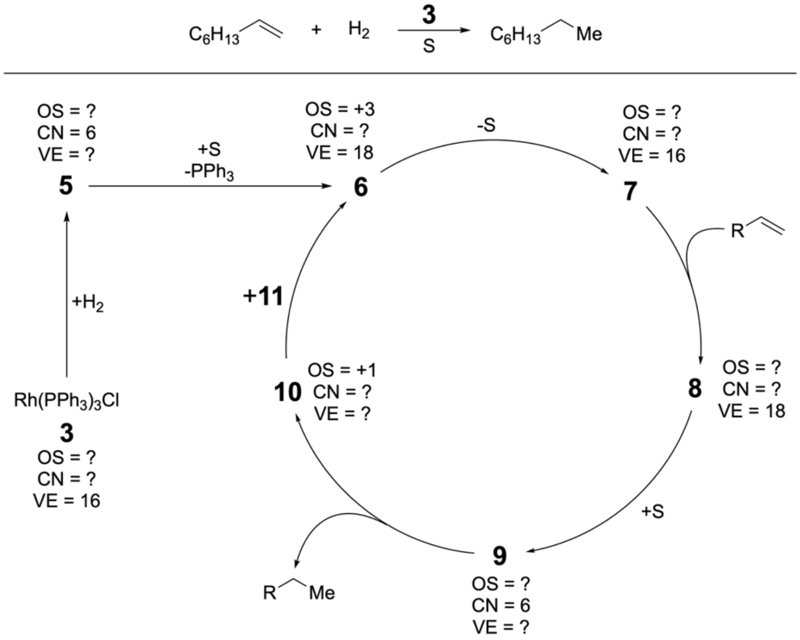


These parameters can change every time a TM complex undergoes different transformations. Below you can find the most important elementary steps that are essential for most TM-catalysed reactions.



2. **Find** the unknown parameters (OS, CN, VE) for complexes that underwent the mentioned elementary steps.

The parameters mentioned are quite important for catalytic transformations mediated by TM complexes. As such, **3** can catalyse the transformation of *n*-octene to *n*-octane via the catalytic cycle depicted below. S is a solvent that can coordinate to a metal with a pair of electrons.

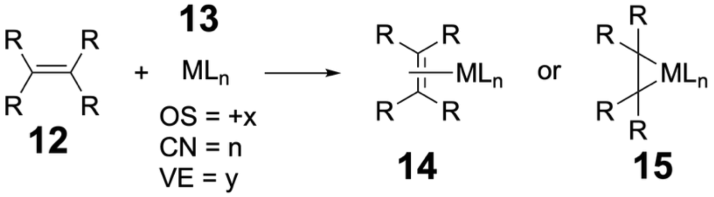


3. **Decipher** the structures **5–10** and **specify** the unknown parameters (OS, CN, VE) for them if the overall charge is zero for all the metal species.
4. **Determine** the structure of molecule **11**.

Oxidative addition of hydrogen can occur not only from complex **3**, but also from another complex. This complex is a neutral 14-electron species (which are generally more reactive), and it is obtained after one ligand leaves complex **3**. Thus, an alternative catalytic cycle can be proposed for the above transformation.

5. **Suggest** an alternative catalytic cycle that involves the aforementioned 14-electron complex. **Specify** the unknown parameters (OS, CN, VE) for the intermediates in your catalytic cycle (VE should be equal or greater than 14 for all the intermediates).

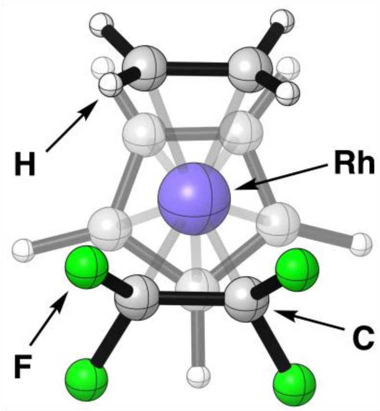
Some ligands can coordinate to metals in different ways. One example of such a case is the coordination of alkenes to a metal. Specifically, alkene **12** can coordinate to complex **13** without much disturbance of the C=C bond to produce **14** or via a metallacyclopropane mode to produce **15**.

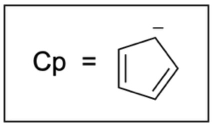


6. **Find** the unknown parameters (OS, CN, VE) for complexes **14** and **15** if the OS, CN, and VE are given for **13** in the scheme above.

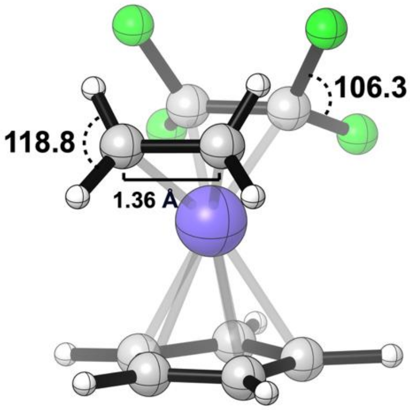
The geometry of complex **16** $RhCp(H_2CCH_2)(F_2CCF_2)$ is shown below. The colour codes of the atoms are as follows: Rh – violet, C – grey, F – green, H – white.

Top view

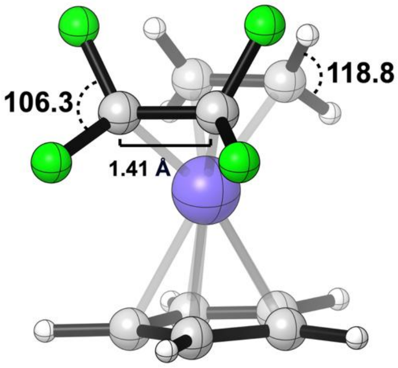




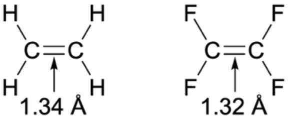
Side view 1



Side view 2



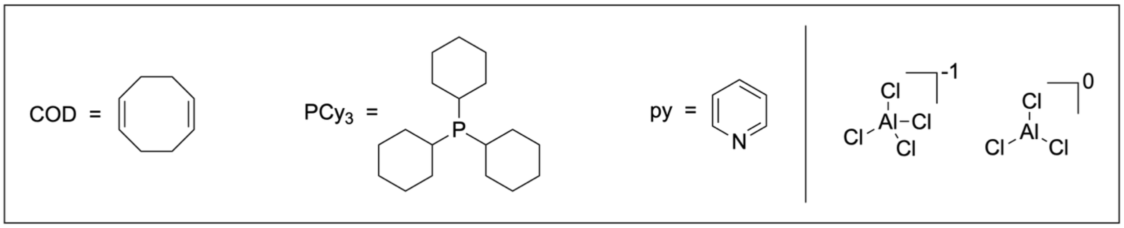
7. Based on this geometry, **find** the unknown parameters (OS and VE) for complex **16**.
Hint: the C–C bond lengths for free olefins are shown below.



8. **Specify** which mode of coordination is utilised by ethylene and tetrafluoroethylene in structure **16**. **Choose** from the following options:
   a) alkene coordination (similar to structure **14**);
   b) metallocyclopropane coordination (similar to structure **15**);
   c) no coordination.

Sometimes, the transition metal complex carries an overall charge (positive or negative).

9. <u>**Compare**</u> the following two complexes: $IrCl(PPh_3)_3$ (**17**) and $Ir(COD)(py)(PCy_3)(PF_6)$ (**18**). <u>**Calculate**</u> the unknown parameters (OS, CN, VE) for **18** if OS, CN, and VE are the same for **17** and **18**. <u>**Specify**</u> the overall charge of the metal part with the charge as presented in the example below for $AlCl_4^-$ and $AlCl_3$.



SET is a single electron transfer process and it it known that metal complexes can take part in it. An example of SET is shown below where donor D has a lone pair and during SET it gives one electron to acceptor A. This process produces a radical cation and a radical anion.

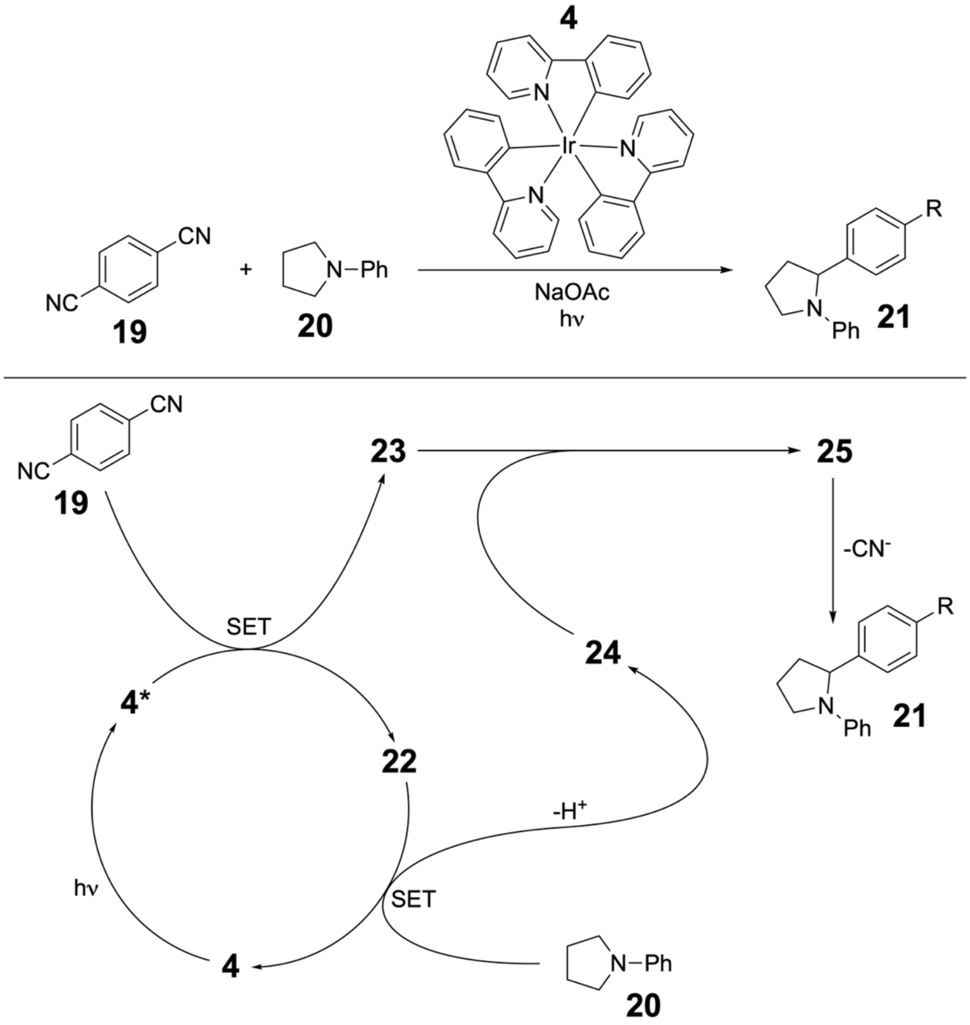
$$\ddot{D} + A \xrightarrow{\text{SET}} \overset{+\bullet}{D} + \overset{-\bullet}{A}$$

