Solutions to Preparatory problems for the 58th International Chemistry Olympiad 2026 Tashkent, Uzbekistan

Theoretical Problems

Contents

Problem 1. Khoja Nasreddin is now a chemist 3
Problem 2. Avicenna 6
Problem 3. Unusual inorganic synthesis 9
Problem 4. Nanozymes 13
Problem 5. Haber–Bosch process 15
Problem 6. Hydrogen fuel 18
Problem 7. 150 Years of Urease 24
Problem 8. Step-by-step 28
Problem 9. Host-guest chemistry of cucurbiturils as Uzbek hospitality 31
Problem 10. Prebiotic life 34
Problem 11. Fluorescence, phosphorescence and quenching 41
Problem 12. Photoacids 43
Problem 13. Interested in micelles? 49
Problem 14. The basics of Metal–Organic Frameworks 53
Problem 15. Microscopy in chemistry 55
Problem 16. Kinase inhibitors 58
Problem 17. Carbohydrate chemistry 60
Problem 18. Inositol 62
Problem 19. Carbohydrates as chiral pools 65
Problem 20. Pacman, Hangman, Cageman 68
Problem 21. Catalyst C 72
Problem 22. Transition metals in organometallic catalysis 75
Problem 23. HAT and XAT (HalAT) 80
Problem 24. Anticancer complexes 84
Problem 25. The equilibria in DNA 86
Problem 26. Sensing of acetaldehyde 88
Problem 27 Archaeal metabolism 91
Problem 28. Non-Archea-typal form of life 96
Problem 29. Archaea-ology 106
Problem 30. Pilaf ingredients 111

Problem 1. Khoja Nasreddin is now a chemist

1. F is dark blue liquid at low temperatures. One well-known substance with this property is ozone. This identification is consistent with the observation that when **F** is passed through a potassium iodide solution, the solution turns brown due to the formation of iodine, while an equal amount of a colourless gas (O_2) is released — a behaviour characteristic of ozone reactions. O_2 and N_2O_3 does not fit these conditions since they cannot oxidise iodide solution with producing another colourless gas with the same amount.

F – O_3.

D has a complex cation, which is obtained by treating **E** with ozone in liquid ammonia. The ligands could be ammonia since the reaction occurs in ammonia. This can be verified. We find the moles of ozone for the synthesis of 1 g of **D**: $\frac{16.3*10^{-3}*0.0998}{2} * 10 = 8.1337 * 10^{-3}$ mol

$M(\mathbf{D}) = \frac{1}{8.1337*10^{-3}} * x = 122.945x \frac{\text{g}}{\text{mol}}$; N(H) = $122.945x * \frac{0.0983}{1.008} = 12x$ which is divisible by 3 without a remainder, which indirectly means that the ligands in it are ammonia molecules.

When assumed that x = 1 (which means the ratio of **D** to ozone is 1), then $M(\mathbf{D})$ = 122.945 $\frac{\text{g}}{\text{mol}}$;

$N(NH_3) = 12 : 3 = 4$

$122.945 - 17 * 4 - 48 = 6.945$; this equals the atomic mass of lithium. This is consistent with the fact that the coordination number of lithium is always 4. The salt is an ozonide. **D**-$[Li(NH_3)_4]O_3$.

4.21% is a very small mass fraction. It may be hydrogen since its atomic mass is the smallest. Let's check this, $M(\mathbf{E}) = \frac{1.008x}{0.0421} = 23.94x \frac{\text{g}}{\text{mol}}$. It suits lithium hydroxide. **E** – LiOH.

In **C** there is the same metal as in **D**, that is, lithium. $M(\mathbf{C}) = \frac{6.94}{0.1263} = 54.95 \frac{\text{g}}{\text{mol}}$; which corresponds to **C** – LiO_3.

In **A**, the anion is also an ozonide. There are no solids left after all the reactions have completed, which suggests NH_4^+ as the cation.

The increase in pressure at the first stage indicates that gas is formed.

Let's find its quantity, $n_1 = \frac{(106.42-100)*1}{8.314*(-70+273.15)} = 3.8 * 10^{-3}$ mol

At a temperature of 128 °C the pressure should be,

$$P = 106.42 * \frac{128 + 273.15}{-70 + 273.15} = 210.14 \text{ kPa} < 260.63 \text{ kPa}$$

indicating that more gas is formed. Finding its quantity,

$$n_2 = \frac{(260.63 - 210.14) * 1}{8.314 * (128 + 273.15)} = 0.01514 \text{ mol}$$

An increase in moles of gases at 128 °C without any reaction indicates that evaporation has occurred. This temperature indicates that it was water evaporation.

$N = \frac{0.01514}{3.8*10^{-3}} = 4$; the ratio of water to the first gas is 4.

The combustion of a smoldering splinter indicates that there is oxygen in the gas mixture.

Oxygen is formed during the decomposition of ozonide, which is logical, since there are not many reactions where oxygen is formed during decomposition.

Then the first gas is oxygen, and during the decomposition of **A**, mixture of H_2O, O_2 and **B** was formed.

$$m(\mathbf{B}) = 1 - 3.8 * 10^{-3} * 32 - 0.01514 * 18 = 0.606 \text{ g}$$

One can find the amount of gases formed during the decomposition of **B**:

$$n_3 = \left\{396.82 - \left[260.63 * \frac{200 + 273.15}{128 + 273.15}\right]\right\} * \frac{1}{8.314 * (200 + 273.15)} = 0.02273 \text{ mol}$$

In reaction: 1**B**→ $y * gases$; M(**B**)=$\frac{0.606y}{0.02273} = 26.66y \frac{\text{g}}{\text{mol}}$; when $y = 3$, M(**B**)=80 $\frac{\text{g}}{\text{mol}}$

This molar mass is suitable for NH_4NO_3 , which decomposes to form 3 moles of gases.

B – NH_4NO_3

$$n(NH_4NO_3) = \frac{0.606}{80} = 7.575 * 10^{-3} \text{ mol}$$

NH_4NO_3	:	H_2O	:	O_2
$7.575 * 10^{-3}$		0.01514		$3.8 * 10^{-3}$
2	**:**	**4**	**:**	**1**

1g $3.8 * 10^{-3}$ mol

z**A** → $2NH_4NO_3$ + $4H_2O$ + $1O_2$

M(**A**)*z 1

$M(\mathbf{A})*z$ = 263.16 g, when z = 4 , $M(\mathbf{A})$ = 65.8 ≈ 66 $\frac{g}{mol}$, which corresponds to NH_4O_3.

A – NH_4O_3

Another way to find **A** is to sum atoms in the products, so for z**A** we have $N_4H_{16}O_{12} = (NH_4O_3)_4$, and it becomes easier to find **A**.

A	B	C	D	E	F
NH_4O_3	NH_4NO_3	LiO_3	$[Li(NH_3)_4]O_3$	LiOH	O_3

2.

[1] $4NH_4O_3 \rightarrow 2NH_4NO_3 + 4H_2O + O_2$

[2] $NH_4NO_3 \rightarrow N_2O + 2H_2O$

[3] $2KI + O_3 + H_2O \rightarrow 2KOH + I_2 + O_2$

[4] $2Na_2S_2O_3 + I_2 \rightarrow Na_2S_4O_6 + 2NaI$

Problem 2. Avicenna

1. Addition of $BaCl_2$ to **Solution 1** leads to the formation of the white precipitate $BaSO_4$, meaning that the mineral should contain a sulfur atom(s), and probably an oxygen atom(s). The formation of deep-blue **Solution 2** when ammonia solution is added to **Solution 1**, is the qualitative reaction for copper (Cu^{2+}) ions, while the brown precipitate indicates the formation of iron hydroxide under basic conditions ($Fe(OH)_3$). There are no more qualitative reactions for any other ions, so the possible elemental composition would be:

Fe-Cu-S; Fe-Cu-S-O.

Both combinations are considered correct.

2. Brown gas **A**, after dissolution of the mineral in concentrated nitric acid, is NO_2:

$$n(NO_2) = \frac{1.65}{22.4} = 0.0737 \text{ mol}$$

White precipitate **B** when Solution 1 reacts with $BaCl_2$, as mentioned above is $BaSO_4$.

$$n(BaSO_4) = \frac{0.929}{233} = 3.99 * 10^{-3} \text{ mol}$$

As **Solution 1** was divided into two parts:

$$n(BaSO_4) = 3.99 * 10^{-3} * 2 = 7.98 * 10^{-3} \text{ mol}$$

$$n(NO_2) = \frac{1.65}{22.4} = 0.0737 \text{ mol}$$

We can make calculations for both combinations of the *mineral*:

Combination Fe-Cu-S:

$$m(S) = 7.98 * 10^{-3} * 32 = 0.255 \text{ g}$$

$$m(\text{residual}) = 1.0 - 0.255 = 0.745 \text{ g}$$

We don't exactly know the oxidation states of all atoms in the mineral, so let's consider them as 0.

Then,

$$n(NO_2)_{\text{metals}} = 0.0737 - 7.98 * 10^{-3} * 6 = 2.58 * 10^{-2} \text{ mol}$$

Then, average equivalent mass of the metals:

$$\frac{0.745}{2.58 * 10^{-2}} = 28.9 \text{ g/mol}$$

As we know that one of the metals is copper, the equivalent mass of the second atom should be lower than 28.9 g/mol and it should form a brown precipitate when reacted with an excess of ammonia. The only option here is iron – Fe.

The brown precipitate formed after reaction of Solution 1 with ammonia is $Fe(OH)_3$ as mentioned above.

Then, $Fe(OH)_3$ gets oxidised by $KBrO_3$ under alkaline conditions to FeO_4^{2-} (**D**). The red precipitate that forms when **D** reacts with $BaCl_2$ is $BaFeO_4$.

$$n(BaFeO_4) = \frac{0.256}{256.8} = 1.0 * 10^{-3} \text{ mol}$$

Remembering that **Solution 1** was divided into two parts:

$$n(BaFeO_4) = 1.0 * 10^{-3} * 2 = 2.0 * 10^{-3} \text{ mol}$$

$$m(Fe) = 2.0 * 10^{-3} * 55.8 = 0.112 \text{ g}$$

$$\text{Then, } m(Cu) = 1 - 0.112 - 0.255 = 0.633 \text{ g}$$

$$n(Cu) = \frac{0.633}{63.5} = 0.01 \text{ mol}$$

Then, *n*(Cu) : *n*(Fe) : *n*(S) = 5 : 1 : 4; with the molecular formula of the mineral as Cu_5FeS_4. This can be further confirmed by NO_2 formation.

Combination Fe-Cu-S-O

$$m(Fe + Cu + O) = 0.745\ g$$

$$m(Fe) = 0.112\ g$$

$$m(Cu + O) = 0.745 - 0.112 = 0.633 \text{ g}$$

$$n(S) : n(Fe) = 7.98 * 10^{-3} : 2.0 * 10^{-3} = 4 : 1$$

Considering *mineral* has only one iron atom:

$$M_r(mineral) = \frac{1.00}{2 * 10^{-3}} = 500\ g/mol$$

Then,

$$m(Cu + O) = 500 - 56 - 32 * 4 = 316\ g/mol$$

The only possible combination of copper and oxygen atoms for this relation is:

$$n(Cu) = 3 \quad n(O) = 7.84 \approx 8$$

$$n(Cu) = 4 \quad n(O) = 3.88 \approx 4$$

However, these combinations can't be further confirmed by NO_2 formation, so these compositions of the mineral are incorrect.

Then,

$$\text{Mineral} - Cu_5FeS_4;\ \mathbf{A} - NO_2;\ \mathbf{B} - BaSO_4;\ \mathbf{C} - Fe(OH)_3;\ \mathbf{D} - K_2FeO_4;\ \mathbf{E} - BaFeO_4$$

3.

$$Cu_5FeS_4 + 50HNO_3 \rightarrow 5Cu(NO_3)_2 + Fe(NO_3)_3 + 4H_2SO_4 + 37NO_2 + 21H_2O$$

$$H_2SO_4 + BaCl_2 \rightarrow BaSO_4 \downarrow + 2HCl$$

$$Fe^{3+} + 3NH_3 + 3H_2O \rightarrow Fe(OH)_3 + 3NH_4^+$$

$$Cu^{2+} + 4NH_3 \rightarrow [Cu(NH_3)_4]^{2+}$$

$$2Fe(OH)_3 + KBrO_3 + 4KOH \rightarrow 2K_2FeO_4 + KBr + 5H_2O$$

$$K_2FeO_4 + BaCl_2 \rightarrow BaFeO_4 + 2KCl$$

Problem 3. Unusual inorganic synthesis

1. Cathode: $2H^+ + 2e^- \rightarrow H_2$.

$$Q = 10 \cdot 0.785 \cdot 0.1^2 \cdot 3600 = 282.6 \text{ C}$$

$$V_{H_2} = \frac{282.6 \cdot 22.4 \cdot 10^3}{96485 \cdot 2} = 32.8 \text{ mL}$$

2. $\mathbf{I_2} + 2S_2O_3^{2-} \rightarrow 2I^- + S_4O_6^{2-}$

3. From the text of the problem it can be understood that **A1-A4** contain the same unknown element in different oxidation states, the oxidation state of the element increases in the series: **A3** < **A4** < **A2** < **A1**.

Let's compare the oxidation potential of **S0** and **S1**:

$$\frac{0.100 \cdot 10.10 \cdot 10^{-3} \cdot 500}{10} = 0.0505 \text{ mol of electrons for } \mathbf{S0}$$

$$\frac{0.100 \cdot 1.35 \cdot 10^{-3} \cdot 100}{10} = 1.35 \cdot 10^{-3} \text{ mol of electrons for } \mathbf{S1}$$

During electrolysis, the following amount of electrons are removed from the anion **A2**:

$$\frac{282.6 \cdot 11.5 \cdot 10^{-2}}{96485} = 3.37 \cdot 10^{-4} \text{ mol of electrons}$$

If simplify the ratio $0.0505 : 1.35 \cdot 10^{-3} : 3.37 \cdot 10^{-4}$, we obtain the ratio 150 : 4 : 1. That means that the first titration determines the concentration of anion **A1** formed in solution **S1**, and the second titration determines the initial concentration of anion **A2** in solution **S0**. So, in the first titration anion **A2** was reduced and removed from solution **S2** in the form **A4**.

Element **A4** is a product of following conproportionation reactions:

Step **S1** → **S2**: **A2** + **A3** → **A4** (yellow),

Step **S3** → **S4**: **A1** + **A3** → **A4** (yellow).

Volatile, yellow element **A4** is Br_2 (this is further confirmed by calculating the atomic mass of the enriched isotope of **X**). So, anion **A3** has an oxidation state less than 0, this is Br^-.

The following diagram shows the redox properties of anion **A1** according to our calculations:

$$\mathbf{A2} \xrightarrow{-1x\ e^-} \mathbf{A1} \xrightarrow{+4x\ e^-} Br^-.$$

In case x = 1, **A1** is BrO_2^- and **A2** is BrO. This is incorrect because **A2** must be an anion.

In case $x = 2$, **A1** is BrO_4^- and **A2** is BrO_3^-. This is correct.

The spontaneous transformation of anion **B** into anion **A1** without the absorption or release of any chemical substances indicates that one of the elements in its composition is undergoing radioactive decay (the possible radioactive process is also indicated by the irradiation of the enriched **X** isotope with thermal neutrons). If we assume that the element emits positively charged particles and the anion **A1** has a charge of "−1" (BrO_4^-), anion **B** must have "0" or "+" charge. It's incorrect. So, the element in anion **B** emits negatively charged particles – electrons. Let's summarize our findings:

$$^{m}X \xrightarrow[3)O_3]{\begin{array}{c}1)\ \text{neutrons}\\2)\ HNO_3\end{array}} {}^{m+1}XO_4^{2-} \underset{\beta^-}{\rightarrow} {}^{m+1}BrO_4^-$$

So, **X** is Se.

$$\frac{0.9 \cdot 1.000}{m} = \frac{1.613}{m + 1 + 16 \cdot 4} \rightarrow m = 82 \text{ (atomic mass of enriched isotope } \mathbf{X}).$$

The atomic mass of the enriched isotope **X** is close to the relative atomic mass of bromine (79.916), which once again confirms our conclusions that substances **A1-A4** contain bromine.

Anion **B** is $^{83}SeO_4^{2-}$.