10. <u>**Find**</u> the unknown parameters (OS, CN, VE) for complexes below that underwent SET process. <u>**Specify**</u> overall charges *k* and *m* for the new complexes.

$$ML_n + A \xrightarrow{\text{SET}} [ML_n]^{k} + A^{\bullet -} \qquad ML_n + D \xrightarrow{\text{SET}} [ML_n]^{m} + D^{\bullet +}$$

ML_n	$[ML_n]^k$	ML_n	$[ML_n]^m$
OS = +x	OS = ?	OS = +x	OS = ?
CN = n	CN = n	CN = n	CN = n
VE = y	VE = ?	VE = y	VE = ?

Some of metal complexes are useful photocatalysts. For instance, previously discussed complex **4** can be applied for C–H-functionalisation of tertiary amines such as **20** with electron poor arenes such as **19**. The catalytic cycle for the synthesis of **21** is shown below and it consists of two SET processes.



11. **Indicate** which compound (**19** or **20**) is likely to be a donor in the SET process and which is likely to be an acceptor.

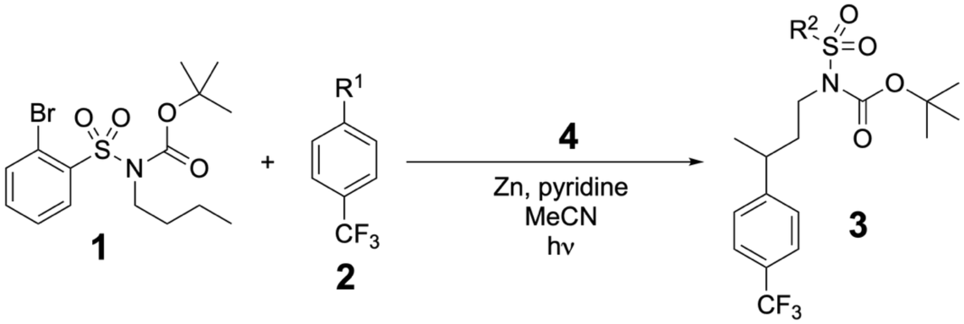
12. **Determine** the unknown structures **22**–**25**. Intermediate **4*** is an excited state of complex **4**. **Calculate** OS, CN, and VE for **22**, and show the overall charge of **22**.

13. **Identify** substituent R in product **21**.

Problem 23. HAT and XAT (HalAT)

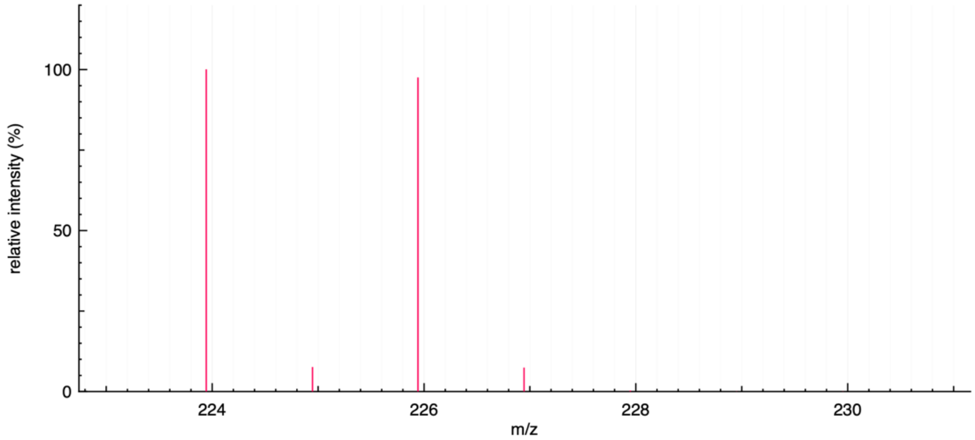
Note: Natural abundances of some isotopes are given at the end of the problem.

Transition metal-catalysed reactions can be versatile and unique. As such, nickel complex **4** can catalyse the transformation of compounds **1** and **2** into product **3** as described below.

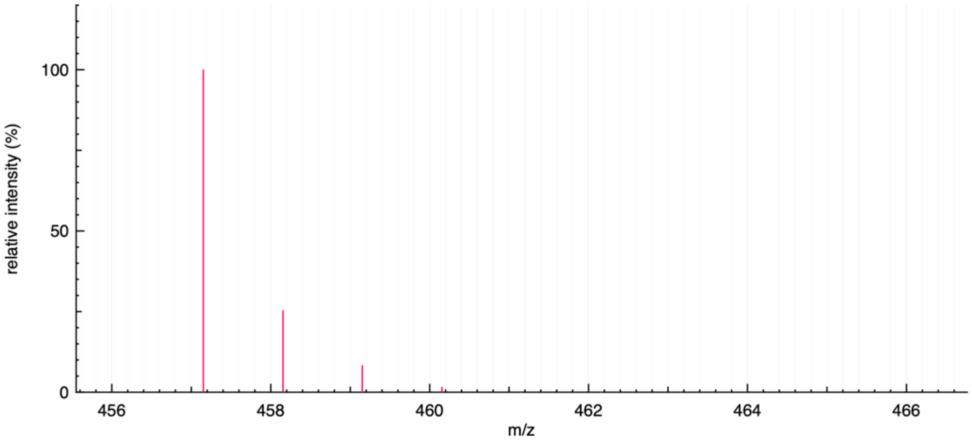


1. The mass spectra (MS) of **2** and **3** are shown (electron impact ionisation method). **Determine** the structures of **2** and **3**.

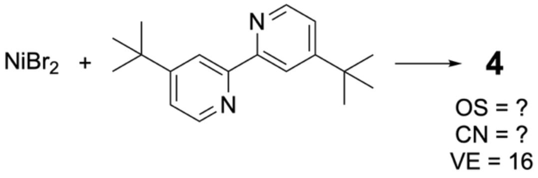
MS of **2**



MS of **3**

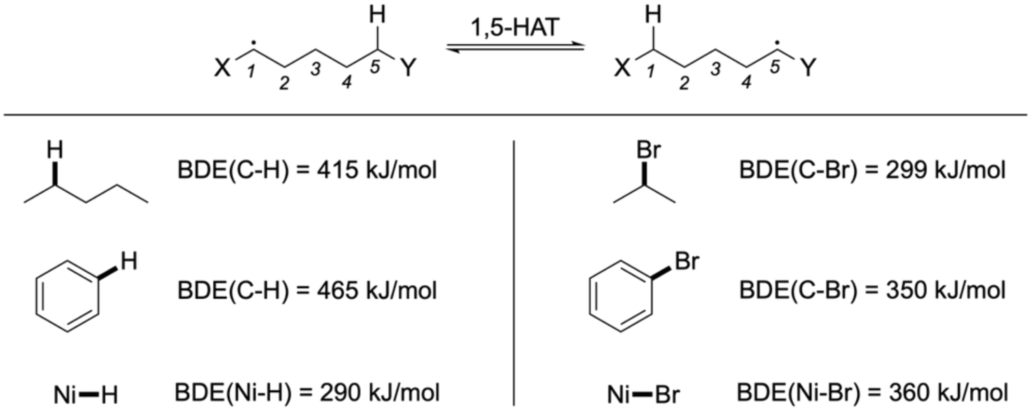


Complex **4** can be obtained via the following reaction. OS is the oxidation state of a metal in a complex; CN is the coordination number, and VE is the total number of valence electrons.

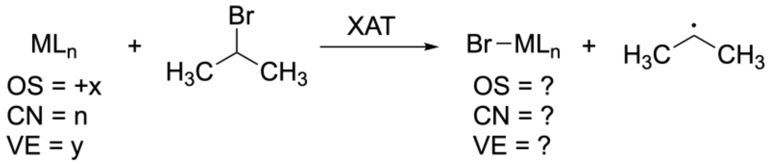


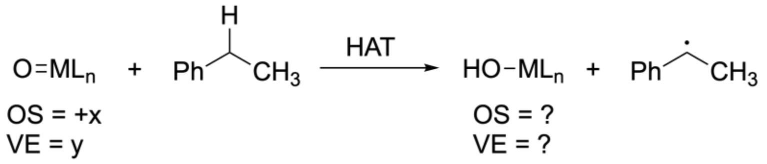
2. **Draw** the structure of **4** and **report** the OS and CN of the metal in it.

HAT is Hydrogen Atom Transfer, while XAT is Halogen Atom Transfer. If atom transfer is intramolecular, it is called 1,n-AT. An example of 1,5-HAT is presented on the scheme below along with some Bond Dissociation Energies (BDE) useful for this task.