A1	A2	A3	A4	X	B
BrO_4^-	BrO_3^-	Br^-	Br_2	^{82}Se	$^{83}SeO_4^{2-}$

Reaction equations:

$^{82}Se + {}^{1}n \rightarrow {}^{83}Se$

$^{83}Se + 4HNO_3 \rightarrow H_2{}^{83}SeO_3 + 4NO_2 + H_2O$

$H_2{}^{83}SeO_3 + O_3 + 2OH^- \rightarrow {}^{83}SeO_4^{2-} + O_2 + 2H_2O$

$^{83}SeO_4^{2-} \rightarrow {}^{83}BrO_4^- + {}_{-1}^{0}\beta$

$BrO_3^- + H_2O - 2e^- \rightarrow BrO_4^- + 2H^+$ (at the anode)

$BrO_3^- + 5Br^- + 6H^+ \rightarrow 3Br_2 + 3H_2O$

$BrO_4^- + 7Br^- + 8H^+ \rightarrow 4Br_2 + 4H_2O$

$Br_2 + 2I^- \rightarrow 2Br^- + I_2$

$BrO_3^- + 6I^- + 6H^+ \rightarrow Br^- + 3I_2 + 3H_2O$

$I_2 + 2S_2O_3^{2-} \rightarrow 2I^- + S_4O_6^{2-}$

4. In the second titration the following reaction occurs:

$$BrO_3^- + 6I^- + 6H^+ \rightarrow Br^- + 3I_2 + 3H_2O$$

$$C_0 = \frac{0.0505}{6 \cdot 3 \cdot 10^{-3}} = 2.8\ \text{M}$$

5. In this case, too, only the β^- decay of the anion $^{83}BrO_4^-$ will lead to the formation of two neutral gaseous elements **C** and **D**. The stoichiometric ratio (1:2) allows us to conclude that **C** is ^{83}Kr and **D** is O_2.

6. The maximum concentration of anion **A1** will occur when the following condition is met:

$$\frac{d[\mathbf{A1}]}{dt} = k_1[\mathbf{B}] - k_2[\mathbf{A1}] = 0$$

Let's express the concentrations of **B** and **A1** through the initial concentration of **B** and time:

$$\frac{d[\mathbf{A1}]}{dt} = k_1[\mathbf{B}]_0 e^{-k_1 t} - \frac{k_2 k_1[\mathbf{B}]_0}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t}) = 0$$

$$k_1[\mathbf{B}]_0 e^{-k_1 t} = \frac{k_2 k_1[\mathbf{B}]_0}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t})$$

$$e^{-k_1 t} = \frac{k_2}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t})$$

$$\frac{k_2 - k_1}{k_2} = 1 - e^{(k_1 - k_2)t}$$

$$\frac{k_1}{k_2} = e^{(k_1 - k_2)t}$$

$$t = \frac{\ln\left(\frac{k_1}{k_2}\right)}{k_1 - k_2} = \frac{\ln\left(\frac{0.031}{0.0048}\right)}{0.031 - 0.0048} = 71.2\ \text{min}$$

An alternative solution method is using the derivative:

$$\left(\frac{k_1[\mathbf{B}]_0}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t})\right)' = 0$$

$$-\frac{k_1^2[\mathbf{B}]_0}{k_2-k_1}e^{-k_1t}+\frac{k_1k_2[\mathbf{B}]_0}{k_2-k_1}e^{-k_2t}=0$$

$$k_2e^{-k_2t}=k_1e^{-k_1t}$$

$$e^{(k_1-k_2)t}=\frac{k_1}{k_2}$$

$$t=\frac{\ln\frac{k_1}{k_2}}{k_1-k_2}=\frac{\ln\left(\frac{0.031}{0.0048}\right)}{0.031-0.0048}=71.2\text{ min}$$

7. For Xe to form because of β^- decay, salt **F** must contain an isotope of iodine:

$$\mathrm{K}\,{}^{p}\mathrm{IE}_n \xrightarrow[\beta^-]{} \mathrm{K}^+ + {}^{p}\mathrm{XeE}_n$$

$$0.1262=\frac{39.1}{39.1+p+nA_\mathrm{E}}\rightarrow A_\mathrm{E}=\frac{270.7-p}{n}$$

If we temporarily assume that p = 126.9, we obtain the following:

n	1	2	3	4	5	6	7	8
A_E	143.8	71.9	47.9	**35.9 ~ Cl**	28.8	23.9	20.5 ~ F	17.9 ~ O or F

Fluorine is excluded as an element in the problem, and oxygen does not match in valence, which means that element E is chlorine and n = 4.

$p=270.7-35.45\cdot4=129.$

So, **E** is $^{129}XeCl_4$ and **F** is $K^{129}ICl_4$.

Problem 4. Nanozymes

1. From the co-precipitation method, **X** has multiple oxidation states and is a transition metal. We can start with the fact that the cation and the anion from **S1** have a common halogen. We can also see that the charge of metal **X** is 3+ in **S1**. By using *m/z* numbers of the cation and the anion we form the following equations:

$$\begin{cases} x + \text{Hal} \times 2 + 18 \times 4 = 198 \\ \text{x} + \text{Hal} \times 4 = 196 \end{cases}$$

From here **Hal = 35** and ***x* = 56,** which are most abundant isotopes for **Cl** and **Fe**. So, **X** is **Fe**, **S1** is the hydrated crystal of **$FeCl_3$**.

For **S2**, the mass losses to turn from green to white indicates the loss of all water from it. If **S2** contains only one atom of **X**, its molar mass and number of hydrated water molecules are as follows:

$$M_{\mathbf{S2}} = \frac{55.85}{0.2009} = 278 \frac{\text{g}}{\text{mol}} \rightarrow N_{\text{H}_2\text{O}} = \frac{278 \times 0.4536}{18.016} = 7.$$

The residual molar mass for the anion of **S2** is $278 - 55.85 - 7 \times 18.016 = 96.038$, which corresponds to the SO_4^{2-} anion. So, **S2** is **$FeSO_4 \cdot 7H_2O$**.

Finally, we need to determine the amount of water molecules of hydration in **S1**. We can assume that the co-precipitation reaction is the following two parallel reactions:

$Fe^{2+} + 2OH^- \rightarrow FeO + H_2O$ (1)

$2Fe^{3+} + 6OH^- \rightarrow Fe_2O_3 + 3H_2O$ (2)

Salts **S1** and **S2** precipitate through reactions (1) and (2) respectively. The mass of FeO formed through reaction (1) is as follows:

$$m_{\text{FeO}} = \frac{1.3901}{278} \times (55.85 + 16.00) = 0.3593 \text{ g}$$

Through the mass of Fe_2O_3, we can calculate the number of water molecules in **S1**:

$$N_{\text{H}_2\text{O}} \times 18.016 = \frac{2.7025}{2 \times \frac{1.1578 - 0.3593}{(2 \times 55.85 + 3 \times 16.00)}} - (55.85 + 3 \times 35.46) \rightarrow N_{\text{H}_2\text{O}} = 6$$

So, **S1** is **$FeCl_3 \cdot 6H_2O$**.

2. The molar ratio of Fe^{II} to Fe^{III} shows the stochiometric relationship for **Fe_3O_4**. The values for n and m are 3 and 4 respectively.
3. $Fe^{2+} + 2Fe^{3+} + 8OH^- \rightarrow Fe_3O_4 + 4H_2O$.
4. So, n = 3 and $\mathbf{X}^{(n-1)+}$ is Fe^{2+}, $\mathbf{X}^{n+}$ is Fe^{3+}. The absence of catalytic activity means that there is no Fe^{2+} ion. So, that oxide is **Fe_2O_3**, in which the mass fraction of Fe is smaller than in Fe_3O_4.

5. To prolong the catalytic activity, we must reduce Fe^{3+} to Fe^{2+}:

$NaIO_4$ is incorrect, because it is an oxidant;

Sodium citrate is not a correct option as it forms a chelate complex with Fe^{2+} hindering the activity;

As Na_3PO_4 can interfere with Fe^{3+} ions, forming the insoluble $FePO_4$ salt, which has a reported solubility product constant around 1×10^{-22}, the catalytic performance is consequently reduced due to slower reduction of Fe^{3+} to Fe^{2+};

$NaBH_4$ is the only **correct option** without negative effects on the system. It reduces back oxidised Fe_2O_3 NPs to Fe_3O_4 NPs.

6. Firstly from the volume we can get $r = \sqrt[3]{\frac{3 \times V_{max}}{4\pi N_A}}$ = 0.355 nm. Using that value, the surface for one mole of $\mathbf{Y}_p$ is $S = 4\pi r^2 N_A = 9.532 \times 10^5$ $m^2\ mol^{-1}$. We can calculate the molar mass of $\mathbf{Y}_p$:

$$M_{\mathbf{Y}_p} = \frac{9.532 \times 10^5\ m^2\ mol^{-1}}{1.324 \times 10^6\ m^2\ kg^{-1}} = 0.720\ kg\ mol^{-1}$$

The only logical solution is C_{60}. So, **Y** is C, "*p*" is 60.

7. We have, $\frac{0.001 \times 10^{-6}}{60} = \frac{3.73 \times 10^{-9}[S]}{1.03 \times 10^{-3} + [S]}$, that gives [S] = 4.62 μM.

8. First, we need to calculate: $r = \frac{3.73 \times 10^{-9} \cdot 0.025 \times 10^{-3}}{1.03 \times 10^{-3} + 0.025 \times 10^{-3}} = 8.84 \times 10^{-11}\ M \cdot s^{-1}$, and from here we calculate $t = \frac{\Delta[P]}{r} = \frac{0.025 \times 10^{-6}}{8.84 \times 10^{-11}} = 283\ s = 4.72\ min$

Problem 5. Haber–Bosch process

1. $N_2 + 3H_2 \rightarrow 2NH_3$

2. $\Delta_r H^o = -46.2 \times 2 = -92.4 \left(\frac{\text{kJ}}{\text{mol}}\right)$

$\Delta_r S^o = 192.7 \times 2 - 191.6 - 130.7 \times 3 = -198.3 \left(\frac{\text{J}}{\text{mol} \cdot \text{K}}\right)$

$K_p = 1 \rightarrow \Delta_r G^o = 0 \rightarrow \boldsymbol{T = \frac{92400}{198.3} = 466\ \text{K}}$

3. $\Delta_r C_p = 35.1 \times 2 - 29.1 - 28.8 \times 3 = -45.3 \left(\frac{\text{J}}{\text{mol} \cdot \text{K}}\right)$

$T = 300 + 273 = 573\ \text{K}$

$\Delta_r H^o_{573} = -92400 - 45.3(573 - 298) = -104857.5 \left(\frac{\text{J}}{\text{mol}}\right)$

$\Delta_r S^o_{573} = -198.3 - 45.3 \ln\left(\frac{573}{298}\right) = -227.9 \left(\frac{\text{J}}{\text{mol} \cdot \text{K}}\right)$

$\Delta_r G^o_{573} = -104857.5 - 573 \times (-227.9) = 25729.2 \left(\frac{\text{J}}{\text{mol}}\right)$

$\boldsymbol{K_p = e^{-\frac{25729.2}{573 \times 8.314}} = 4.5 \times 10^{-3}}$

4. $n(H_2) = 0.9 \times 3 =$ **2.7 mol**.

5. **a)**

	N_2	+	$3H_2$	→	$2NH_3$
Initially	1		3		0
Reacted	$-x$		$-3x$		$+2x$
Equilibrium	$1-x$		$3-3x$		$2x$

$$K_p = \frac{\left(\frac{P_{NH_3}}{P^\circ}\right)^2}{\left(\frac{P_{H_2}}{P^\circ}\right)^3 \times \frac{P_{N_2}}{P^\circ}} = \frac{\left(\frac{2x}{4-2x} \times \frac{P}{P^\circ}\right)^2}{\left(\frac{3-3x}{4-2x} \times \frac{P}{P^\circ}\right)^3 \times \frac{1-x}{4-2x} \times \frac{P}{P^\circ}};\ P^\circ = 1\ \text{bar};\ P = 150\ \text{bar}$$

$$K_p = 4.5 \times 10^{-3} = \frac{(2x)^2 \times (4-2x)^2}{(3-3x)^3 \times (1-x) \times 150^2}$$

$$x = 0.733; \quad \textbf{yield} = \boldsymbol{73.3\%}$$

b) The total pressure will be higher on the reactant side and vice versa. As the total pressure increases at a constant temperature, the equilibrium will shift toward the product according to Le Chatelier's principle, which in turn increases the equilibrium yield.

6. **Stage 2→3** is limiting because the triple bond of nitrogen is breaking, which requires the largest activation energy. In addition, **stage 5→6** also has a large activation energy, because two hydrogen molecules dissociate at a time, so, this answer is also accepted.

Note: In this problem, the simple kinetic model was considered to understand the mechanism. In real experiments, the rate determining step is the dissociative chemisorption of nitrogen, where absorption and dissociation are the intrinsically coupled parts of a single event.

7. The catalyst is produced by reducing Fe_3O_4. However, the reducing agent cannot reduce the entire volume of Fe_3O_4, since its concentration decreases as it penetrates the catalyst and does not reach its entire interior. Thus, the catalyst remains unreduced inside. According to this logic, there should be two layers, but an additional layer forms, an under-reduced form of Fe_3O_4. It is known for certain that the oxidation state of Fe in this layer must be lower than that of Fe_3O_4 and greater than 0. Only **FeO** is suitable for this option.

8. $n(\text{Fe}) = \frac{10}{55.85} = 0.1791 \text{ mol}$

$$n(\text{Fe})_{\text{in } Fe_3O_4} = \frac{4}{3} \times 3.141 \times (50 \times 10^{-7})^3 \times \frac{5 \times 3}{231.55} = 3.3913 \times 10^{-17} \text{ mol}$$

$$n(\text{Fe})_{\text{in FeO}} = \frac{4}{3} \times 3.141 \times (60 \times 10^{-7\times3} - 50 \times 10^{-7\times3}) \times \frac{5.7}{71.85} =$$
$$= 3.0234 \times 10^{-17} \text{ mol}$$

$$n(\text{Fe})_{\text{in Fe}} = \frac{4}{3} \times 3.141 \times (100 \times 10^{-7\times3} - 60 \times 10^{-7\times3}) \times \frac{7.9}{55.85} =$$
$$= 4.6444 \times 10^{-16} \text{ mol}$$

$$\sum n(\text{Fe})_{\text{in nanoparticle}} = 3.3913 \times 10^{-17} + 3.0234 \times 10^{-17} + 4.6444 \times 10^{-16}$$
$$= 5.2859 \times 10^{-16} \text{ mol}$$

$$N(\text{nanoparticle}) = \frac{0.1791}{5.2859 \times 10^{-16}} = 3.3883 \times 10^{14}$$

$$S(\text{nanoparticle}) = 4 \times 3.141 \times (100 \times 10^{-9})^2 = 1.2564 \times 10^{-13} \text{ m}^2$$

$$\mathbf{S_{total}} = 1.2564 \times 10^{-13} \times 3.3883 \times 10^{14} = \mathbf{42.57 \text{ m}^2}$$

9. $V_{\text{cube}} = \frac{10}{7.9} = 1.266\ \text{cm}^3$

$S_{\text{cube}} = 1.266^{\frac{2}{3}} \times 6 = 7.02\ \text{cm}^2 = 7.02 \times 10^{-4}\ \text{m}^2$

$\frac{42.57}{7.02 \times 10^{-4}} =$ **60640 times faster**

Problem 6. Hydrogen fuel

1.

a) $H_{2(g)} + 0.5O_{2(g)} \rightarrow H_2O_{(l)}$

b) $CH_{4(g)} + 2O_{2(g)} \rightarrow CO_{2(g)} + 2H_2O_{(l)}$

c) $C_8H_{18(l)} + 12.5O_{2(g)} \rightarrow 8CO_{2(g)} + 9H_2O_{(l)}$

$$\Delta_r H^\circ = \Sigma \Delta_f H^\circ_{\text{products}} - \Sigma \Delta_f H^\circ_{\text{reagents}}$$

$\Delta_c H^\circ(H_2) = \mathbf{-285.8\ kJ/mol}$

$\Delta_c H^\circ(CH_4) = (-393.5 + (-285.8) \times 2) - (-74.8) = \mathbf{-890.3\ kJ/mol}$

$\Delta_c H^\circ(C_8H_{18}) = (-393.5 \times 8 + (-285.8) \times 9) - (-260.2) = \mathbf{-5460\ kJ/mol}$

2. Energy density (H_2) = $285.8 \times \frac{1000}{2} \times \frac{1}{1000} = 142.9$ MJ/kg

Energy density (CH_4) = $890.3 \times \frac{1000}{16} \times \frac{1}{1000} = 55.6$ MJ/kg

Energy density (C_8H_{18}) = $5460 \times \frac{1000}{114} \times \frac{1}{1000} = 47.9$ MJ/kg

Hydrogen is the best fuel, considering energy density.