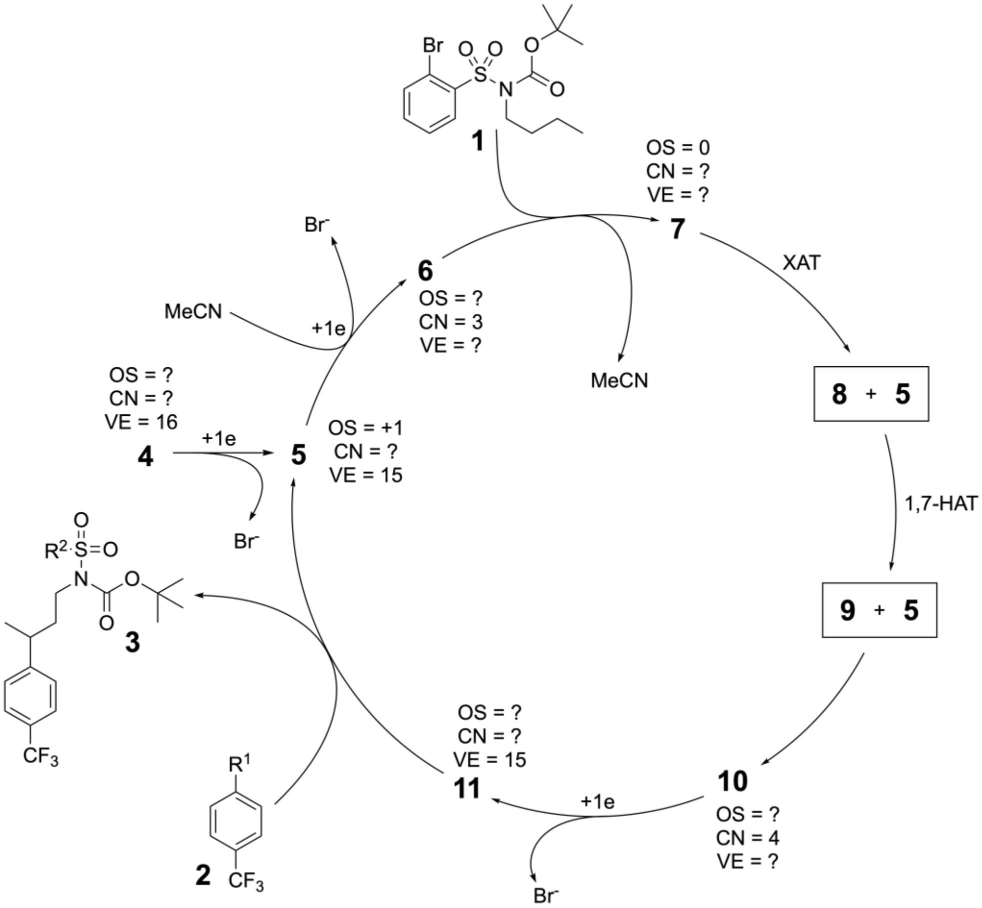
Metal complexes can participate in HAT and XAT processes. Two examples are given below:





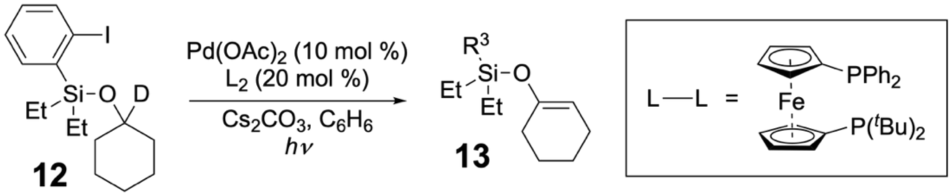
3. **Find** the missing parameters (OS, CN, and VE) for the complexes produced in HAT and XAT processes.

The catalytic cycle for the transformation of **1** and **2** to product **3** is described below:

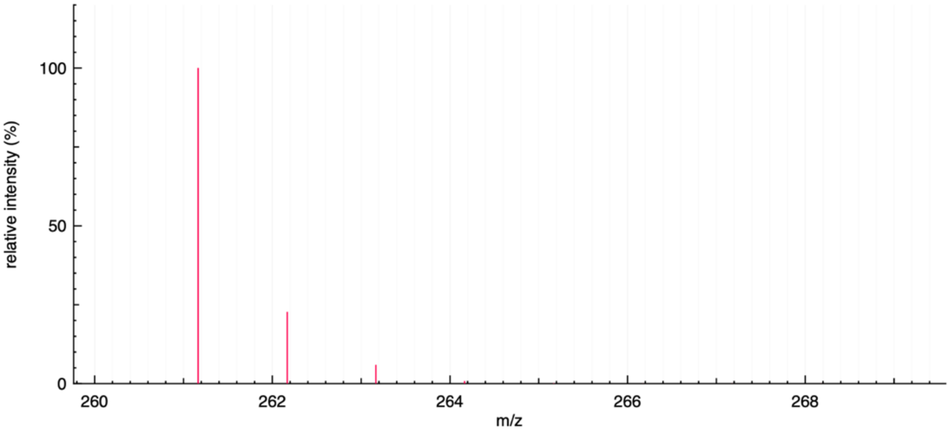


4. **Find** the structures of **5**–**11** in the cycle. For Ni-containing structures **5**–**7** and **10**–**11**, **find** the missing parameters (OS, CN, and VE).
5. **Analyse** the given reaction conditions for the synthesis of compound **3** and **determine** the electron donor in the above catalytic cycle.

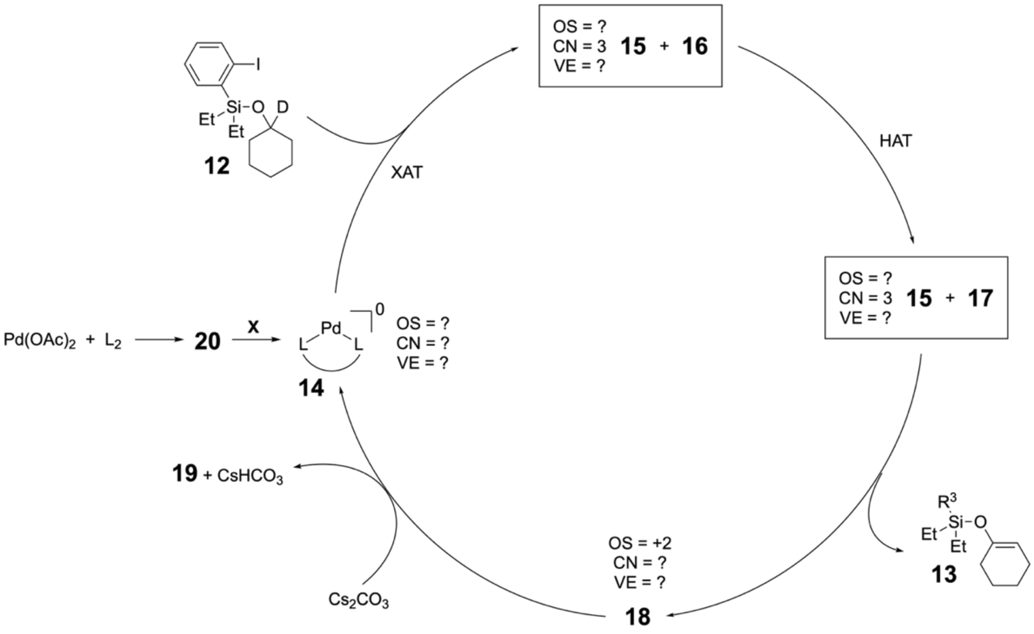
A conceptually similar catalytic reaction is a desaturation of silylated alcohols such as **12**. In the example below a deuterium label was installed for mechanistic studies.



6. The mass spectrum of **13** is presented below (electron impact ionisation method). **Determine** its structure.



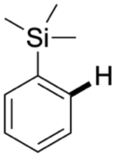
The catalytic cycle for the transformation of **12** into **13** is shown below:



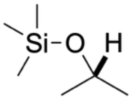
7. **Determine** the structures of **15**–**19**. For Pd-containing intermediates **14**, **15**, and **18**, **find** the missing parameters (OS, CN, VE).

Active species **14** is not particularly stable, thus, it is practically easier to start the reaction with a bench-stable palladium complex $Pd(OAc)_2$. The first step towards **14** is reaction with the phosphine ligand.

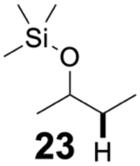
8. **Draw** a possible structure for **20**. **Suggest** what type of process should occur in the second step marked as **X** on the scheme. **Analyse** the given reaction conditions for the synthesis of **13** and **suggest** what reagent can participate in the step **X**.



21



22



23

9. Based on the information in the task, **analyse** BDE(C–H) of highlighted bonds in compounds **21–23**. **Compare** the values in pairwise manner using symbols <, > or =:

$\mathrm{BDE(C–H)}_{\mathbf{21}}$ ______ $\mathrm{BDE(C–H)}_{\mathbf{22}}$

$\mathrm{BDE(C–H)}_{\mathbf{21}}$ ______ $\mathrm{BDE(C–H)}_{\mathbf{23}}$

$\mathrm{BDE(C–H)}_{\mathbf{22}}$ ______ $\mathrm{BDE(C–H)}_{\mathbf{23}}$

Table of Natural Abundances

Isotope	Natural Abundance / %
^{1}H	99.9885
^{2}H	0.0115
^{12}C	98.93
^{13}C	1.07
^{14}N	99.632
^{15}N	0.368
^{16}O	99.757
^{17}O	0.038
^{19}F	100

Isotope	Natural Abundance / %
^{32}S	94.93
^{33}S	0.76
^{34}S	4.29
^{36}S	0.02
^{35}Cl	75.78
^{37}Cl	24.22
^{79}Br	50.69
^{81}Br	49.31
^{127}I	100

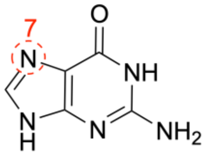
Problem 24. Anticancer complexes

The World Health Organisation reports that 20% of population will suffer from cancer in their lifetime. Today, cancer treatment is increasingly personalised being a combination of surgery, radiation, and chemotherapy with biomarker-guided targeted drugs and immunotherapies that can give long-lasting responses for some cancers. In this problem, we will consider some types of drugs used in cancer treatment.

Part 1. Platinum complexes

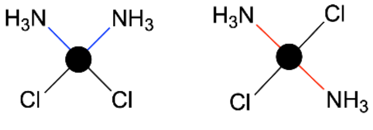
Platinum complexes show antitumor activity by covalently binding to DNA, thus preventing its replication. The first-generation platinum drug, *cis*-diaminedichloroplatinum(II) (cisplatin), is efficient against cancer, whereas its *trans*-isomer (transplatin) is not. One reason for this difference is that cisplatin and transplatin form different types of crosslinks with DNA strands.

There are two possibilities for a complex to bind DNA: intrastrand and interstrand crosslinks through the N7 position of guanine. Intrastrand crosslinks form between two bases of the same DNA strand, while interstrand crosslinks connect one base from one strand to a base on the opposite strand.

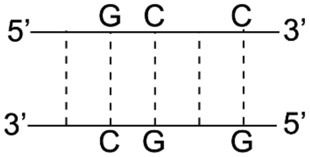


guanine

1. Using the dot notation shown below, **draw** the intrastrand and interstrand crosslinks on the given DNA template. **State** the kind of crosslink that is formed by a) cisplatin; b) transplatin.

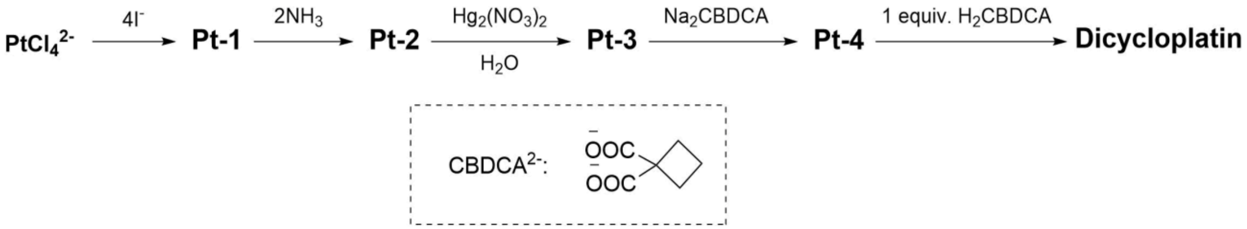


cisplatin transplatin



2. **Deduce** the effect of the formation of intrastrand and interstrand crosslinks on the melting temperature, T_m. T_m is the temperature at which half of DNA duplexes dissociate into single-stranded DNA.

Dicycloplatin, a third-generation platinum-based anticancer drug, was developed to reduce the dose-limiting toxicities that restrict the clinical use of cisplatin. Dicycloplatin crystals precipitate from a saturated aqueous solution containing **Pt-4** and 1,1-cyclobutanedicarboxylic acid (H_2CBDCA) in a 1:1 molar ratio. Dicycloplatin was found to have four hydrogen bonds in one structural unit.



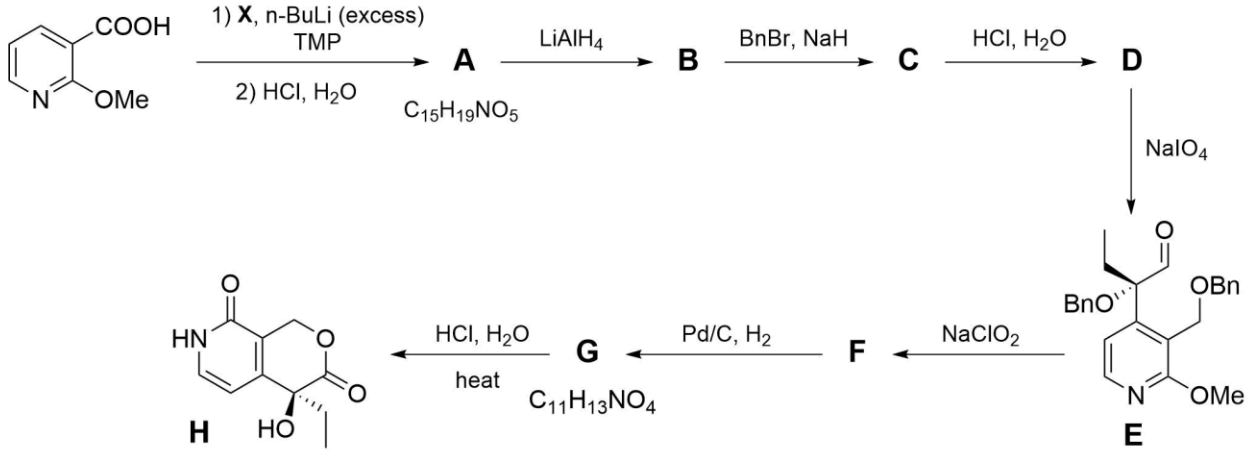
3. **Draw** the structural formulae of particles **Pt-1**, **Pt-2**, **Pt-3** and **Pt-4**. The mass percentage of platinum in **$PtCl_4^{2-}$** to **Pt-4** changes as 57.9% → 27.7% → 40.4% → 73.6% → 52.6%, respectively.
4. **Draw** the structural formula of dicycloplatin showing the hydrogen bonds.

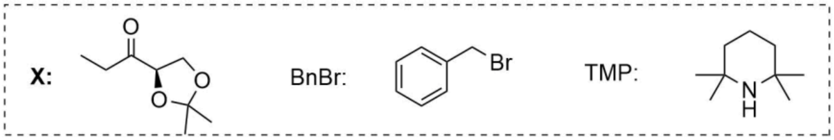
Pt-4 (carboplatin) was introduced into clinical use in the late 1980s and has been widely used in cancer treatment. One practical challenge with platinum-based drugs is that, in aqueous solution, they can undergo hydrolysis, which may affect their stability and reactivity.

5. **Explain** which drug, **Pt-4** or dicycloplatin, is more resistant to hydrolysis in aqueous solution.

Part 2. Camptothecin and its derivatives

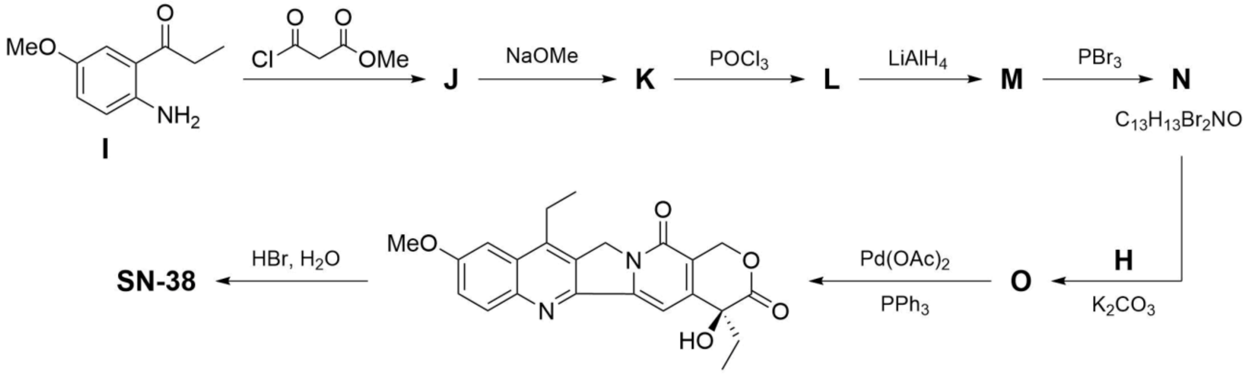
Alkaloid camptothecin and its derivatives form another group of drugs that show antitumor activity by inhibiting topoisomerase I, the enzyme that 'untangles' the DNA helix. Total synthesis of camptothecin derivative, **SN-38**, was carried out by coupling two building blocks **H** and **N**.





6. **Draw** the structural formulae of compounds **A**–**G** showing stereochemistry, if the step from **B** to **C** requires an excess of benzyl bromide.

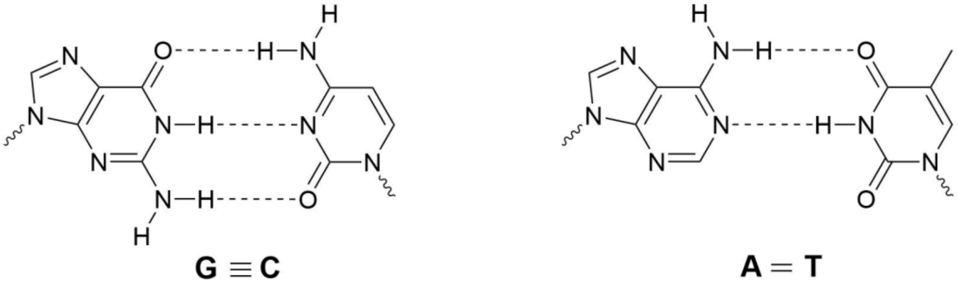
The second part of the synthesis is given below:



7. **Draw** the structural formulae of compounds **J**–**O** and **SN-38** showing stereochemistry where applicable.

Problem 25. The equilibria in DNA

A single-stranded DNA (ss-DNA) molecule can dimerise into a double-stranded one (ds-DNA) in the case that the ss-DNA is complementary to itself. Complementarity arises due to the ability of nitrogenous bases to form hydrogen bonds with each other. As shown below, guanine (G) forms three hydrogen bonds with cytosine (C), whereas adenine (A) forms two hydrogen bonds with thymine (T):



Tetranucleotide 5'-CCGG-3' can dimerise as follows:

$$\begin{array}{c} \text{5'-CCGG-3'} \\ + \\ \text{3'-GGCC-5'} \end{array} \rightleftharpoons \begin{array}{c} \text{5'-CCGG-3'} \\ |\,|\,|\,| \\ \text{3'-GGCC-5'} \end{array}$$

1. **Deduce** how many tetranucleotides exist that are capable of dimerisation with pairing of all 4 nucleotides in the strands.