3. $E = 1 \times \frac{3600}{0.36} = 10{,}000$ kJ

$$n(O_2) = \frac{10000}{285.8} \times 0.5 = 17.5 \text{ mol}$$

$$V(\text{air}) = \frac{17.5}{0.2} \times \frac{RT}{P} = 87.5 \times \frac{8.314 \times 298}{100} = 2170 \text{ L}$$

4. Total distance = $3.75 \times 10^6 \times 22000 = 8.25 \times 10^{10}$ km

Methane used for : $8.25 \times 1010 \times 0.65 = 5.3625 \times 10^{10}$ km

Gasoline used for : $8.25 \times 1010 \times 0.35 = 2.8875 \times 10^{10}$ km

$$V(CH_4) = \frac{5.3625 \times 10^{10} \times 5}{100} = 2.68125 \times 10^9 \text{ m}^3$$

$$n(CO_2)_{\text{from methane}} = \frac{2.68125 \times 10^9 \times 10^5}{8.314 \times 298} = 1.08221 \times 10^{11} \text{ mol}$$

$$V(\text{isooctane}) = \frac{2.8875 \times 10^{10} \times 9}{100} = 2.5988 \times 10^9 \text{ dm}^3$$

$$n(CO_2)_{\text{from isooctane}} = \frac{2.5988 \times 10^9 \times 1000 \times 0.72 \times 8}{114} = 1.31305 \times 10^{11} \text{ mol}$$

$$n(CO_2)_{\text{total}} = 1.08221 \times 10^{11} + 1.31305 \times 10^{11} = 2.39526 \times 10^{11} \text{ mol}$$

$m(CO_2) = \frac{2.39526 \times 10^{11} \times 44}{10^{12}} = \mathbf{10.539\ million\ tons\ per\ year}$

This is approximately 7-8% of the total annual CO_2 emissions in Uzbekistan.

5. $285800 = 75.4 \times (373 - 298) + (44000 + (-38) \times (373 - 298)) + 37.4 \times (800 - 373) +$

$$+(29.1 \times \boldsymbol{n(N_2)} + 29.4 \times \left(\frac{\boldsymbol{n(N_2)}}{4} - 0.5\right)) \times (800 - 298)$$

$$\boldsymbol{n(N_2)} = 12.6 \text{ mol}$$

For simplicity, the reaction: $H_{2(g)} + 0.5O_{2(g)} \rightarrow H_2O_{(g)}$ (T = 298 K, P = 1 bar) can be used for the calculations. Considering the reaction with liquid water would introduce an additional term in the energy balance, but it does not affect the final result. $\Delta_r H° = -241800$ J/mol

$$241800 = \left(37.4 \times 1 + \boldsymbol{n(N_2)} \times 29.1 + \left(\frac{\boldsymbol{n(N_2)}}{4} - 0.5\right) \times 29.4\right) \times (800 - 298)$$

$$\boldsymbol{n(N_2)} = 12.6 \text{ mol}$$

6.

a) $241800 = 37.4 \times 1 \times (T_f - 298)$

$T_f = 6763.24$ K

b) $\Delta_r H°$ and $\Delta_r S°$ as functions of temperature:

- $\Delta_r H° = -241800 - 6.1(T - 298)$
- $\Delta_r S° = -44.6 - 6.1 \ln\left(\frac{T}{298}\right)$

K_p for the reaction:

$$\boldsymbol{K_p = \exp\left\{-\frac{[(-241800 - 6.1(T - 298)) + T\left(44.6 + 6.1 \ln\left(\frac{T}{298}\right)\right]}{RT}\right\}} \quad (1)$$

	$H_{2(g)}$ +	$0.5O_{2(g)}$ →	$H_2O_{(g)}$
Initially	-	-	1
Reacted	x	$0.5x$	$-x$
Equilibrium	x	0.5	$1-x$

$$\boldsymbol{\sum n = 1 + 0.5x}$$

Since the equilibrium constant K_p is "large", for simplification, it is assumed that the reaction initially proceeds to completion and then slightly reverses to reach equilibrium.

$$K_p = \frac{\frac{P_{H_2O}}{P°}}{\frac{P_{H_2}}{P°} \times \sqrt{\frac{P_{O_2}}{P°}}} = \frac{\frac{1-x}{1+0.5x} \times \frac{P}{P°}}{\frac{x}{1+0.5x} \times \frac{P}{P°} \times \sqrt{\frac{0.5x}{1+0.5x} \times \frac{P}{P°}}}; \; P = P° = 1 \text{ bar}$$

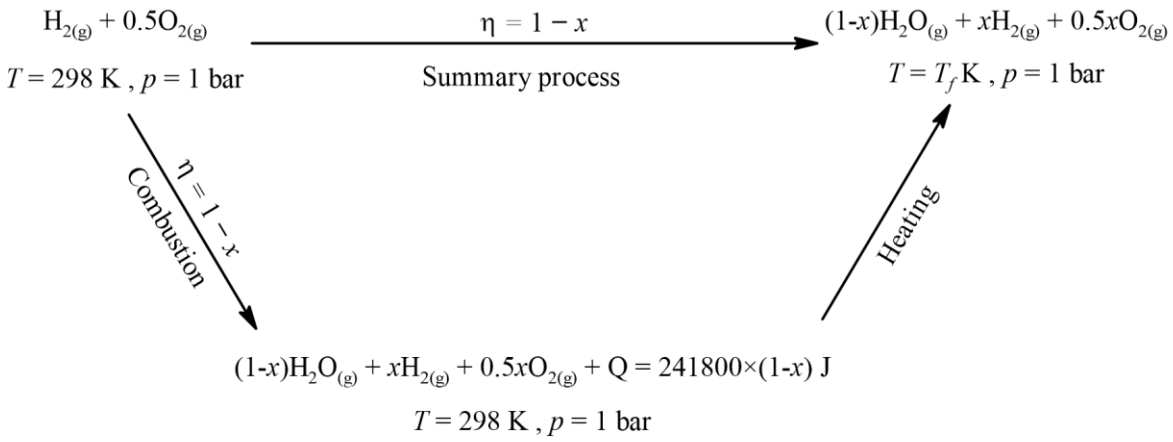
$$\boldsymbol{K_p = \frac{(1-x) \times \sqrt{1 + 0.5x}}{x \times \sqrt{0.5x}}} \quad (2)$$

Using equations **(1)** and **(2)** with T = 6763.24 K → $\boldsymbol{x}$ = 0.9614.

Composition of mixture: 0.9614 H_2 : 0.0386 H_2O : 0.4807 O_2, that corresponds to

$\chi(H_2)$ = 0.6493; $\chi(H_2O)$ = 0.0261; $\chi(O_2)$ = 0.3246.

c) One can construct a simple thermodynamic cycle for the combustion process:



$$241800 \times (1-x) = \left[(1-x) \times 37.4 + x \times 28.8 + x \times \frac{29.4}{2}\right](T_f - 298)$$

$$T_f = \frac{241800 \times (1-x)}{\left((1-x) \times 37.4 + x \times 28.8 + x \times \frac{29.4}{2}\right)} + 298 \qquad (3)$$

For x = 0.9614, T_f = 513.73 K. This value is not even near to 6763.24 K because the composition of the mixtures is very different.

d) The composition of the gas mixture at the final temperature is initially unknown, which is required to determine the final temperature, while simultaneously, the final temperature is unknown for calculating the mixture composition. Therefore, an iterative procedure is necessary, adjusting the temperature until it is consistent with all other parameters. The iteration can be performed as follows:

1. Determine the composition of the gas mixture at a given temperature
2. Calculate the final temperature using the energy balance based on the determined composition.
3. Compare the calculated temperature with the given temperature: if the given temperature is lower than the calculated value, increase it; if it is higher, decrease it.
4. Repeat steps 1–3 until the given and calculated temperatures converge.

The iterative procedure using equations (1), (2) and (3):

First iteration step:

Assume $T = 3900$ K. Then, $K_p = 2.42$, $x = 0.4836$, and the calculated temperature, $T_f = 3392.6$ K < 3900 K, indicating that the temperature should be decreased.

Second iteration step:

Assume $T = 3850$ K. Then, the calculated temperature $T_f = 3533.2$ K < 3850 K, indicating that the temperature should be decreased.

The following table summarises the calculated values:

T	K_p	x	T_f
3900	2.42	0.4836	3392.6
3850	2.6893	0.4619	3533.2
3800	2.9967	0.4400	3676.1
3750	3.3485	0.4180	3820.6
3770	3.202	0.4268	3762.7
3768	3.2163	0.4259	3768.6

T_f = 3768.6 K

[ ] 3900 K [ ] 3850 K [ ] 3800 K [X] 3750 K [ ] 3700 K [ ] 3650 K

In reality T_f is about 2400 K. The significantly higher temperatures obtained in calculations arise from the assumptions that the heat capacities (C_p) are independent from temperature and that the system consists only of H_2, O_2 and H_2O. In practice, C_p also depends on temperature, and additional species such as H, OH, and O are present, which affect the energy balance and reduce the final temperature.

7.

$H_2 + 0.5O_2 \rightarrow H_2O$ K_1

$N_2 + O_2 \rightarrow 2NO$ K_2

K_p values for these reactions at 3200 K :

$$K_1 = \exp\left\{-\frac{\left(-241800 - 6.1(3200 - 298)\right) + 3200\left(44.6 + 6.1\ln\left(\frac{3200}{298}\right)\right]}{R \times 3200}\right\} = 14.12$$

$$K_2 = \exp\left\{-\frac{\left(180600 + 1.3(3200 - 298)\right) - 3200\left(24.8 + 1.3\ln\left(\frac{3200}{298}\right)\right]}{\mathrm{R} \times 3200}\right\} = 0.028$$

If 1 mol of hydrogen is used for combustion, $n(O_2)_{initial}$ = 0.5 mol and $n(N_2)_{initial}$ = 2 mol

	$H_{2(g)}$	+	$0.5O_{2(g)}$	→	$H_2O_{(g)}$
Initially	1		0.5		-
Reacted	$-x$		$-0.5x-y$		x
Equilibrium	$1-x$		$0.5-0.5x-y$		x

	$N_{2(g)}$	+	$O_{2(g)}$	→	$2NO_{(g)}$
Initially	2		0.5		-
Reacted	$-y$		$-0.5x-y$		$2y$
Equilibrium	$2-y$		$0.5-0.5x-y$		$2y$

It can be seen that $K_1 \gg K_2$, so it can be assumed that most of the oxygen is lost as a result of reaction 1.

	$H_{2(g)}$	+	$0.5O_{2(g)}$	→	$H_2O_{(g)}$
Initially	1		0.5		-
Reacted	$-x$		$-0.5x$		x
Equilibrium	$1-x$		$0.5-0.5x$		x

$$\sum n = 1 + 0.5 + 2 - 0.5x = 3.5 - 0.5x$$

$$K_p = \frac{\frac{P_{H_2O}}{P^\circ}}{\frac{P_{H_2}}{P^\circ} \times \sqrt{\frac{P_{O_2}}{P^\circ}}} = \frac{\frac{x}{3.5-0.5x} \times \frac{P}{P^\circ}}{\frac{1-x}{3.5-0.5x} \times \frac{P}{P^\circ} \times \sqrt{\frac{0.5-0.5x}{3.5-0.5x} \times \frac{P}{P^\circ}}}; \qquad P = P^\circ = 1 \text{ bar}$$

$$K_p = \frac{x \times \sqrt{3.5-0.5x}}{(1-x) \times \sqrt{0.5-0.5x}} = 14.12$$

$$x = 0.742 \text{ mol}$$

(without $K_1 \gg K_2$ assumption, x = 0.719 mol)

$$0.028 = \frac{(2y)^2}{(2-y) \times (0.5 - 0.5 \times 0.742 - y)}$$

$$y = 0.036 \text{ mol}$$

(without $K_1 \gg K_2$ assumption, y = 0.038 mol)

$$\textbf{Mole fraction}(\textbf{NO}) = \frac{2 \times 0.036}{3.5 - 0.5 \times 0.742} = 0.023$$

(without $K_1 \gg K_2$ assumption, mole fraction (NO) = 0.024)

8.

[**X**] Lower the combustion temperature.
[ ] Decrease the excess air ratio significantly.
[ ] Raise the compression ratio in the engine.
[ ] Increase the fuel flow speed.

Problem 7. 150 Years of Urease

1. The rate of product formation is as follows:

$$\frac{d[\mathrm{P}]}{dt} = k_2[\mathrm{Urease} * \mathrm{CO(NH_2)_2}]$$

Use the steady-state approximation for $[\mathrm{Urease} * \mathrm{CO(NH_2)_2}]$:

$$\frac{d[\mathrm{Urease} * \mathrm{CO(NH_2)_2}]}{dt} = \\ = k_1[\mathrm{Urease}][\mathrm{CO(NH_2)_2}] - k_{-1}[\mathrm{Urease} * \mathrm{CO(NH_2)_2}] \\ - k_2[\mathrm{Urease} * \mathrm{CO(NH_2)_2}] = 0$$

Derive an equation for $[\mathrm{Urease} * \mathrm{CO(NH_2)_2}]$:

$$[\mathrm{Urease} * \mathrm{CO(NH_2)_2}] = \frac{k_1}{k_{-1}+k_2}[\mathrm{Urease}][\mathrm{CO(NH_2)_2}]$$

Express the rate of product formation in terms of the equilibrium concentrations: [Urease] and $[\mathrm{CO(NH_2)_2}]$:

$$\frac{d[\mathrm{P}]}{dt} = \frac{k_1k_2}{k_{-1}+k_2}[\mathrm{Urease}][\mathrm{CO(NH_2)_2}]$$

Use mass balance for $[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}$ considering $[\mathrm{CO(NH_2)_2}]_{\mathrm{total}} \gg [\mathrm{Urease}]_{\mathrm{total}}$:

$$[\mathrm{CO(NH_2)_2}]_{\mathrm{total}} = [\mathrm{CO(NH_2)_2}] + [\mathrm{Urease} * \mathrm{CO(NH_2)_2}] \approx [\mathrm{CO(NH_2)_2}]$$

Use mass balance for $[\mathrm{Urease}]_{\mathrm{total}}$ considering acid-base equilibrium:

$$[\mathrm{Urease}]_{\mathrm{total}} = \frac{[\mathrm{Urease}]}{\alpha_{\mathrm{HU^-}}} + \frac{[\mathrm{Urease}*\mathrm{CO(NH_2)_2}]}{\alpha_{\mathrm{HU^-}}} = \frac{[\mathrm{Urease}]+\frac{k_1}{k_{-1}+k_2}[\mathrm{Urease}][\mathrm{CO(NH_2)_2}]}{\alpha_{\mathrm{HU^-}}},$$

$$[\mathrm{Urease}] = \frac{[\mathrm{Urease}]_{\mathrm{total}}\alpha_{\mathrm{HU^-}}}{1+\frac{k_1}{k_{-1}+k_2}[\mathrm{CO(NH_2)_2}]}$$

Determine $\alpha_{\mathrm{HU^-}}$:

$$\alpha_{\mathrm{HU^-}} = \frac{[\mathrm{HU^-}]}{[\mathrm{H_2U}]+[\mathrm{HU^-}]+[\mathrm{U^{2-}}]} = \frac{[\mathrm{HU^-}]}{\frac{[\mathrm{H^+}][\mathrm{HU^-}]}{K_{a1}}+[\mathrm{HU^-}]+\frac{K_{a2}[\mathrm{HU^-}]}{[\mathrm{H^+}]}} = \frac{1}{\frac{[\mathrm{H^+}]}{K_{a1}}+1+\frac{K_{a2}}{[\mathrm{H^+}]}}.$$

Express the rate of product formation in terms of the total concentrations of Urease, $\mathrm{CO(NH_2)_2}$ and $\mathrm{H^+}$:

$$\frac{d[\mathrm{P}]}{dt} = \frac{k_1k_2}{k_{-1}+k_2} \cdot \frac{[\mathrm{Urease}]_{\mathrm{total}}}{1+\frac{k_1}{k_{-1}+k_2}[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}} \cdot \frac{1}{\frac{[\mathrm{H^+}]}{K_{a1}}+1+\frac{K_{a2}}{[\mathrm{H^+}]}}[\mathrm{CO(NH_2)_2}]_{\mathrm{total}} =$$

$$= \frac{k_2[\mathrm{Urease}]_{\mathrm{total}}[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}}{\left(\frac{k_{-1}+k_2}{k_1}+[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}\right)\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)}$$

2.

a) $\frac{d[\mathrm{P}]}{dt} = \frac{k_2[\mathrm{Urease}]_{\mathrm{total}}[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}}{\left(\frac{k_{-1}+k_2}{k_1}\right)\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)}$. So n(Urease) = 1 and $n(\mathrm{CO(NH_2)_2})$ = 1.

b) $\frac{d[\mathrm{P}]}{dt} = \frac{k_2[\mathrm{Urease}]_{\mathrm{total}}[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}}{[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)} = \frac{k_2[\mathrm{Urease}]_{\mathrm{total}}}{\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)}$. So n(Urease) = 1 and $n(\mathrm{CO(NH_2)_2})$ = 0.

3.

$$r = \frac{d[\mathrm{P}]}{dt} = \frac{k_2[\mathrm{Urease}]_{\mathrm{total}}[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}}{\left(\frac{k_{-1}+k_2}{k_1}+[\mathrm{CO(NH_2)_2}]_{\mathrm{total}}\right)\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)} = \frac{\mathrm{const}}{\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)}.$$

$$\text{At } r_{\max} \rightarrow \frac{dr}{d[\mathrm{H^+}]} = 0 \rightarrow r' = \frac{-\mathrm{const}\cdot\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)'}{\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right)^2} = 0 \rightarrow \frac{1}{K_{a1}} - \frac{K_{a2}}{[\mathrm{H^+}]^2} = 0$$

Or we can just differentiate the denominator (it should have minimum):

$$\frac{d\left(1+\frac{[\mathrm{H^+}]}{K_{a1}}+\frac{K_{a2}}{[\mathrm{H^+}]}\right))}{d[\mathrm{H^+}]} = 0 \rightarrow \frac{1}{K_{a1}} - \frac{K_{a2}}{[\mathrm{H^+}]^2} = 0$$

$$[\mathrm{H^+}] = \sqrt{K_{a1}K_{a2}} \rightarrow \mathrm{p}H = \frac{\mathrm{p}K_{a1}+\mathrm{p}K_{a2}}{2} = \frac{5.62+9.07}{2} = 7.35$$

4.

a) Products of decomposition of 0.005 M $\mathrm{CO(NH_2)_2}$ are 0.005 M $\mathrm{CO_2}$ and 0.010 M $\mathrm{NH_3}$. Firstly, ammonium bicarbonate is formed, because the equilibrium constant for this reaction is sufficiently large:

$$\mathrm{CO_2 + NH_3 + H_2O \overset{K}{\underset{\leftarrow}{\rightarrow}} HCO_3^- + NH_4^+},$$

$$K = \frac{K_bK_{a1}}{K_w} = \frac{[\mathrm{NH_4^+}][\mathrm{OH^-}]}{[\mathrm{NH_3}]}\cdot\frac{[\mathrm{HCO_3^-}][\mathrm{H^+}]}{[\mathrm{CO_2}]}\cdot\frac{1}{[\mathrm{H^+}][\mathrm{OH^-}]} = \frac{[\mathrm{NH_4^+}][\mathrm{HCO_3^-}]}{[\mathrm{NH_3}][\mathrm{CO_2}]} = \frac{10^{-4.75}\cdot 10^{-3.75}}{10^{-14}} = 3.16\cdot 10^5\ .$$

After this reaction, we can assume that the solution contains 0.005 M $\mathrm{HCO_3^-}$, 0.005 M $\mathrm{NH_4^+}$ and 0.005 M $\mathrm{NH_3}$. Residual ammonia can react with bicarbonate to form carbonate as follows:

$$\mathrm{HCO_3^- + NH_3 \overset{K}{\underset{\leftarrow}{\rightarrow}} CO_3^{2-} + NH_4^+},$$

$$K = \frac{K_bK_{a2}}{K_w} = \frac{[\mathrm{NH_4^+}][\mathrm{OH^-}]}{[\mathrm{NH_3}]}\cdot\frac{[\mathrm{CO_3^{2-}}][\mathrm{H^+}]}{[\mathrm{HCO_3^-}]}\cdot\frac{1}{[\mathrm{H^+}][\mathrm{OH^-}]} = \frac{[\mathrm{NH_4^+}][\mathrm{CO_3^{2-}}]}{[\mathrm{NH_3}][\mathrm{HCO_3^-}]} = \frac{10^{-4.75}\cdot 10^{-10.33}}{10^{-14}} = 8.32\cdot 10^{-2}.$$

If x M carbonate is formed:

$$8.32\cdot 10^{-2} = \frac{x\cdot(x+0.005)}{(0.005-x)\cdot(0.005-x)} \rightarrow x = 3.39\cdot 10^{-4}$$

$$\text{So } [\mathrm{H^+}] = \frac{K_{a2}[\mathrm{HCO_3^-}]}{[\mathrm{CO_3^{2-}}]} = \frac{10^{-10.33}\cdot(0.005-3.39\cdot 10^{-4})}{3.39\cdot 10^{-4}} = 6.43\cdot 10^{-10} \quad \text{or}$$

$[H^+] = \frac{K_w[NH_4^+]}{K_b[NH_3]} = \frac{10^{-14}\cdot(0.005+3.39\cdot10^{-4})}{10^{-4.75}\cdot(0.005-3.39\cdot10^{-4})} = 6.44\cdot10^{-10} \rightarrow pH = 9.19$.

b) Let's simplify the kinetic equation as follows:

$$r = \frac{d[P]}{dt} = \frac{k_2[\text{Urease}]_{total}[CO(NH_2)_2]_{total}}{\left(\frac{k_{-1}+k_2}{k_1}+[CO(NH_2)_2]_{total}\right)\left(1+\frac{[H^+]}{K_{a1}}+\frac{K_{a2}}{[H^+]}\right)} = \frac{const}{\left(1+\frac{[H^+]}{K_{a1}}+\frac{K_{a2}}{[H^+]}\right)}$$

At pH = 7.00: $r(7.00) = \frac{const}{\left(1+\frac{10^{-7.00}}{10^{-5.62}}+\frac{10^{-9.07}}{10^{-7.00}}\right)} = 0.9522\cdot const.$

At pH = 9.19: $r(9.19) = \frac{const}{\left(1+\frac{10^{-9.19}}{10^{-5.62}}+\frac{10^{-9.07}}{10^{-9.19}}\right)} = 0.4313\cdot const.$

So, $\frac{r(7.00)}{r(9.19)} = \frac{0.9522\cdot const}{0.4313\cdot const} = 2.21.$

5. (1) – (c); (2) – (a); (3) – (b).

(1) When Urease is in the buffer solution with $CO(NH_2)_2$ the concentration of the product will increase at a constant pH until $CO(NH_2)_2$ runs out. This corresponds to option (c).

(2) When Urease is in the neutral solution with $CO(NH_2)_2$ and without any buffer, as the concentration of the product increases, the pH of the solution will shift towards an alkaline environment, which will inhibit Urease and slow down the enzymatic reaction. This corresponds to option (a).

(3) When Urease is in the liposomes with permeable membranes with $CO(NH_2)_2$ and without any buffer inside, first, in the liposome $CO(NH_2)_2$ decomposes under the action of Urease, thereby increasing the pH of the liposome and inhibiting Urease, but then the products leave the liposome (Urease remains in the liposome!), returning the pH and Urease activity to their original state. This corresponds to option (b).

6. The solution in question 4a contains $[CO_3^{2-}] = 3.39\cdot10^{-4}\ M$.
So, $[Ca^{2+}]\cdot3.39\cdot10^{-4} > 3.36\cdot10^{-9} \rightarrow [Ca^{2+}] > 9.91\cdot10^{-6}$ M

7. The rate of product formation from equilibrium concentrations of [Urease] and $[CO(NH_2)_2]$ remains as:

$$\frac{d[P]}{dt} = \frac{k_1k_2}{k_{-1}+k_2}[\text{Urease}][CO(NH_2)_2]$$

For $[CO(NH_2)_2]_{total}$, the mass balance equation doesn't change:

$$[CO(NH_2)_2]_{total} = [CO(NH_2)_2] + [\text{Urease} * CO(NH_2)_2] \approx [CO(NH_2)_2]$$

Use mass balance for $[B(OH)_3]_{total}$ considering $[B(OH)_3]_{total} \gg [\text{Urease}]_{total}$:

$$[B(OH)_3]_{total} = [B(OH)_3] + [\text{Urease} * B(OH)_3] \approx [B(OH)_3]$$

Use mass balance for $[\text{Urease}]_{total}$:

$$[\text{Urease}]_{\text{total}} = \frac{[\text{Urease}]}{\alpha_{\text{HU}^-}} + \frac{[\text{Urease}*\text{CO(NH}_2)_2]}{\alpha_{\text{HU}^-}} + \frac{[\text{Urease}*\text{B(OH)}_3]}{\alpha_{\text{HU}^-}} =$$

$$= \frac{[\text{Urease}] + \frac{k_1}{k_{-1}+k_2}[\text{Urease}][\text{CO(NH}_2)_2] + \frac{[\text{Urease}][\text{B(OH)}_3]}{K_\text{I}}}{\alpha_{\text{HU}^-}},$$

$$[\text{Urease}] = \frac{[\text{Urease}]_{\text{total}}\alpha_{\text{HU}^-}}{1 + \frac{k_1}{k_{-1}+k_2}[\text{CO(NH}_2)_2] + \frac{[\text{B(OH)}_3]}{K_\text{I}}},$$

$$\alpha_{\text{HU}^-} = \frac{[\text{HU}^-]}{[\text{H}_2\text{U}]+[\text{HU}^-]+[\text{U}^{2-}]} = \frac{[\text{HU}^-]}{\frac{[\text{H}^+][\text{HU}^-]}{K_{a1}}+[\text{HU}^-]+\frac{K_{a2}[\text{HU}^-]}{[\text{H}^+]}} = \frac{1}{\frac{[\text{H}^+]}{K_{a1}}+1+\frac{K_{a2}}{[\text{H}^+]}}.$$

Express the rate of product formation in terms of the total concentrations of Urease, $CO(NH_2)_2$, $B(OH)_3$, and H^+:

$$\frac{d[\text{P}]}{dt} = \frac{k_1k_2}{k_{-1}+k_2} \cdot \frac{[\text{Urease}]_{\text{total}}}{1+\frac{k_1}{k_{-1}+k_2}[\text{CO(NH}_2)_2]_{\text{total}}+\frac{[\text{B(OH)}_3]_{\text{total}}}{K_\text{I}}} \cdot \frac{1}{\frac{[\text{H}^+]}{K_{a1}}+1+\frac{K_{a2}}{[\text{H}^+]}} \cdot [\text{CO(NH}_2)_2]_{\text{total}} =$$

$$= \frac{k_2[\text{Urease}]_{\text{total}}[\text{CO(NH}_2)_2]_{\text{total}}}{\left(\frac{k_{-1}+k_2}{k_1}\left(1+\frac{[\text{B(OH)}_3]_{\text{total}}}{K_\text{I}}\right)+[\text{CO(NH}_2)_2]_{\text{total}}\right)\left(1+\frac{[\text{H}^+]}{K_{a1}}+\frac{K_{a2}}{[\text{H}^+]}\right)}.$$

8. Make the linear function as follows:

$$r = \frac{k_2[\text{Urease}]_{\text{total}}[\text{CO(NH}_2)_2]_{\text{total}}}{\left(\frac{k_{-1}+k_2}{k_1}\left(1+\frac{[\text{B(OH)}_3]_{\text{total}}}{K_\text{I}}\right)+[\text{CO(NH}_2)_2]_{\text{total}}\right)\left(1+\frac{[\text{H}^+]}{K_{a1}}+\frac{K_{a2}}{[\text{H}^+]}\right)} = \frac{r_{\max}[\text{CO(NH}_2)_2]_{\text{total}}}{K_\text{M}\left(1+\frac{[\text{B(OH)}_3]_{\text{total}}}{K_\text{I}}\right)+[\text{CO(NH}_2)_2]_{\text{total}}},$$

$$\frac{1}{r} = \frac{K_\text{M}\left(1+\frac{[\text{B(OH)}_3]_{\text{total}}}{K_\text{I}}\right)}{r_{\max}} \cdot \frac{1}{[\text{CO(NH}_2)_2]_{\text{total}}} + \frac{1}{r_{\max}} \rightarrow \text{slope} = \frac{K_\text{M}\left(1+\frac{[\text{B(OH)}_3]_{\text{total}}}{K_\text{I}}\right)}{r_{max}}.$$

Using the ratio of the slopes we find the value of K_I:

$$2.48 = \frac{1+\frac{3.0\cdot10^{-4}}{K_\text{I}}}{1+\frac{7.5\cdot10^{-5}}{K_\text{I}}} \rightarrow K_\text{I} = 7.7\cdot10^{-5}.$$

9. It is necessary to solve the following system of equations:

$$\begin{cases} \Delta_\text{r}H° - 298.15\cdot\Delta_\text{r}S° = -R\cdot298.15\cdot ln\left(\frac{1}{7.7\cdot10^{-5}}\right) \\ \Delta_\text{r}H° - 308.15\cdot\Delta_\text{r}S° = -R\cdot308.15\cdot ln\left(\frac{1}{1.57\cdot10^{-4}}\right) \end{cases} \rightarrow \begin{cases} \Delta_\text{r}H° = -54.42\ \text{kJ mol}^{-1} \\ \Delta_\text{r}S° = -103.8\ \text{J mol}^{-1}K^{-1} \end{cases}$$

Problem 8. Step-by-step

1. The overall kinetic equation is:

$$r = k[\mathrm{OH}]^x[\mathrm{COOH}]^y$$

We will use the values from the table to evaluate the values for *k*, *x*, and *y*.

$\frac{k[0.2]^y}{k[0.2]^x[0.2]^y} = \frac{6}{4}$, so, we have $\left(\frac{0.3}{0.2}\right)^x = \frac{6}{4}$, so $\boldsymbol{x = 1}$.

We do the same for *y*,

$\frac{k[0.2][0.2]^y}{k[0.1][0.1]^y} = 4$, or $2 \cdot \frac{[0.2]^y}{[0.1]^y} = 4$, so $\boldsymbol{y = 1}$.

Now we need to find the value of *k* by simply substituting the values of *x* and *y*.

$$2.7 \cdot 10^{-4} = k[0.1]^1[0.1]^1$$

$$\boldsymbol{k = 0.027\ \mathrm{M^{-1}s^{-1}}}$$

2. $z = x+y = 2$.

For a second order reaction with the equal initial concentrations:

$$\frac{1}{C} - \frac{1}{C_0} = kt$$

3. The degree of conversion is as follows:

$$p = 1 - \frac{C}{C_0}$$

$DP = \frac{1}{1-p}$ (Carothers equation).

4. Degree of polymerisation is as follows:

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

$\frac{C_0}{C} - 1 = C_0kt$, so $DP(t) = C_0kt + 1$, thus a linear equation.

5. Therefore, $\frac{1}{1-p(t)} = C_0kt + 1$ – a linear relationship between $\frac{1}{1-p}$ and *t*.

6. Equilibrium constants for these reactions are expressed as the ratio between forward and reverse rate constants:

$$K_\mathrm{F} = \frac{[\mathrm{ArF} \cdots 2\mathrm{B}]}{[\mathrm{ArF}][\mathrm{B}]^2} = \frac{k_{1,\mathrm{F}}}{k_{-1,\mathrm{F}}}$$

$$K_{\mathrm{Cl}} = \frac{[\mathrm{ArCl} \cdots \mathrm{B}]}{[\mathrm{ArCl}][\mathrm{B}]} = \frac{k_{1,\mathrm{Cl}}}{k_{-1,\mathrm{Cl}}}$$

7. Rate of reaction is determined by the final step,

$r = \frac{d[P]}{dt} = k_{\mathrm{F}}[\mathrm{ArF} \cdots 2\mathrm{B}]$ for DFDPS,

using the equilibrium constant, we can express the intermediate concentration,

$$[\mathrm{ArF} \cdots 2\mathrm{B}] = K_{\mathrm{F}}[\mathrm{ArF}][\mathrm{B}]^2$$

So, the rate law is:

$$r = \frac{d[P]}{dt} = k_{\mathrm{F}}K_{\mathrm{F}}[\mathrm{ArF}][\mathrm{B}]^2$$

with $k_{\mathrm{obs}} = k_{\mathrm{F}}K_{\mathrm{F}}$.