Dimerisation of tetranucleotide 5'-CCGG-3' was studied at different temperatures by measuring the absorbance, A, of solutions at 280 nm, with total initial concentration of single strands $[\text{ss-DNA}]_0 = 10\ \mu\text{M}$. The results of the experiments were as follows: $A = 0.89$ at 10 °C and $A = 1.17$ at 30 °C.

2. **Calculate** $\Delta_r H°$ and $\Delta_r S°$ values for the dimerisation reaction of tetranucleotide 5'-CCGG-3', if they are independent of temperature and $\varepsilon_{\text{ss-DNA}} = 2.44 \times 10^4\ \text{cm}^{-1}\cdot\text{M}^{-1}$, $\varepsilon_{\text{ds-DNA}} = 2.93 \times 10^4\ \text{cm}^{-1}\cdot\text{M}^{-1}$ at 280 nm for tetranucleotide 5'-CCGG-3', $l = 5$ cm.

The melting temperature of ds-DNA, T_m, is the temperature at which half of ds-DNA dissociates into ss-DNA.

3. **Calculate** T_m for the double strand of tetranucleotide 5'-CCGG-3' when the total concentration of single strands, $[\text{ss-DNA}]_0 = 10\ \mu\text{M}$.

Another driving force for ds-DNA formation is the π-π-stacking between the nitrogenous bases in addition to hydrogen bonds. To estimate the contribution of the π-π-stacking to the stabilisation energy of ds-DNA, the dimerisation of oligonucleotides 5'-ACCGG-3', 5'-CCGGT-3', and 5'-ACCGGT-3' was investigated.

4. **Deduce** what dimers do oligonucleotides 5’-ACCGG-3’, 5’-CCGGT-3’, 5’-ACCGGT-3’ form. **Draw** them schematically.

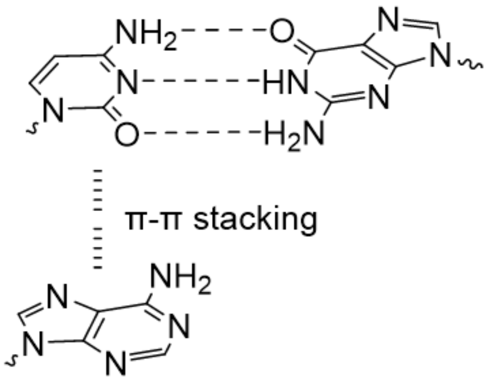
The T_m of these oligonucleotides has the following dependence on the total concentration of single strands:

$$\frac{1}{T_\mathrm{m}} = k \times \ln[ss - \mathrm{DNA}]_0 + b$$

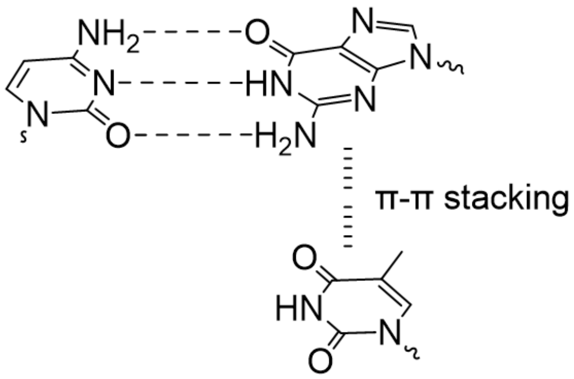
The following table shows k and b parameters for some oligonucleotides:

Oligonucleotide	***k* / K^{-1}**	***b* / K^{-1}**
5’-ACCGG-3’	-5.15×10^{-5}	2.76×10^{-3}
5’-CCGGT-3’	-4.80×10^{-5}	2.78×10^{-3}

5. **Calculate** the $\Delta_r H°$ values for the dimerisation reactions of oligonucleotides 5’-ACCGG-3’ and 5’-CCGGT-3’, assuming they are independent of temperature.
6. **Determine** the π-π-stacking energies for A-CG and T-GC stackings (see figure below). **Assume** that there is no π-π stacking in ss-DNA.



A-CG π-π stacking

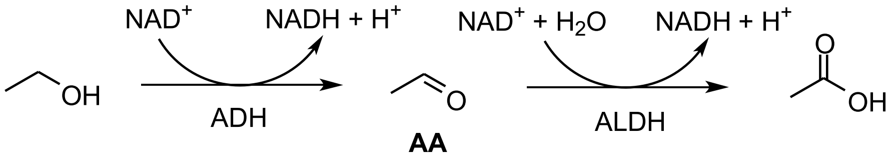


T-GC π-π stacking

7. **Predict** the $\Delta_r H°$ value for the dimerisation reaction of oligonucleotide 5’-ACCGGT-3’, if the total hydrogen bond energy in an A=T pair is 28842.0 J·mol^{-1}.

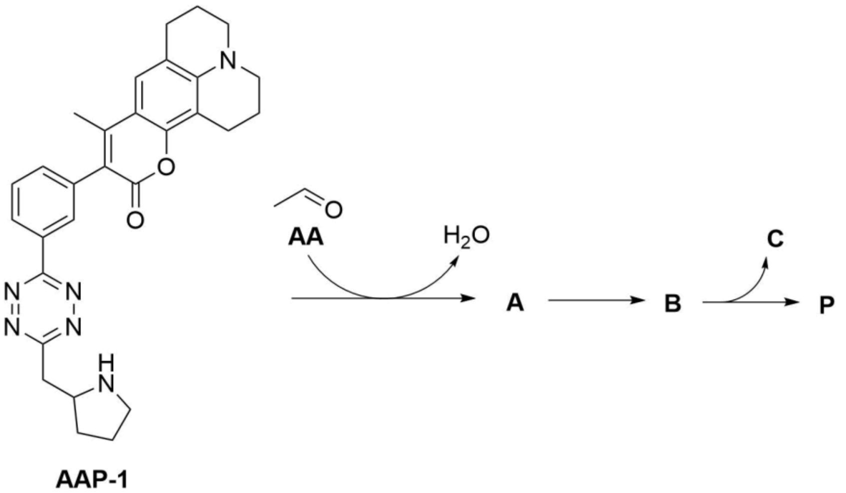
Problem 26. Sensing of acetaldehyde

Acetaldehyde (**AA**) is a reactive aldehyde primarily produced in cells as a metabolic intermediate during ethanol oxidation:

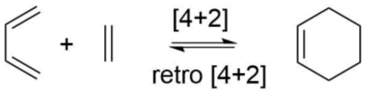


ADH – alcohol dehydrogenase, ALDH – aldehyde dehydrogenase.

Excess **AA**, often resulting from impaired **AA** detoxification, leads to aberrant DNA, protein, and/or lipid damage, and increases risk of diseases such as cancer, hepatitis, and cirrhosis. Recently, G. Li, C. Chang *et al.* reported the development of a highly selective method for the detection of **AA** in living cells. In this system, **AAP-1** reacts with **AA** to form a fluorescent product **P** by the following mechanism:



A is an enamine, **C** is an inorganic byproduct, and **P** contains the same number of rings as **AAP-1**. The above mechanism includes [4+2] and retro [4+2] cycloaddition reactions:



1. **<u>Draw</u>** the structures of intermediates **A**–**B**, product **P**, and determine the inorganic byproduct **C**. *Note: Molecular ions of **AAP-1** and **P** in mass spectra have peaks with masses of 494 and 492, respectively.*

Although biological fluids contain different types of aldehydes and ketones, **AAP-1** reacts only with **AA**.

2. **Give** the reason why formaldehyde (methanal) and methylglyoxal (2-oxopropanal) do not react with **AAP-1**.

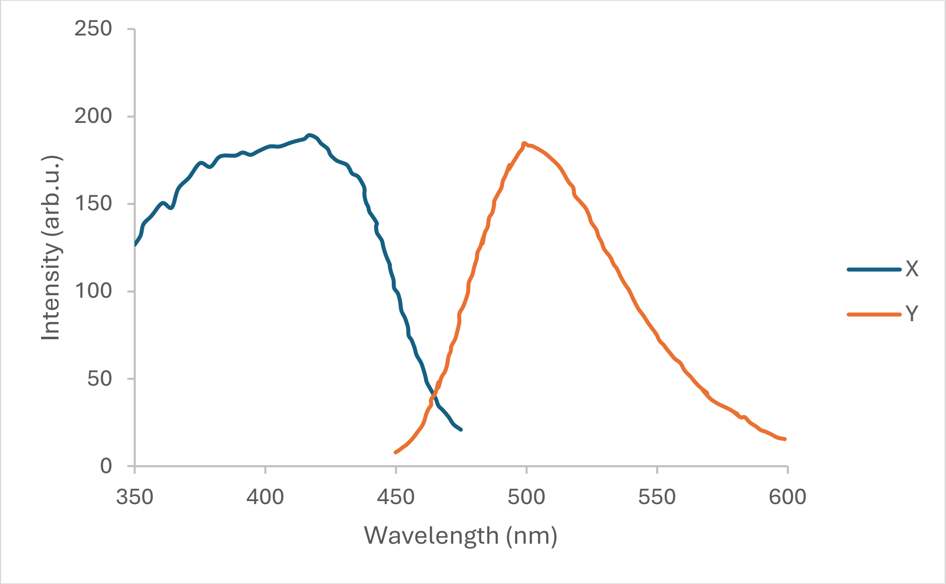
Aldehydes and ketones on relatively large molecules also do not react with **APP-1** due to the steric effects during transformation of enamine **A** to [4+2] cycloaddition product **B**. The following table shows the [4+2] cycloaddition reaction participants' Gibbs energies relative to reagents for **AA** (**A**, transition state #1 and **B**) and hexanal (**A'**, transition state #2 and **B'**):

Species	**ΔG / kJ·mol^{-1}**	**Species**	**ΔG / kJ·mol^{-1}**
enamine **A**	−8.36	enamine **A'**	−12.96
transition state #1	63.95	transition state #2	68.13
cycloaddition product **B**	−21.74	cycloaddition product **B'**	−19.23

3. **Calculate** the Gibbs energy changes (kJ·mol^{-1}) for processes **A→B** and **A'→B'**. **Conclude** whether these reactions are spontaneous.

4. **Calculate** the ratio of the rate constants corresponding to the **A→B** and **A'→B'** transformations at 310 K. **Assume** that in both reactions pre-exponential constant is the same, and that activation energy is equal to activation Gibbs energy.

The excitation and emission spectra of product **P** are shown below:



5. **Match** the spectral lines **X** and **Y** with excitation and emission processes.

The fluorescence intensity, F, of **P** depends linearly on its concentration: $F = F_0 + k[\mathbf{P}]$. The following table shows the fluorescence intensities (in arbitrary units) of solutions with different concentrations of **P**:

F / arb.u.	92.50	97.50	102.50	107.50
[**P**] / μM	10	20	30	40

6. **Calculate** the values of F_0 and k.

The lower detection limit of the fluorescence intensity of **P** is 90.65 arb.u. In the absence of ethanol, the concentration of **AA** can very quickly decrease due to enzyme ALDH ($r_{max} = 4.1$ mM·min^{-1}, $K_M = 0.2$ μM).

7. **Calculate** the time needed (in seconds) for the fluorescence intensity of the product of **AA** with **AAP-1** to decrease from 137.50 arb.u. to 90.65 arb.u. in the absence of ethanol. *Note*: *ALDH obeys Michaelis–Menten kinetics.*

Adding inhibitors to cell cultures where the concentration of **AA** is determined using **AAP-1** affects the intensity of fluorescence:

Control (no inhibitor, no EtOH)	EtOH	EtOH + Fomepizole (ADH inhibitor)	EtOH + CVT-10216 (ALDH inhibitor)
F_1	F_2	F_3	F_4

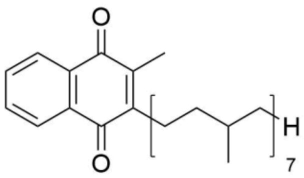
8. **Arrange** F_1–F_4 in order of increasing fluorescence intensity.

Problem 27. Archaeal metabolism

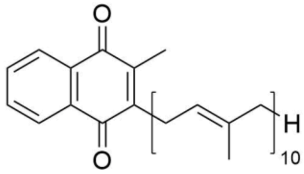
Archaea, along with bacteria and eukaryotes, constitute a distinct domain of living organisms. Archaea were previously grouped with bacteria, but this classification is now considered obsolete, as it has been established that archaea have their own independent evolutionary history and are characterised by many biochemical features that distinguish them from other life forms. Archaea appeared on Earth 3.5–4 billion years ago during the Archean Eon and have evolved continuously since then.

1. **Select** from the options provided a measurable, planetary-scale quantitative parameter that has not changed within reasonable accuracy throughout the entire evolution of archaea and to changes to which they can quickly adapt:
   a) *Chemical parameters*: the concentration of a gaseous substance in the atmosphere (oxygen, nitrogen or carbon dioxide), the composition of the environment (reservoirs) as the ratio of various compounds (salinity, pH, buffer capacity), etc.
   b) *Physical parameters*: temperature, atmospheric pressure, radioactivity, gravitational force, total luminosity of the Sun.

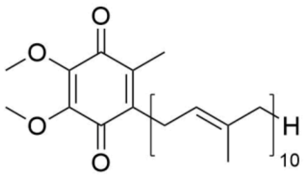
To demonstrate the differences between the three main domains of living organisms, one can cite as an example the quinone derivatives that play a role in oxidation-reduction processes in archaea (menaquinone MK-7), bacteria (menaquinone MK-10), and eukaryotes (ubiquinone, presented in the form of coenzyme Q_{10}, CoQ_{10}):



MK-7



MK-10



coenzyme Q_{10}

2. **Draw** the structure of the two-electron reduction product of MK-7.

The standard biochemical redox potentials, *E*°', of some redox pairs are given in the table below.

Redox pair	***E*°' / V**
$0.5\ O_2/H_2O$	0.82
cytochrome *c* (+3)/ cytochrome *c* (+2)	0.22
$CoQ_{10}/CoQ_{10}H_2$	0.10
fumarate/succinate	0.03
MK-10/MKH_2-10	–0.07
acetaldehyde/ethanol	–0.20

3. Based on the information provided, qualitatively **explain** why menaquinone was displaced by ubiquinone in the ancestors of modern eukaryotes and is currently not found in eukaryotes, with a few exceptions.

Natural ubiquinol (reduced form of ubiquinone), obtained from eukaryotic cell culture, is a hydrophobic unsaturated compound that can be incorporated into the standard method for determining bromine value (strictly in the dark!).

4. **Suggest** at least two chemical structural features of natural samples of ubiquinol that distinguish it from, for example, the triacylglycerol of linoleic acid, often called linolein, and that should be considered when analysing the obtained results of interaction with molecular bromine.

The unique habitat of archaea, from their very inception, has led some species to possess an enzyme involved in the metabolism of extremely intriguing substrates – substances **X**, **Y**, and **Z**, which exhibit similar chemical properties. Thus, structural isomers **X** and **Y** react *in vitro* with menaquinol (MKH$_2$-7), catalysed by an enzyme commonly referred to as **Z**-reductase, according to the following reaction equations:

$$\mathbf{X} + MKH_2 \rightarrow \mathbf{X1} + \mathbf{X2} + MK \qquad (1)$$

$$\mathbf{Y} + MKH_2 \rightarrow \mathbf{Y1} + \mathbf{Y2} + MK \qquad (2)$$

For the molar masses of unknown substances rounded to the nearest integer, three strict inequalities hold: $M(\mathbf{X}) < 250$ g/mol; $M(\mathbf{X1}) > M(\mathbf{X2})$; $M(\mathbf{Y1}) > M(\mathbf{Y2})$.

5. Using only the information above, **find** the upper limit of the range of possible integer molar masses of **X2**.

X1 and **X2** each contain 22% of element **E1** and 1.0% of element **E2** by mass, while **X** contains 23.4% oxygen by mass.

6. **Calculate** the percentage by mass of **E2** in **X** to the nearest whole number.

7. Without disclosing any elements other than **E2**, **narrow** the range of possible whole-number molecular masses of **X2** as much as possible, taking into account the mass fractions given above.

8. **Suggest** two potential explanations for the apparent logical inconsistency between the equal mass fractions of elements in two compounds despite different whole-number molecular masses; **discriminate** against one of them for the purposes of the problem.

Despite the lack of any direct information, the problem statement provides sufficient information to unambiguously determine the structures of **Y1** and **Y2**.

9. **Draw** the molecular formulae of **Y1** and **Y2**, and propose the structural formulae of substrates **X** and **Y**.

Compounds **X** and **Y**, unlike compound **Z**, are not found in nature.

10. Based on the principles of enzymatic catalysis, **propose** the molecular formula of **Z**.

Surprisingly, there is a substrate **W** that, when catalysed by the enzyme **Z**-reductase, yields two products with identical molecular masses when rounded to the nearest whole number.

11. **Suggest** two variants of the molecular formula of **W**, both containing **E1**.

12. **Describe** the metabolic characteristics and habitat of the archaea species whose metabolic fragment was discussed above.

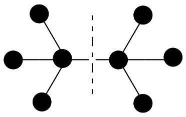
Most of the IChO 2026 problem authors are recent medalists of the International Chemistry Olympiads. Try testing your skills by creating an Olympiad problem based on the composition you've just solved. Your goal is to simplify this problem somewhat, guiding the participants' thinking toward the correct answers.

13. **Analyse** the reaction equations (1) and (2) and, based on the interesting coincidences/patterns you've identified, **offer** two hints to make the problem easier for the solvers without oversimplifying it.

Problem 28. Non-archea-typal form of life

Molecule **A**, with the formula $C_xH_yO_z$, where $x, y, z \in \mathbb{N}$, is found in some fossil deposits (such as oil), and serves as a biomarker for the presence of archaea in a particular region of the planet in the past.

Molecule **A** can be represented using a simplified version of knot notation: each covalent bond, regardless of the type of atoms forming it, is represented by a strand, with the atoms themselves being the knots at the ends of one or more strands. The ethane molecule, written in this notation, looks like this:



Using scissors, it takes 123 cuts to completely separate molecule **A** into individual nodes along the strands (the same operation with ethane requires 7 cuts). Note that two nodes are separated by the number of cuts equal to the number of strands between them (for example, a double bond requires two cuts). It is known that there is a single cut that bisects the molecule **A** into two identical fragments (similarly to ethane, as shown by the dashed line in the figure above).

1. Based on the data presented above, **decide** whether the indices x, y, and z are even or odd numbers or can be either of these. Is it possible to decide wither $x \leq 60$ or not? For $x = 60$, **suggest** two fundamentally different theoretically possible structures that meet the conditions (ignore the reality of their existence; this is just a theoretical concept).

Hint: Using the method of incomplete mathematical induction, try to consider the dependence of the number of bonds on the molecular formula for the first members of the homologous series of alkanes or saturated alcohols, as well as the acetylene-ethylene-ethane series.

2. **Deduce** whether there exists a theoretical acyclic form of **A** that can be divided into two identical fragments: a) by two cuts; b) by three cuts. If you answered yes, **propose** a molecular formula supporting your choice.

It is known that **A** has a molecular mass, rounded to the nearest whole number, of 595 g/mol.

3. Without guessing, uniquely **find** the molecular formula of **A**.

Molecule **A**, containing 8 stereocentres, can be cut along the strands into two types of fragments (fragments of the same type are defined as those that can be superimposed). Note that the cuts are made at equal distances from each other (the shortest number of strands between them, excluding branches), and one of the cuts coincides in location with the one that bisects **A** into two identical fragments as described above. Based on the experiment, **A** can be represented as a sequence of two fragments, *c* and *d*, as *dccccccccd*.