Similarly, for DCDPS:

$$r = \frac{d[P]}{dt} = k_{\mathrm{Cl}}K_{\mathrm{Cl}}[\mathrm{ArCl}][\mathrm{B}]$$

with $k_{\mathrm{obs}} = k_{\mathrm{Cl}}K_{\mathrm{Cl}}$.

8. In the case of DFPDS (order of the reaction $n = 3$) $[\mathrm{ArF}] = [\mathrm{B}] = [\mathrm{M}]$:

$$\frac{1}{[\mathrm{M}]^2} - \frac{1}{[\mathrm{M}]_0^2} = 2k_{\mathrm{F}}K_{\mathrm{F}}t$$

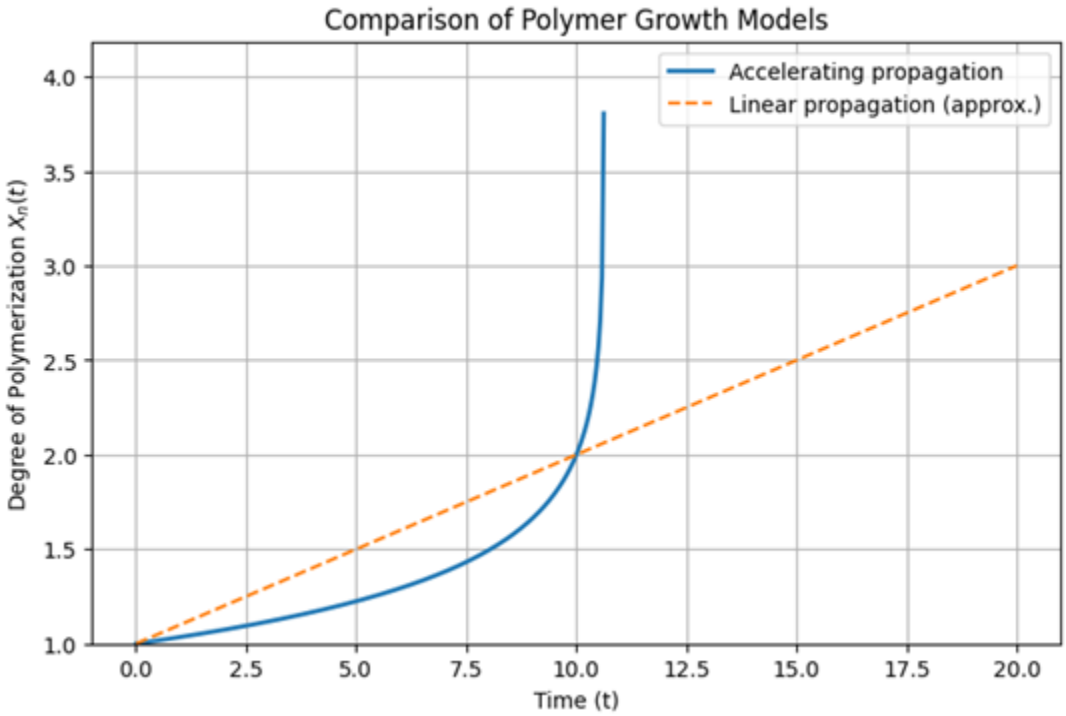
Degree of polymerisation: $DP = \frac{[\mathrm{M}]_0}{[\mathrm{M}]}$

$$\frac{[\mathrm{M}]_0{}^2 - [\mathrm{M}]^2}{[\mathrm{M}]^2} = [M]_0{}^2 2k_{\mathrm{F}}K_{\mathrm{F}}t$$

According to the rate equation, $k_{\mathrm{obs}} = \mathrm{k}_{\mathrm{F}}K_{\mathrm{F}}$, then

$DP^2 - 1 = [M]_0{}^2 2k_{\mathrm{F}}K_{\mathrm{F}}t = 0.1633^2 \times 2 \times 7023 \times 1 \Rightarrow DP = 19.4.$

9. If we want to sketch qualitatively, then we start from DP = 1 in both cases:



In the second case ($\xi > 1$, blue line), propagation step accelerates by rate ratio ξ after every step of propagation, so eventually reaction proceeds so fast that it obtains exponential character.

10. So, for normal Carothers equation we have:

$$DP = \frac{1}{1-p}$$

How much time is needed to form a chain of length n? Let's find it:

$$T_n = \frac{1}{kc_0} + \frac{1}{k\xi c_0} + \frac{1}{k\xi^2 c_0} + \cdots = \sum_{i=1}^{n-1} \frac{1}{k\xi^{i-1} c_0} = \frac{1}{kc_0} \sum_{i=1}^{n-1} \frac{1}{\xi^{i-1}} = \frac{1}{kc_0} \sum_{i=0}^{n-2} \frac{1}{\xi^i}$$

So we came up with the sum of geometric progression:

$$T_n = \frac{1-\xi^{-(n-1)}}{kc_0(1-\xi^{-1})} = \frac{\xi-\xi^{-n+2}}{kc_0(\xi-1)}$$

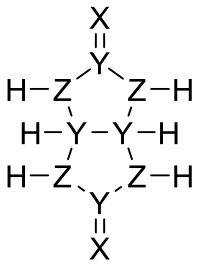
$$\xi^{-n+2} = \xi - kc_0(\xi-1)T_n$$

$$(-n+2)\ln\xi = \ln(\xi - kc_0(\xi-1)T_n)$$

$$DP = n = 2 - \frac{\ln(\xi - kc_0(\xi-1)T_n)}{\ln\xi} = 2 - \frac{\ln(\xi - kc_0(\xi-1)t)}{\ln\xi}$$

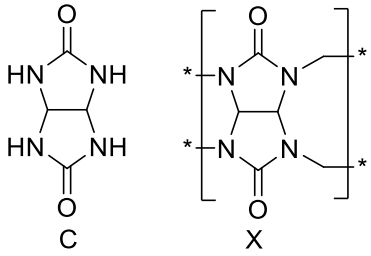
Problem 9. Host-guest chemistry of cucurbiturils as Uzbek hospitality

1. Mass spectra, one type of carbon and hydrogen and positive Tollens reaction of **B** correspond to the aldehyde group - CHO. This means **B** is glyoxal. From the figure, you can see that cucurbit[6]uril has six rotating links. Formaldehyde connects these fragments. Then **C** has the structure:



where most likely X is oxygen, Y is carbon and Z is nitrogen. In the synthesis of **C**, 1 mol of glyoxal and 2 mol of **A** are used. If we take this from the structure of **C**, then **A** should contain eight atoms, of which one atom is most likely oxygen and one carbon atom bound to it, and two are nitrogen, which condenses with the aldehyde. Then four atoms remain for hydrogen. Thus, the formula of **A** should be NH_2CONH_2. This corresponds to urea, meaning salt **Y** will be NH_4OCN.

Y: NH_4OCN, **A:** $CO(NH_2)_2$, **B:** $(CHO)_2$.



2. The molar mass of tropicamide is 284 g/mol, and for CB_8 1328 g/mol. 285 corresponds to $[TRH]^+$ (m/z = (284+1)/1). The peak with 665 with z = 1 does not have a suitable answer, then with z = 2 m = 1330 which corresponds to $[CB_8+2H]^{2+}$. Similarly, for peak 807 with z = 1 there is no answer, and for z = 2, m = 1614. This corresponds to $[TR{\cdot}CB_8+2H]^{2+}$.

3. The equilibria that occur:

$TRH^+ \rightleftharpoons TR + H^+$ $\qquad pK_a$

$TRH^+ + CB_8 \rightleftharpoons TRH^+{\cdot}CB_8$ $\qquad K_{TRH^+}$

$TR + CB_8 \rightleftharpoons TR{\cdot}CB_8$ $\qquad K_{TR}$

$TRH^+ \cdot CB_8 \rightleftharpoons TR \cdot CB_8 + H^+$ $\qquad pK_a'$

Let's start with a pure TR solution. At pH ≈ 3, the TRH^+ form predominates, and at pH ≈ 7, the TR form predominates, and accordingly, the absorbances A_{max} and A_{min} depend only on their concentrations respectively. At the point $(A_{max} + A_{min})/2$, $[TRH^+] = [TR]$, therefore $K_a = [H^+]$, and from the graph, $pK_a \approx 5.2$.

Similarly, for the complex, $pK_a' \approx 6.7$

From the thermodynamic cycle we find $K_{TRH^+} = K_{TR} \cdot \frac{K_a}{K_a'} = 4.42 \cdot 10^5$.

4. At pH = 4.0 the molar fraction of protonated form of tropicamide is higher:

$$\alpha_{TRH^+} = \frac{[TRH^+]}{[TRH^+]+[TRH]} = \frac{[H^+]}{[H^+]+K_a} = \frac{10^{-4.0}}{10^{-4.0}+10^{-5.2}} = 0.941$$

So, biologically active form is TRH^+.

5. Without CB_8:

$$\alpha_{TRH^+} = \frac{[TRH^+]}{[TRH^+]+[TRH]} = \frac{[H^+]}{[H^+]+K_a} = \frac{10^{-7.0}}{10^{-7.0}+10^{-5.2}} = 0.0156$$

With CB_8:

$$\alpha_{[TRH^+ \cdot CB_8]} = \frac{[TRH^+ \cdot CB_8]}{[TRH^+ \cdot CB_8]+[TR \cdot CB_8]} = \frac{[H^+]}{[H^+]+K_a'} = \frac{10^{-7.0}}{10^{-7.0}+10^{-6.7}} = 0.334$$

The molar fraction of the protonated form of tropicamide increases by $\frac{0.334}{0.0156} = 21.4$ times in the presence of CB_8.

6. Reaction order for CPD is 2 due to the linear dependence of 1/[CPD] on reaction time. This dependence is as follows:

$$\frac{1}{[CPD]} = \frac{1}{[CPD]_0} + k_{obs}t \rightarrow \text{slope} = k_{obs} = \frac{850 - 300}{4000 - 0} = 0.14\ M^{-1}\ s^{-1}$$

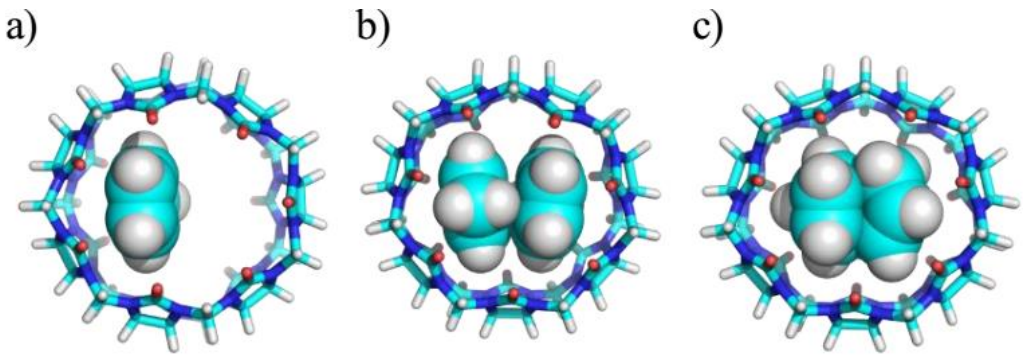
7. The rate of the dimerisation reaction is as follows:

$$r = \frac{d[DCPD \cdot CB_7]}{dt} = k[(CPD)_2 \cdot CB_7] = kK_2[CPD][CPD \cdot CB_7] = kK_1K_2[CPD]^2[CB_7].$$

8. The rate of CPD dimerisation in the presence of CB_5, which is too small to encapsulate CPD, remains similar to that in water. There is neither complexation with CPD nor DCPD.

CB_6 is sufficiently large to encapsulate an individual CPD molecule. A CB_6-isolated CPD substrate is inaccessible to dimerisation and, unsurprisingly, addition of CB_6 markedly inhibits the reaction.

The intermediary-sized CB_7 accelerates the reaction by *ca.* five orders of magnitude. The dimerisation inside CB_7 proceeds in the tight-packing regime. This allows for transition-state stabilisation to occur.

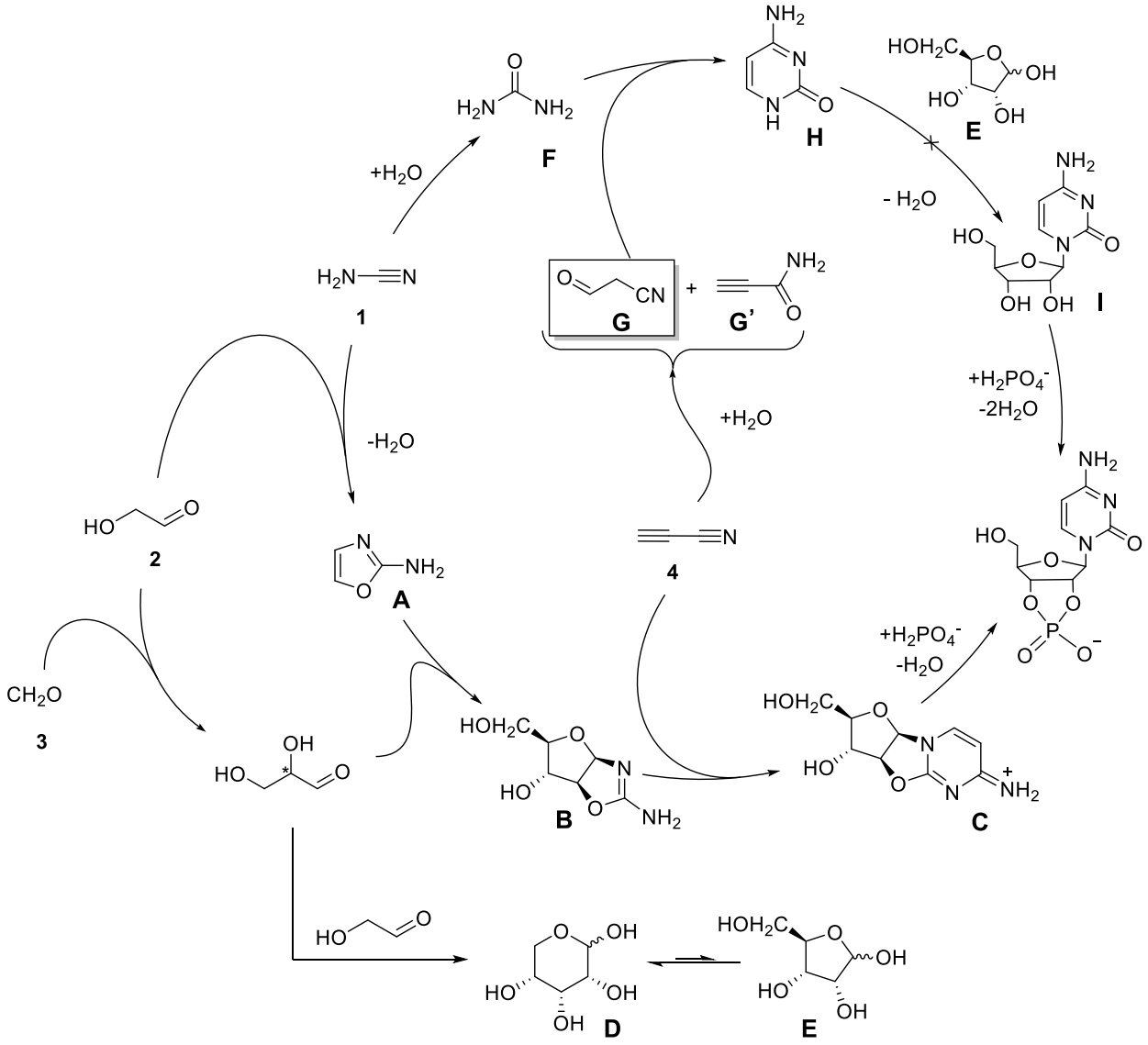


a) one CPD unit, b) two CPD units, and c) DCPD inside CB_7

For CB_8, the reaction proceeds in the loose-packing regime. In this case, a further contraction in the guest volume along the reaction coordinate would decrease rather than increase the overall complex stability, and no rate acceleration occurs for this homologue.

Problem 10. Prebiotic life

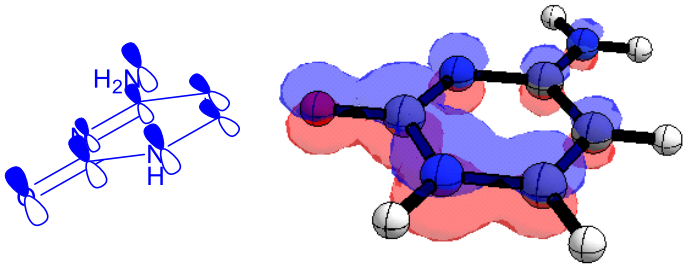
1. This is the Powner scheme:



The acyclic form of compound **D** is also considered as the correct answer.

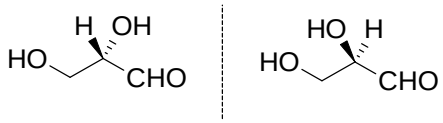
2.

- **electron delocalisation due to the aromaticity of pyrimidine ring**
- extreme ring strain in cytosine that interrupts the reaction
- the positive charge due to formation of ammonium salts that repels other molecules.



π-conjugation hinders nucleophilic activity of nitrogen atom

3.



L-glyceraldehyde D-glyceraldehyde

4.

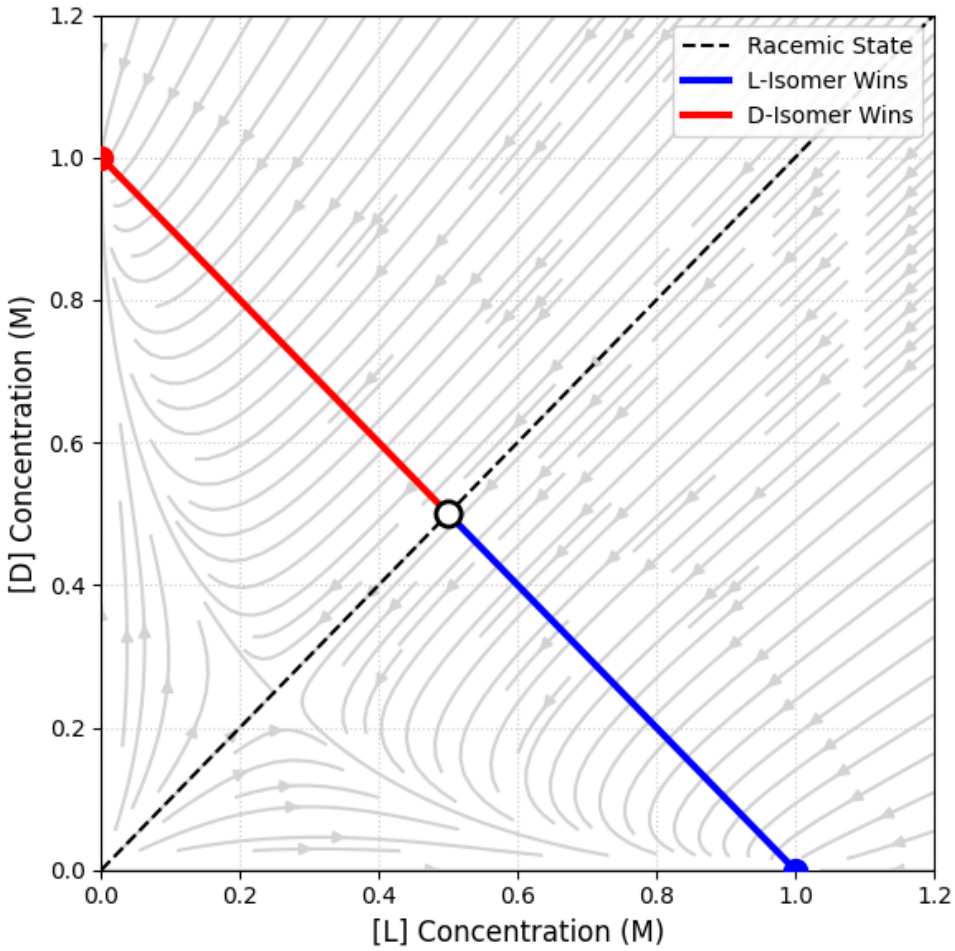
$$\frac{k_{\mathrm{L}}}{k_{\mathrm{D}}} = \exp\left(\frac{2H_{\mathrm{SO}}}{k_{\mathrm{B}}T}\right) \Rightarrow ee = \frac{[\mathrm{LY}] - [\mathrm{DY}]}{[\mathrm{LY}] + [\mathrm{DY}]} = \frac{(k_{\mathrm{L}} - k_{\mathrm{D}})}{(k_{\mathrm{L}} + k_{\mathrm{D}})} = \frac{\exp\left(\frac{2H_{\mathrm{SO}}}{k_{\mathrm{B}}T}\right) - 1}{\exp\left(\frac{2H_{\mathrm{SO}}}{k_{\mathrm{B}}T}\right) + 1}$$

$$\frac{k_{\mathrm{L}}}{k_{\mathrm{D}}} = \exp\left(2 \times 5 \times 10^{-4} \times 1.602 \times \frac{10^{-19}}{313 \times 1.38 \times 10^{-23}}\right) = 1.038$$

$$ee = \frac{1.038 - 1}{1.038 + 1} = 0.0185 = 1.85\% \approx 2\%$$

Very low magnitudes of spin-orbit coupling can lead to symmetry breaking in a racemic system.

5. So a trajectory that should be depicted needs to fit such criteria as starting from (0.5;0.5) point and ending in (1;0) and (0;1) points:



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

The total concentration of initial AA is defined as:

$$[1] = [L_1] + [D_1], [2] = [L_2] + [D_2] \Rightarrow \text{for instance } [L_1] = \frac{1+ee_1}{2}[1]$$

$$ee_{\text{homo}} = \frac{[LL]-[DD]}{[LL]+[DD]} = \frac{[L_1][L_2]-[D_1][D_2]}{[L_1][L_2]+[D_1][D_2]}$$

$$= \frac{\frac{1+ee_1}{2}\times\frac{1+ee_2}{2}\times[1][2] - \frac{1-ee_1}{2}\times\frac{1-ee_2}{2}\times[1][2]}{\frac{1+ee_1}{2}\times\frac{1+ee_2}{2}\times[1][2] + \frac{1-ee_1}{2}\times\frac{1-ee_2}{2}\times[1][2]}$$

$$= \frac{(1+ee_1)(1+ee_2)-(1-ee_1)(1-ee_2)}{(1+ee_1)(1+ee_2)+(1-ee_1)(1-ee_2)} = \frac{ee_1+ee_2}{1+ee_1ee_2}$$

7.

$$ee_{\text{homochiral}} = \frac{0.02 + 0.02}{0.02^2 + 1} = 0.04$$

So, the *ee* of the system rises. For more significant initial conditions, *ee* will drastically converge to 1 when chain length is increased.

For *ee* taken as 5%, the answer is 9.98%≈ 10%.

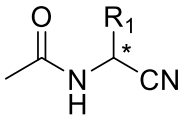
8. For final $ee_{\text{homo}} = \frac{[\text{L}_\text{n}]-[\text{D}_\text{n}]}{[\text{L}_\text{n}]+[\text{D}_\text{n}]} > 99\% \Rightarrow \chi_{\text{L}_\text{n}} > 99.5\%, \chi_{\text{D}_n} < 0.5\%$

$$\left(\frac{0.525}{0.475}\right)^n > \frac{99.5}{0.5}$$

$$n > 52.89$$

So, $n_{\text{min}} = 53$

9.



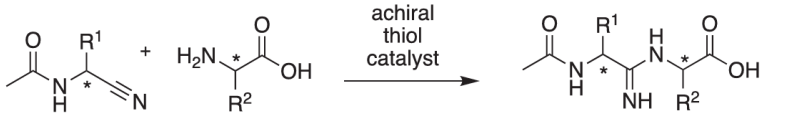
K

So, the full scheme of the model looks like this:

1A transamination

A *pyridoxamine* **aa** (**L** or **D**) *pyridoxal*

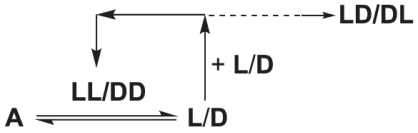
1B catalytic peptide ligation



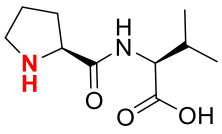
aa$_1$ (ee$_1$) **aa**$_2$ (ee$_2$) **aa**$_1$-**aa**$_2$ (**LL**, **DD**, **LD**, or **DL**)

$$ee = \frac{\mathbf{L} - \mathbf{D}}{\mathbf{L} + \mathbf{D}}$$

1C combined autocatalytic network

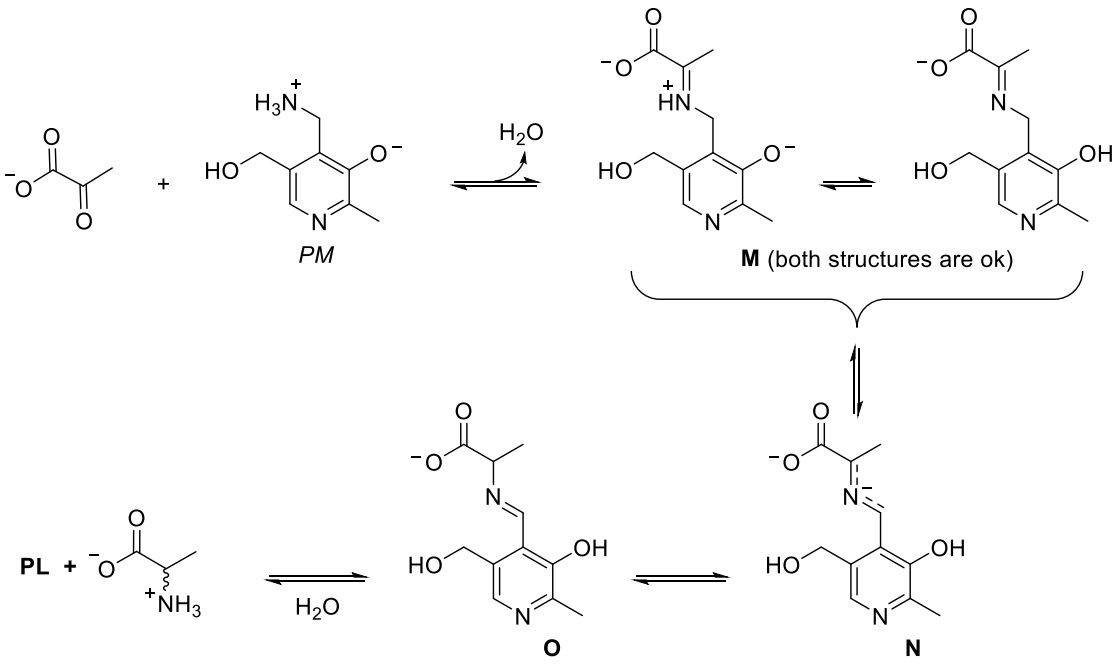


10. Proline imino groups, as they can be involved in acid-base interactions.



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

11.



12.

- **Hydrogen bonds**
- π-stacking
- **Steric repulsion**
- Alignment of stereocentres

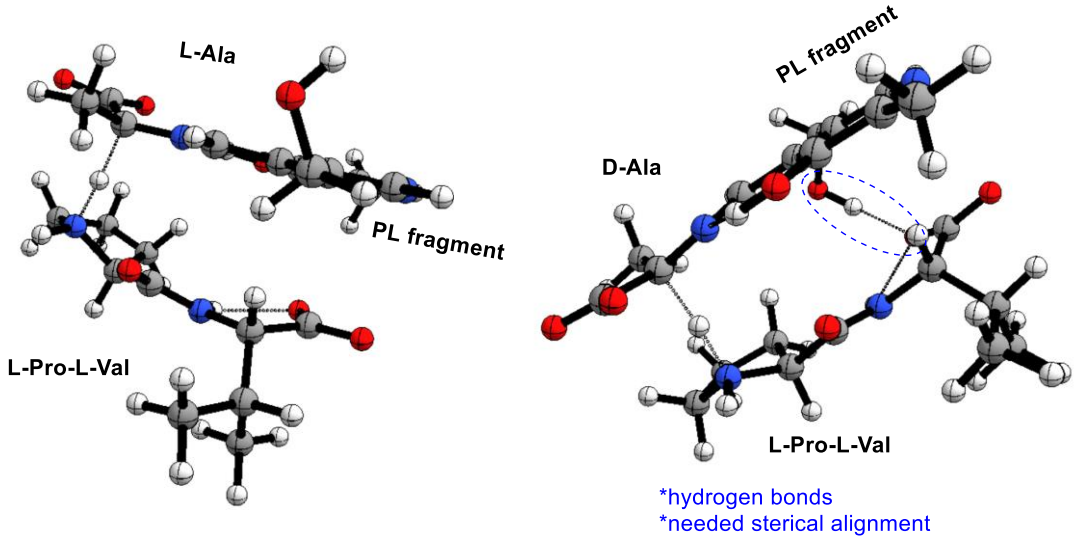
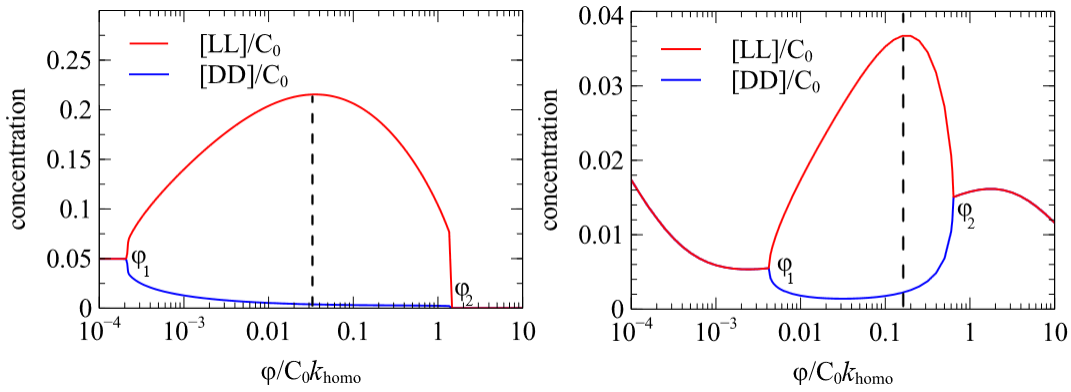


Figure 1. TS of deprotonation of alpha-hydrogen atom by L-Pro-L-Val for L- and D-Ala.

13. The maximum chiral asymmetry means the maximal difference between [LL] and [DD] values. On the given graphs, approximate lines corresponding to the maxima can be drawn at $\varphi/C_0k_{\text{homo}}$ = 0.04 for the constructive mechanism and $\varphi/C_0k_{\text{homo}}$ = 0.2 for the destructive mechanism:



14.

a) When φ is small, the system is in equilibrium and it's predominantly racemic with little noise.

b) Here we see symmetry breaking. The LL concentration rockets while the DD concentration falls. For the destructive mechanism, we feed the system with amino acids in the L- and D-form, but we've seen that generally D-amino acids get destroyed faster and amplify the *ee*, so at some flow rate the difference becomes significant. For the constructive mechanism, we feed the system with ketoacids and expect synthesis of amino acids. The symmetry breaking here is based on the same fact of D-amino acids being destroyed faster and that from insignificant ee_{monomer} of the initial amino acids we get a larger *ee* of the dipeptides (See Q6).

c) If the flow rate becomes too high, the monomers can be washed out from the system, and the dimer concentrations drop to zero. In case of the destructive mechanism, there is a significant excess of L- and D-monomers, so it prevents the appearance of any *ee*.

Problem 11. Fluorescence, phosphorescence and quenching

1. The absorbed photon excites the electron from the ground state (S_0) to an excited state (S_1). During this process, the molecule absorbs high-energy (short-wavelength) light. Before fluorescence occurs, the excited molecule loses some energy non-radiatively (as heat or vibration). The correct answer is lower than that of fluorescence.