4. **State** whether the hydroxy groups in **A** can be: a) secondary; b) tertiary? If so, **draw** the corresponding structural formula(e) representing the cuts by dashed lines. Showing stereochemistry is not required.
5. **Draw** all possible structural formulae of **A** that form the *dccccccccd* pattern with cuts spaced apart by at least three intact strands.
6. Knowing the prevalence of certain structural biochemical motifs, **choose** the variant(s) of **A** (from those found in question 5) that best fit(s) to the metabolites of living things.

A is formed by the hydrolysis of compound **B** ($C_kH_lO_m$), which is involved in the formation of membranes in most archaeal species. The number of covalent bonds (n) in **B** is in the range $261 \leq n \leq 265$. The hydrolysis process is abiotic in nature and is a classic example of diagenesis (the long-term decomposition of organic material in geological deposits). The reaction equation for the decomposition of **B** can be written as follows:

$$\mathbf{B} + 4H_2O \rightarrow 2\mathbf{A} + 2\mathbf{C}$$

Note:

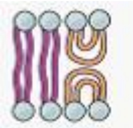
- *Optically inactive **C** has a symmetry axis passing through one of the carbon atoms;*
- *The molecular formula of both **C** residues in **B** is the same.*

7. **Prove**, without determining the formulae, that $l \leq 2k$.
8. **Find** all structural formulae of **C** that satisfy the above condition.

By combining pairwise the structures of **A** obtained in question 6 with the structures of **C** from question 8, we can see that one variant of **C** yields more structural isomers of **B** than all the other variants of **C** combined. This variant of **C** is found in archaeal lipids.

9. **Find** the final structure of **C** and determine the number of structural isomers of **B**. Draw the structural formulae of any two isomers of **B**.
10. **Explain** why the hydrolysis of **B** takes long and is not catalysed by enzymes of other microorganisms.

It is known that esterified derivatives of **B** are located in the membranes of some species of archaea in two fundamentally different ways (*O*- and *U*-conformations), shown below:



11. Of the **B** isomers proposed earlier, **select** those that can form similar membrane types. **Give** the advantage that such membranes have over those formed by bacterial and eukaryotic cells.
12. **Suggest** two explanations why the evolution chose **C** for the synthesis of **B**, rather than compounds from the same group with lower molecular weights.
13. Based on the information obtained in questions 11 and 12, **narrow down** as much as possible the range of isomers of **B** that meet the biological requirements of archaea.

Archaea of the phylum *Thaumarchaeota* adapt to rising environmental temperatures through a process virtually unique to them. It is known as homeoviscous adaptation, which involves modifying the structure of **B** by forming saturated cycles of either only five or only six carbon atoms (but not spirocycles) and, as a consequence, the biosynthesis of a separate family of lipids (**D**).

It is known that this modification of the structure of **B** is catalysed by two enzymes that are extremely similar in function and structure and exhibit strict positional specificity. Theoretically, based on the number of suitable loci, there could be up to four such enzymes, which would lead to an expansion of the **D** family (however, the third and fourth enzymes are not present in nature).

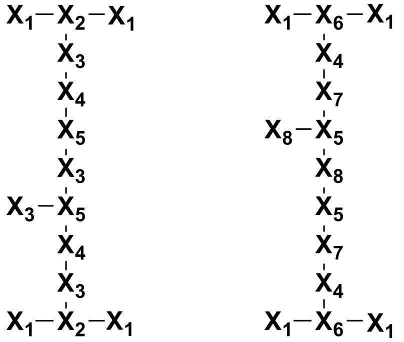
14. **<u>Determine</u>** the molecular formula of lipid **D1**, found in archaea, which withstands extreme temperature conditions better than other members of the family.

It is known that symmetrical molecule **D1** contains 26 stereocentres.

15. **<u>How many</u>** structural isomers that meet the condition can be proposed for **D1**? **<u>Draw</u>** any one of them. In this case, keep in mind that it is necessary to carefully analyse which carbon atoms can participate in the closure of the hydrocarbon ring, whether this option leads to an increase in the number of chiral centres and, if so, by what value.

Problem 29. Archaea-ology

Structural differences between biopolymers form one of fundamental criteria by which life forms are classified into separate domains. Biopolymers **A** (found in some archaea) and **B** (found in bacteria) have molecular weights over 100 kDa. The biopolymer fragments important for formation of spatial structures are shown below as I-shaped sequences of monomers ($\mathbf{X_1}$–$\mathbf{X_8}$) linked by covalent bonds.



fragment from **A** fragment from **B**

Each of the above fragments can be separated from the respective polymers (**A** or **B**) by cleaving four bonds.

1. The molecular masses of all $\mathbf{X}_i$ were rounded to the nearest integer. **Determine** whether the difference between the molecular masses of the fragment from **A** and the fragment from **B** is an even or odd number or can be any of these.

The ratio of all monomers forming the biopolymer **B** is close to equimolar.

2. **Propose** the regular two-dimensional structures of biopolymers **A** and **B**. Note that **A** and **B** contain no monomers other than those shown in the figure.

Hint: It may be useful to redraw the I-shaped fragments with $\mathbf{X_1}$–$\mathbf{X_8}$ *depicted as dots of different colours.*

A set of 6 molecular formulae is needed to be able to describe any $\mathbf{X}_i$ of $\mathbf{X_1}$–$\mathbf{X_8}$. *Note: the X_1–X_8 monomers might be both D- and L-compounds.*

3. Based only on the information given above, **choose** the theoretically correct statements from the given hereunder:
   a) $\mathbf{X_1}$–$\mathbf{X_8}$ can belong to different classes of biomolecules;
   b) Some of $\mathbf{X_1}$–$\mathbf{X_8}$ are isomers;
   c) Some of $\mathbf{X_1}$–$\mathbf{X_8}$ contain phosphorus;
   d) Some of $\mathbf{X_1}$–$\mathbf{X_8}$ are interclass isomers;
   e) 50% of $\mathbf{X_1}$–$\mathbf{X_8}$ are homologs.

In fact, only the first two statements are correct for $\mathbf{X_1}$–$\mathbf{X_8}$.

4. **State** whether it is now possible to determine unambiguously the nature of biopolymers **A** and **B**.

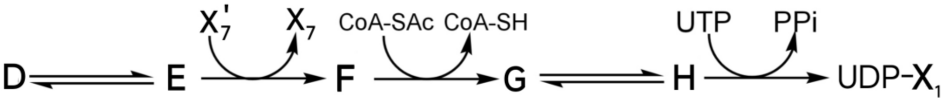
Any of the fragments (from **A** or **B**) can be divided into three pairs of smaller fragments containing at least two monomers linked together. The molecular mass of any of these resulting fragments (after rounding) belongs to the set: 217, 218, 275, 642 and 700. In **A** and **B**, monomers with identical molecular masses do not form covalent bonds with each other, and $M(\mathbf{X_6}) > M(\mathbf{X_2})$.

5. **Determine** unambiguously the molecular formulae of as many monomers as possible. If you fail doing so, **determine** the molecular masses of the monomers without assigning them to specific $\mathbf{X}_i$.

The sums of the molecular masses of all fragments formed from **A** and **B**, rounded to the nearest integer, are the same, 2270 g/mol.

6. **Explain** why these sums are equal. **Choose** the correct statement(s):
   a) occasionally;
   b) because the fragments from **A** and **B** are isomers;
   c) due to evolution reasons (only polymers with such molecular masses are functional).
7. **Determine** the molecular formulae of the yet unknown monomers.

Monomer $\mathbf{X_1}$ is synthesised from the product of photosynthesis, substance **D**, via the following biochemical pathway:



where CoA is coenzyme A, UDP is uridine diphosphate, UTP is uridine triphosphate, and PPi is pyrophosphate.

It is known that:

- **F** reacts with the Tollens reagent;
- $\mathbf{X_1}$ molecule is formed by atoms of four elements (C, H, N, and O);
- $\mathbf{X_1}$ contains 8 carbon atoms;
- Only two carbon atoms are involved in transformations;
- Three pairs can be arranged from compounds **D**, **E**, **F**, $\mathbf{X_7}$, and $\mathbf{X_7}$'. The molecular masses of these pairs differ by 1.

8. **Draw** the structures of **D**–**H** and $\mathbf{X_1}$. **Explain** how you decided which isomer is encoded by **D** and which by **E**.

To finally identify the structures of polymers **A** and **B**, previously unidentified fragments **A-1** and **B-1** were isolated. They were methylated on alcoholic hydroxy groups, followed by exhaustive hydrolysis. In the case of **B**, this resulted in the formation of the 3,4,6-tri-*O*-methyl derivative of $\mathbf{X_1}$, the 1,4,6-tri-*O*-methyl derivative of $\mathbf{X_1}$, and the 6-mono-*O*-methyl derivative of $\mathbf{X_6}$. For **A**, this resulted in the formation of the 1,4,6-tri-*O*-methyl derivative of $\mathbf{X_1}$, the 3,4,6-tri-*O*-methyl derivative of $\mathbf{X_1}$, and the 4-mono-*O*-methyl derivative of $\mathbf{X_2}$.

9. **Draw** the structures of **A-1** and **B-1**, designating the unknown substituents as R_i.

Complete acidic hydrolysis of **A-1** and **B-1**, followed by oxidation with HIO_4, leads to the formation of acetaldehyde in the amount equimolar to that of the starting fragment for only one of **A-1** and **B-1**.

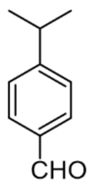
10. **Propose** the complete structural formulae of **A-1** and **B-1**.

Problem 30. Pilaf ingredients

Pilaf is one of the most delicious, traditional meals in Uzbekistan. We asked a chef who works in *Besh Qozon*, a famous restaurant, how they make pilaf so delicious. The chef described the procedure as follows:

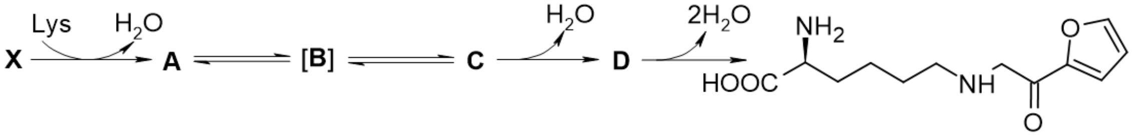
"Oil is first heated in a kazan. Chopped ***onions*** *are fried until they reach a uniform golden-brown color. Pieces of* ***meat*** *are then added and cooked until they develop a caramelised surface. Afterwards, sliced* ***carrots*** *are introduced, and the mixture is seasoned with* ***black pepper*** *and* ***cumin****. Water is added, whole* ***garlic*** *bulbs are placed on top, and the mixture is simmered until the meat becomes tender. Subsequently, pre-washed* ***rice*** *is added in a separate upper layer without stirring. Additional water is poured carefully to cover the rice. The mixture is cooked until all the water is absorbed. Finally, the kazan is covered, and the pilaf is allowed to steam for approximately 30 minutes. After gentle mixing, the pilaf is ready to be served".*

One of the key aromatic components of pilaf is ***cumin***. Its characteristic flavour is largely due to the presence of the compound cuminaldehyde. This aldehyde can be obtained synthetically starting from benzene, isopropyl chloride, dimethylformamide, and appropriate inorganic reagents.


Cuminaldehyde

1. **Suggest** the synthesis scheme for cuminaldehyde using the reagents mentioned above.

The monomer **X** of the major polymeric compound present in ***rice*** undergoes several transformations during cooking. These reactions lead to the formation of volatile and non-volatile products that contribute to the characteristic aroma and flavour of the dish. The synthesis of one of these products is shown below:



Hints:

- *Among* ***A–D*** *only compound* ***D*** *contains a ring.*
- ***D*** *does not have any acyclic tautomers.*
- ***A***, ***B*** *and* ***C*** *are tautomers.*

2. **Draw** the structures of compounds **A–D** and **X** without stereochemistry details.

The hydrolysis of the amide bond in the main component **Z** of ***black pepper*** produces compounds **E** and **F**. Upon oxidative ozonolysis, acyclic compound **F**, having one oxygen-containing group, yields an equimolar mixture of $C_{\alpha}H_{\beta}O_{\gamma}$ (**F1**) and $C_{\delta}H_{\varepsilon}O_{\alpha}$ (**F2**), where $\alpha \neq \beta \neq \gamma \neq \delta \neq \varepsilon$.

It is also known that:

- The indexes α, β, and γ are members of a geometric progression (not exactly in this order) and $\delta + \alpha = \varepsilon$;
- The total number (N) of all atoms in molecules **F1** and **F2** is an even number less than 45.

3. a) Using mathematical logic and chemical considerations, rigorously **prove** that the coefficients in compounds **F1** and **F2** (α, β, γ, δ and ε) are even numbers.

b) **Find** the molecular formulae of **F1** and **F2**.

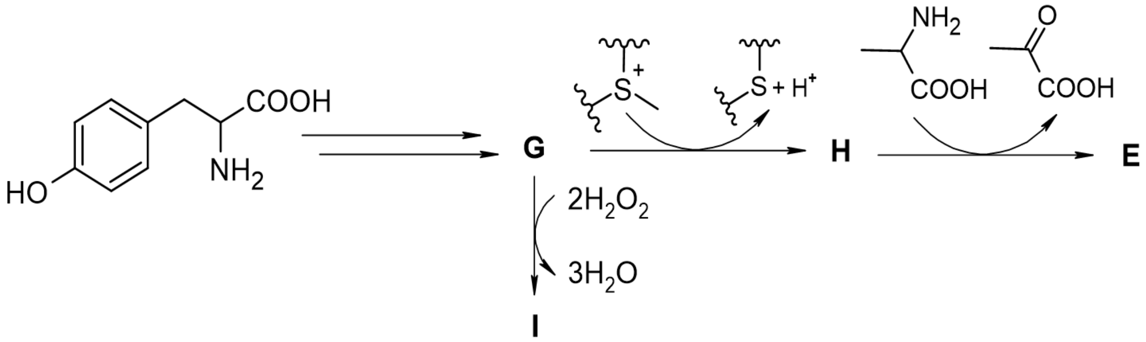
The **F1** and **F2** structures contain three types of hydrogen atoms and at least one plane of symmetry, while these compounds lack stereocentres and alcohol groups.

4. **Draw** the structures of compounds **F1**, **F2**, and **F**.

Combustion of 1.000 g of **E** at high temperature gives 3.0385 g of a gaseous mixture. Passing this mixture through solid sodium hydroxide increases the mass of the solid by 2.9470 g. **E** (w(O) = 20.92%) has a trisubstituted aromatic ring.

5. **Determine** the empirical formula of **E**.

The biosynthesis of **E** from L-tyrosine is shown below:



Hints:

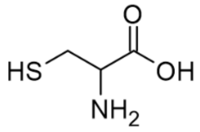
- ***G** contains an aldehyde group.*
- *The initial hydroxy group in L-tyrosine is unchanged during the biosynthesis of **E**.*
- ***I** is an ortho-quinone derivative.*

6. **Draw** the structures of compounds **G**, **H**, **I**, **E**, and **Z**.

Numerous peptides and amino acids are contained in ***meat***. A non-canonical α-amino acid **W** is one of these. The reaction of **W** with $NaIO_4$ yields a mixture of formaldehyde (CH_2O) and **K**. The oxidation of **K** with Fehling`s reagent, gives the anion of acid **L**. Acid **L** is achiral and has three different hydrogen atom environments.

7. **Draw** the structures of compounds **J**, **K**, **L**, and **W**.

Another notable non-canonical amino acid is alliin, **M**, found in ***garlic***, which serves as the precursor of the compounds responsible for its characteristic odour. Alliin, which contains a stereogenic sulfur atom, can be synthesised from cysteine in two steps. Firstly, cysteine reacts with allyl bromide under mildly basic conditions. The resulting intermediate, **N**, is then subjected to oxidation of the sulfur atom with H_2O_2, yielding **M**.



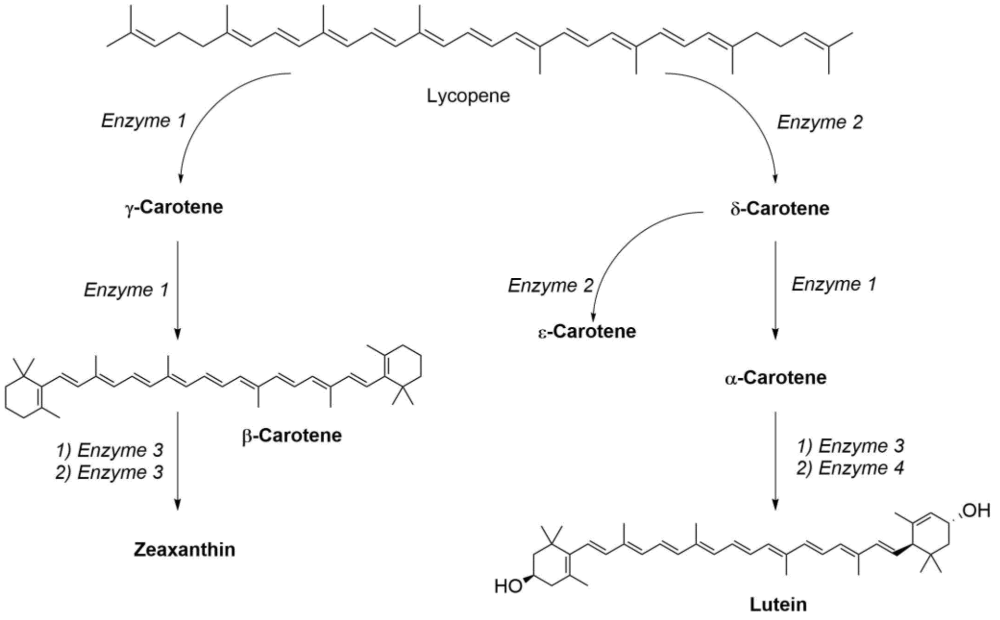
Cysteine

8. **Draw** the structures of compounds **M** and **N**.

Garlic also contains several other interesting molecules, e.g. four compounds **P_1**–**P_4**, which exhibit structural similarity. Their molar masses are members of an arithmetic progression, in which the first member is **P_1**. Molar masses of the second and third molecules are 146 g·mol^{-1} and 178 g·mol^{-1} respectively. Each of them contains at least one terminal C=C bond and possesses two planes of symmetry.

9. **Draw** the structures of compounds **P_1**–**P_4**, if it is known that **P_1** is biosynthesised from alliin.

There are a lot of dishes in the world using the combination of meat, rice, and carrots. However, Pilaf is the only one which utilises yellow ***carrot***. The reason behind the different colour of these carrots is the presence of two pigments – β-carotene and lutein. In plants, they are synthesised from the tetraterpene, lycopene.



10. **Draw** the structures of **α-**, **γ-**, **δ-**, **ε-carotenes** and **zeaxanthin**.

11. **Determine** whether **zeaxanthin** is optically active or not, **justify** your answer.

12. **Classify** Enzymes 1–4 according to the main enzyme classes below:

- *Oxidoreductases* – Oxidation/reduction reactions.
- *Transferases* – Transfer of a methyl-, acyl-, amino-, or phosphate group from one substance to another; kinases forming a subclass catalyse the transfer of phosphate group(s) from high-energy phosphorylated species to accepting substrate(s).
- *Hydrolases* – Hydrolytic formation of two products from a substrate.
- *Lyases* – Non-hydrolytic addition or removal of groups from substrates, with C–C, C–N, C–O, or C–S bonds cleavage in the latter case.
- *Isomerases* – Intramolecular rearrangement.
- *Ligases* – Joining together two molecules by formation of new C–O, C–S, C–N, or C–C bonds with simultaneous breakdown of ATP.

The chef in *Besh Qozon* did not mention the most important spice used in the preparation of pilaf.

13. **Identify** this unmentioned spice in Tashkent's pilaf.