2. ISC efficiency increases when the energy gap between S_1 and T_1 is small, because the smaller the gap, the easier the spin-orbit coupling is, hence the more probable transition between the two states. The correct answer is lower.

3. The rate of decay: $-\frac{d[S^*]}{dt} = k_F \times [S^*] + k_{IC} \times [S^*] + k_{ISC} \times [S^*]$

$=(k_F + k_{IC} + k_{ISC}) \times [S^*]$

So, $n([S^*]) = 1$ and $k_{obs} = (k_F + k_{IC} + k_{ISC})$

According to the first order kinetic equation:

$\frac{[S^*]}{e} = [S^*] \times e^{-k_{obs}t} = [S^*] \times e^{-t/\tau_0} \rightarrow \tau_0 = \frac{1}{k_{obs}} = \frac{1}{k_F+k_{IC}+k_{ISC}}$

4. $\varphi_0 = \frac{r_F}{r_F+r_{IC}+r_{ISC}} = \frac{k_F \times [S^*]}{(k_F+k_{IC}+k_{ISC})\times[S^*]} = \frac{k_F}{k_F+k_{IC}+k_{ISC}}$

5. $\varphi = \frac{r_F}{r_F+r_{IC}+r_{ISC}+r_Q} = \frac{k_F \times [S^*]}{(k_F+k_{IC}+k_{ISC}+k_Q[Q])\times[S^*]} = \frac{k_F}{k_F+k_{IC}+k_{ISC}+k_Q[Q]}$.

6. $\frac{\varphi_0}{\varphi} = \frac{k_F+k_{IC}+k_{ISC}+k_Q[Q]}{k_F+k_{IC}+k_{ISC}} = 1 + \frac{k_Q}{k_F+k_{IC}+k_{ISC}}[Q]$

7. The emission lifetime (τ) depends on the concentration of Q as follows:

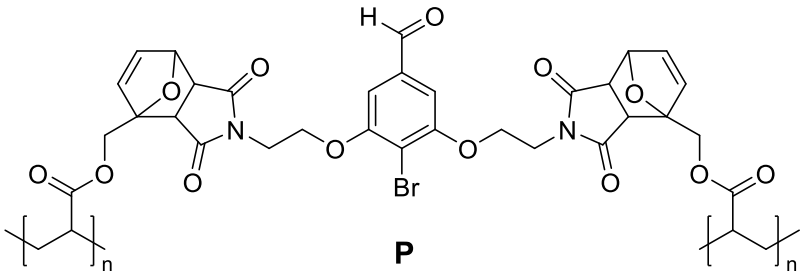
$$\tau = \frac{1}{k_{obs}} = \frac{1}{k_F + k_{IC} + k_{ISC} + k_Q[Q]} \rightarrow \frac{1}{\tau} = (k_F + k_{IC} + k_{ISC}) + k_Q[Q]$$

Using the obtained linear dependence:

$k_F + k_{IC} + k_{ISC} = 3.85 \times 10^8\ s^{-1}, k_Q = 1.26 \times 10^{10}\ M^{-1}s^{-1}$

8. $0.5 = \frac{k_Q[Q]}{k_F+k_{IC}+k_{ISC}+k_Q[Q]} = \frac{1.26 \times 10^{10} \times [Q]}{3.85 \times 10^8 + 1.26 \times 10^{10} \times [Q]} \rightarrow [Q] = 0.0306\ M$

9.



P

10. At high temperatures (the blue dashed line), the total decay process of the T_1^G state is dominated by T-T energy transfer to the polymer matrix (rate constant k_Q), while at low temperatures (the red dashed line) non-radiative decay to the singlet state (rate constant k_N) prevails.

at low temperature (red dashed line): $\ln(k_N + k_Q) \approx \ln k_N = \ln A_N - \frac{E_N}{R} \cdot \frac{1}{T}$.

at high temperature (blue dashed line): $\ln(k_N + k_Q) \approx \ln k_Q = \ln A_Q - \frac{E_Q}{R} \cdot \frac{1}{T}$.

We find the slope of the blue dashed line, which equals to $-\frac{E_Q}{R}$:

slope $= -\frac{E_Q}{R} = -4800 \rightarrow \Delta E = E_Q = 40$ kJ / mol.

Problem 12. Photoacids

1. $2.1 \times 10^{-3} \times \frac{5}{25} = 4.2 \times 10^{-4}$ mol / L

2. $A = \varepsilon l c \Rightarrow \varepsilon = \frac{A}{lc}$

$$\varepsilon_{\mathrm{ArOH}} = \frac{0.219}{1 \cdot 4.2 \cdot 10^{-4}} = 521 \text{ cm}^{-1}\text{ M}^{-1}$$

$$\varepsilon_{\mathrm{ArO^-}} = \frac{0.44}{1 \cdot 4.2 \cdot 10^{-4}} = 1048 \text{ cm}^{-1}\text{ M}^{-1}$$

3. Let's determine the pH of solutions **3-5**:

$$K_b = \frac{[\mathrm{NH_4^+}][\mathrm{OH^-}]}{[\mathrm{NH_3}]} \Rightarrow [\mathrm{H^+}] = \frac{K_w}{K_b} \times \frac{[\mathrm{NH_3}]}{[\mathrm{NH_4^+}]}$$

$$\mathrm{pH} = 14 - \mathrm{p}K_b + \log\frac{[\mathrm{NH_3}]}{[\mathrm{NH_4^+}]}$$

$$\mathrm{pH_3} = 14 - 4.75 + \log\frac{1}{3} = 8.77$$

$$\mathrm{pH_4} = 14 - 4.75 + \log\frac{1}{2} = 8.95$$

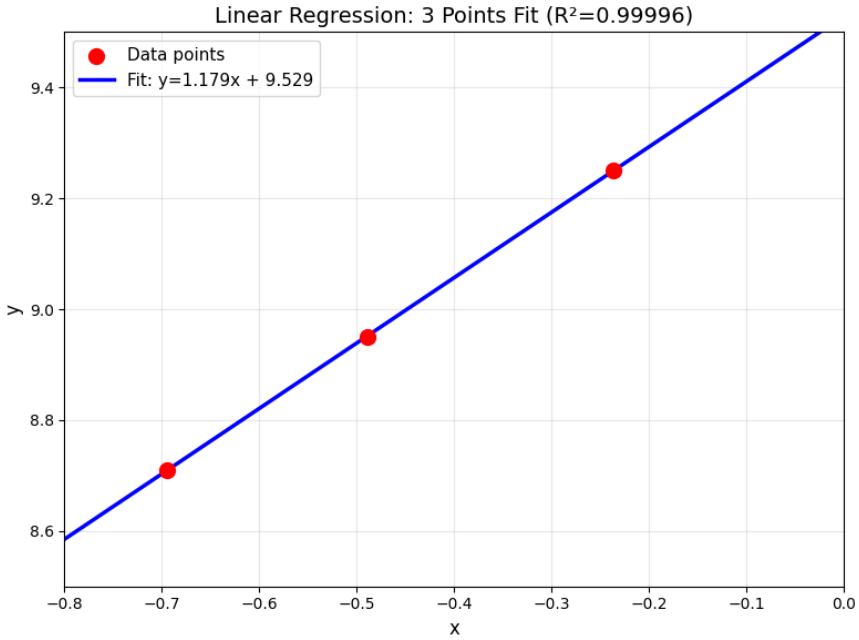
$$\mathrm{pH_5} = 14 - 4.75 = 9.25$$

The $\log[\mathrm{ArO^-}]/[\mathrm{ArOH}]$ values for solutions **3-5** can be found from three systems of linear equations:

$$\begin{cases} \varepsilon_{\mathrm{ArOH}} c_{n,\mathrm{ArOH}} + \varepsilon_{\mathrm{ArO^-}} c_{n,\mathrm{ArO^-}} = A_n \\ c_{n,\mathrm{ArOH}} + c_{n,\mathrm{ArO^-}} = c_{\mathrm{total}} \end{cases}$$

Then:

Solution	$[\mathrm{ArOH}]$	$[\mathrm{ArO^-}]$	$[\mathrm{ArO^-}]/[\mathrm{ArOH}]$	$\log\frac{[\mathrm{ArO^-}]}{[\mathrm{ArOH}]}$
3	3.49×10^{-4}	7.05×10^{-5}	0.202	-0.695
4	3.20×10^{-4}	1.00×10^{-4}	0.324	-0.489
5	2.66×10^{-4}	1.54×10^{-4}	0.579	-0.237



From the intercept we obtain the $pK_a = 9.53$ (9.50 in literature)

4. $\Delta G° = \Delta H - T\Delta S = -RT\ \ln K_a$

$\Delta G°^* = \Delta H^* - T\Delta S = -RT\ \ln K_a^*$

$N_A h\upsilon_1 + \Delta H^* = N_A h\upsilon_2 + \Delta H$

$-RT\ \ln K_a - (-RT\ \ln K_a^*) = \Delta H - \Delta H^*$

$RT(-\ln K_a + \ln K_a^*) = N_A h\upsilon_1 - N_A h\upsilon_2$

$\ln x = 2.303 \log x$

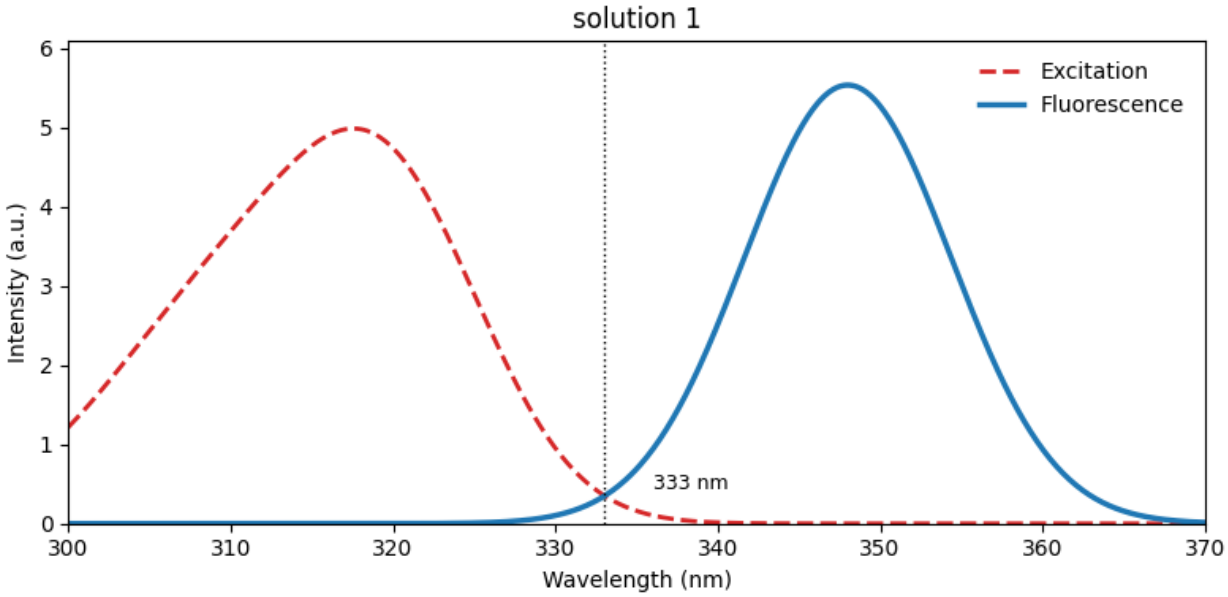
$2.303RT(pK_a - pK_a^*) = N_A h(\upsilon_1 - \upsilon_2)$

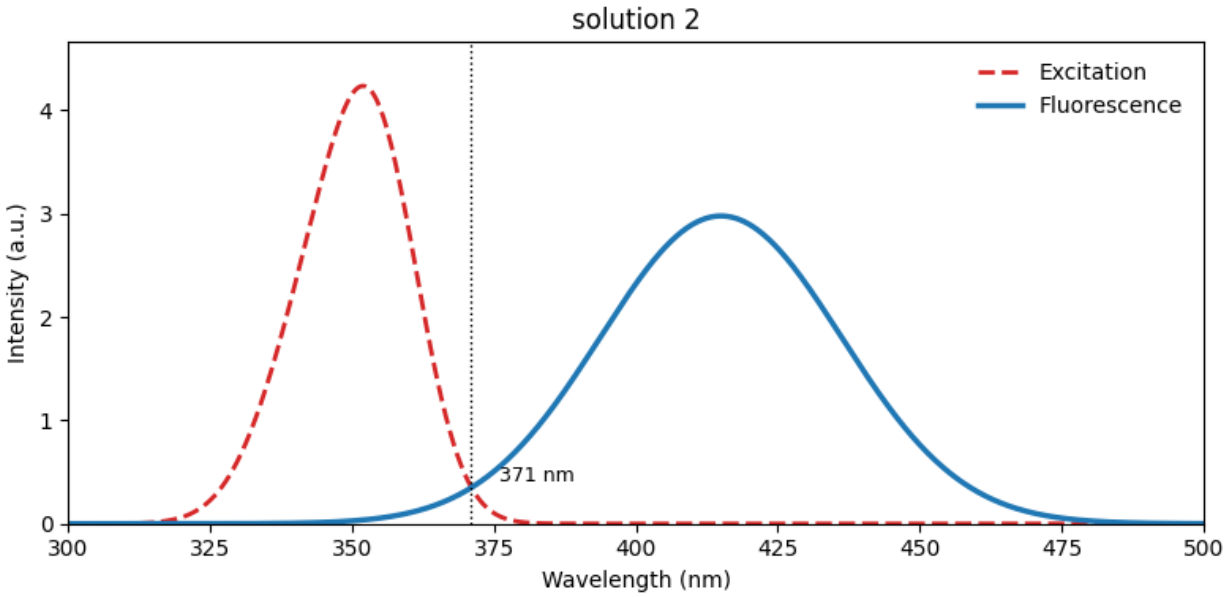
$pK_a - pK_a^* = \Delta pK_a = \frac{N_A h}{2.303RT}(\upsilon_1 - \upsilon_2)$

And finally, converting frequencies to wavelengths:

$$\Delta pK_a = \frac{N_A hc}{2.303RT}\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right)$$

5. Based on the intersection of the two spectra we can find the minimal energy of excitation:





$$\Delta \mathrm{p}K_a = \frac{N_\mathrm{A} hc}{2.303RT}\left(\frac{1}{333\times10^{-9}} - \frac{1}{371\times10^{-9}}\right)$$

$$\Delta \mathrm{p}K_a = \mathrm{p}K_a - \mathrm{p}K_a^* = 6.45$$

$$\mathrm{p}K_a^* = 9.53 - 6.45 = 3.08$$

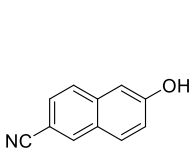
At pH 5:

$$5 = 3.08 + \log\frac{[\mathrm{ArO^{-*}}]}{[\mathrm{ArOH^{*}}]}$$

$$\frac{[\mathrm{ArOH^{*}}]}{[\mathrm{ArO^{-*}}]} = 0.012$$

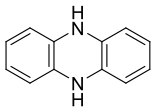
Note: we can't use fluorescence maxima wavelengths $\lambda_{max1}, \lambda_{max2}$ for the calculations of pK_a^*. The necessity of using intersection of absorption and fluorescence spectra is born because of vibrational relaxation (and this is why the wavelength shift occurs – this is also called the Stokes shift), that's why the intersection of the two spectra shows the minimum <u>electronic energy</u> difference needed for a molecule to reach its excited state.

6. Based on the rule, we find out that **a, c, and d** are Baird-antiaromatic, and thus acidic forms of these molecules will try to relieve antiaromaticity and show photoacidic nature.



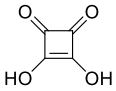
(a)

Huckel-aromatic
Baird-antiaromatic



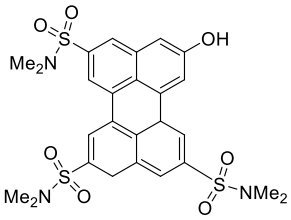
(b)

Huckel-antiaromatic
Baird-aromatic



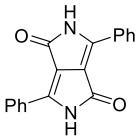
(c)

Huckel-aromatic
Baird-antiaromatic



(d)

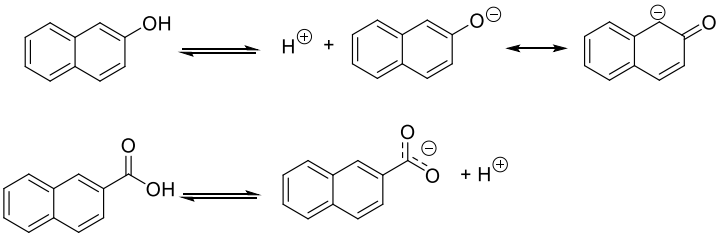
Huckel-aromatic
Baird-antiaromatic



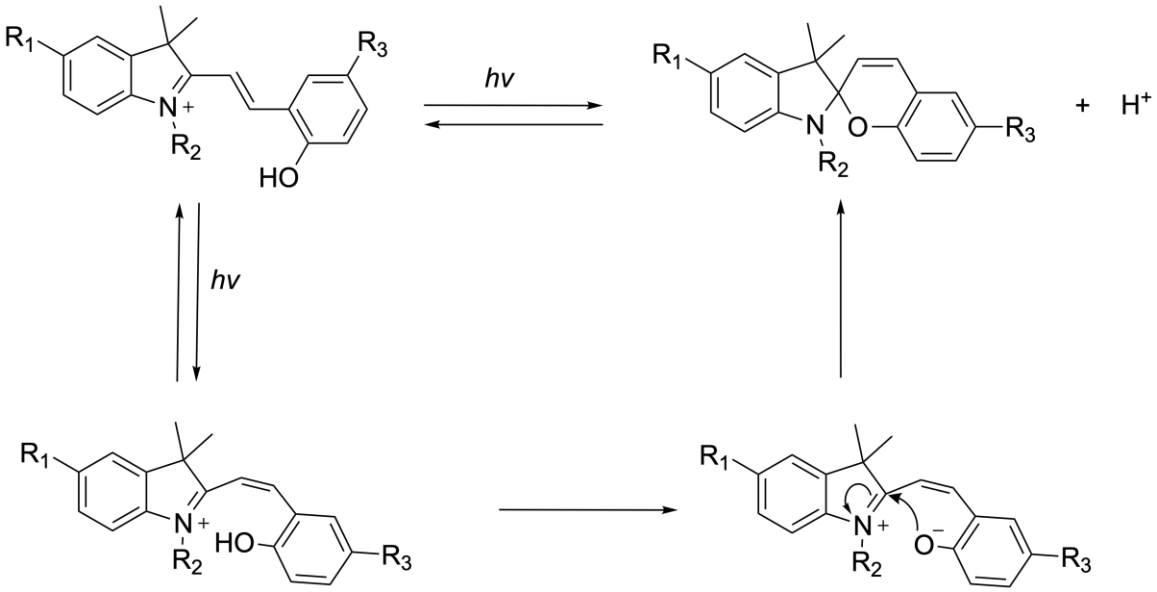
(e)

Huckel-antiaromatic
Baird-aromatic

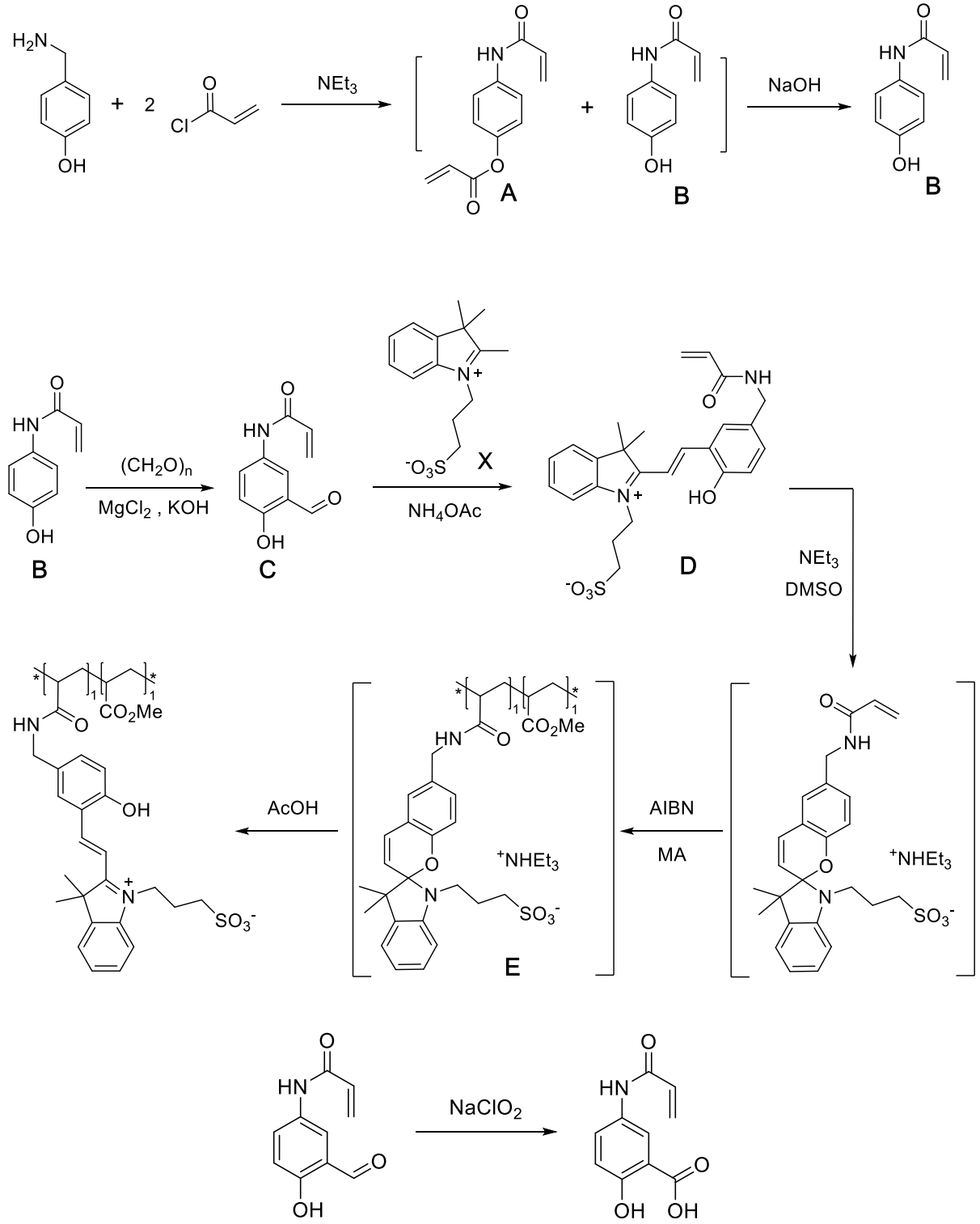
7. This is the result of the small charge redistribution on the ring. In the case of 2-naphthol, charge is partially redistributed to the aromatic core. But in the case of 2-naphthoate, the charge is mostly on the two oxygen atoms and that's why it doesn't help to lessen the antiaromaticity of naphthalene ring in the excited state.



8.

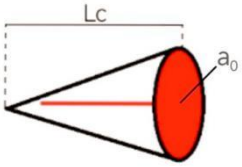
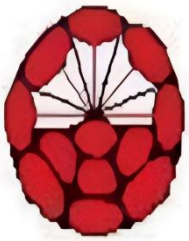
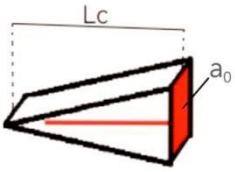
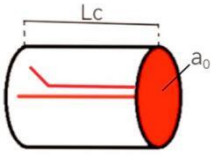
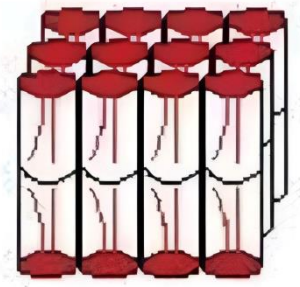


9.

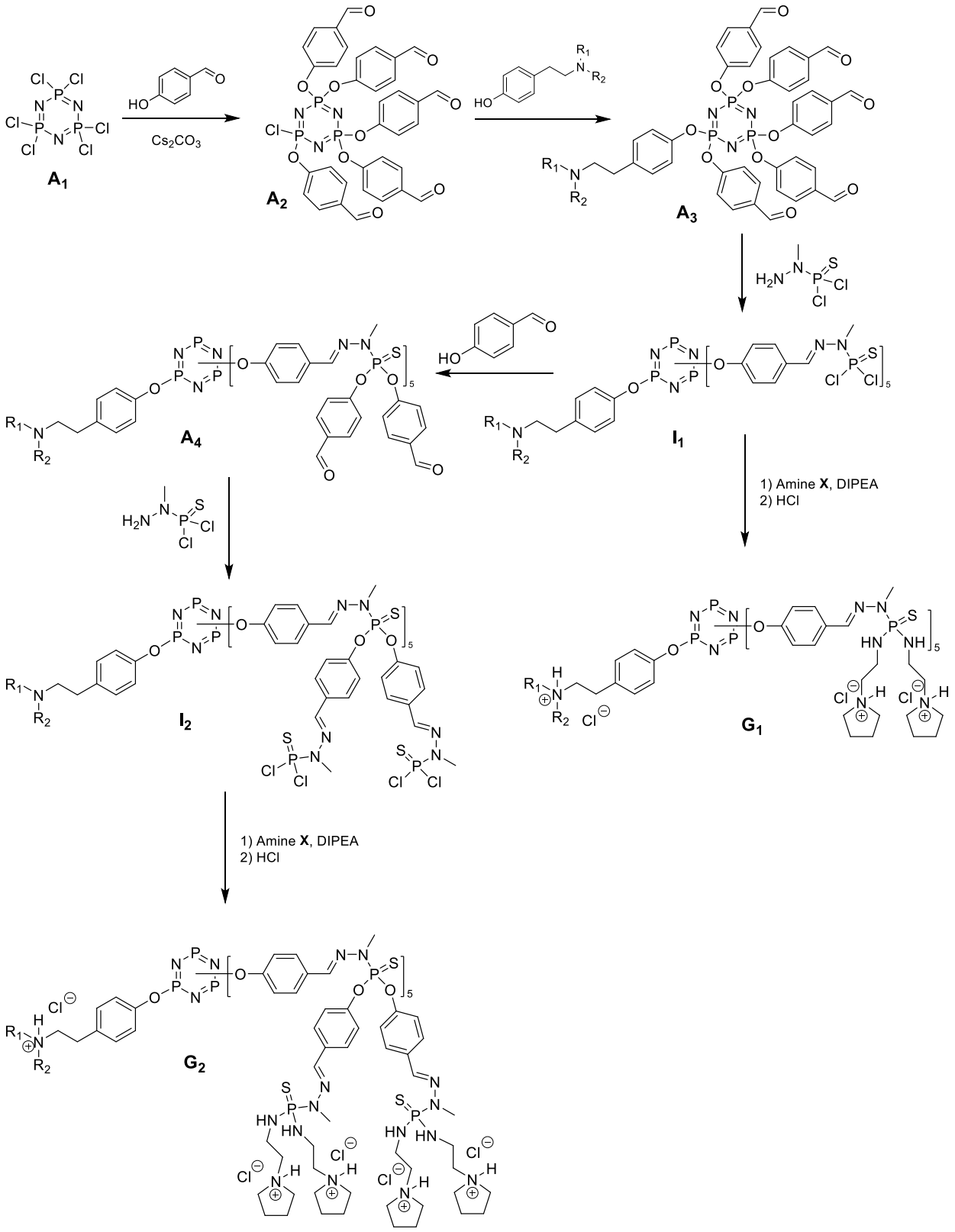

NEt3
NaOH
A
B
(CH2O)n
MgCl2 , KOH
NH4OAc
X
C
D
NEt3
DMSO
AIBN
MA
AcOH
E
NaClO2

Problem 13. Interested in micelles?

1.

For cone-shaped molecules: 	$\frac{V_C}{L_C} = \frac{\frac{1}{3}\pi r^2 L_C}{L_C} = \frac{1}{3}a_0$ This will form a spherical micelle.	
For rectangular wedge-shaped surfactants: 	$\frac{V_C}{L_C} = \frac{\frac{1}{2}a_0 L_C}{L_C} = \frac{1}{2}a_0$ This will form a cylindrical micelle.	
For cylinder-shaped molecules: 	$\frac{V_C}{L_C} = \frac{\pi r^2 L_C}{L_C} = a_0$ This will form a bilayer structure.	

2. We can find the ratio of atoms in **A_1** – 26.73/31:12.09/14:61.18/35.45 = 1:1:2. So, the empirical formula of **A_1** is $PNCl_2$. The structure $N{\equiv}PCl_2$ does not have a threefold symmetry axis, but the cyclic trimer of it demonstrates possesses such a symmetry axis. So, **A_1** is $P_3N_3Cl_6$.


A1
Cs2CO3
A2
A3
A4
I1
I2
G1
G2
1) Amine X, DIPEA
2) HCl
1) Amine X, DIPEA
2) HCl

3. In the final step, hydrochloric acid protonates the external amino group to form an ammonium salt. The amino groups inside are too densely packed to be protonated . The number of cationic groups we can find: $5\times2^{10}+1=5121$.

4. As can be seen from the structures of **G_1** and **G_2**, they are likely to have cone shapes and assemble into spherical micelles.

5. The correct answer is: [c] to create a water-soluble, positively charged form that will show biological activity.

HCl protonates the amine groups, creating cationic dendrons (*e.g.*, P1G1·HCl). This is essential for water solubility (enabling micelle formation) and provides the positive surface charge needed for cellular uptake and anticancer activity.

6. The form of micelle is dependent on the generation of the head groups. Since **Gen_3** has the largest hydrophilic surface area and smallest N_s, it forms a spherical micelle, whereas **Gen_2** and **Gen_1** with a moderate and smaller hydrophilic surface accumulate in cylindrical and bilayer structures respectively.

7. The initial concentration of monomers can be described as:

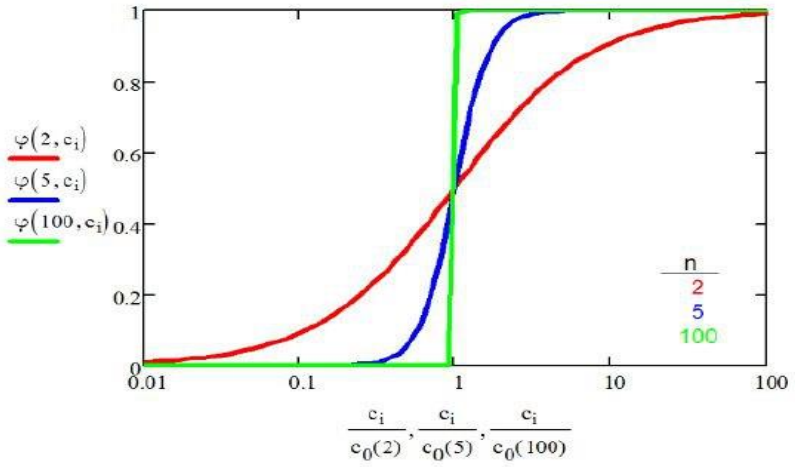
$$C_i = [\mathrm{M}] + n \times [\mathrm{M}_n]$$

So, the fraction of monomers present in micelles:

$$\varphi_{n\times[\mathrm{M}_n]} = \frac{n \times [\mathrm{M}_n]}{C_i} = \frac{K_n[\mathrm{M}]^n \times n}{[\mathrm{M}] + K_n[\mathrm{M}]^n \times n} = \frac{K_n[\mathrm{M}]^{n-1} \times n}{1 + K_n[\mathrm{M}]^{n-1} \times n}$$

8. According to the equation derived in question 7, the maximum number monomers will be present inside the micelles when the concentration of monomers is high and the aggregation number is low, as the concentration of the monomers is lower than 1 mol/dm^3. So the correct answer is [c].

9. The function in the solution of question 7 has an inflection point, c_0, at the CMC, which steepens with increasing aggregation number:



At the point where $\varphi_{n\times[M_n]} = 0.5$ CMC is reached so:

$$1 = K_n \times [\mathrm{M}]^{n-1} \times n \rightarrow \mathrm{CMC} = [\mathrm{M}] = (K_n \times n)^{\frac{1}{1-n}}$$

10. As there are aggregates formed, the specific conductivity of the solution falls. This feature corresponds to the moment that monomer concentration is equal to the CMC (slope of line reduces). Using the CMC's dependence on the aggregation number, we can derive the equation for calculating the equilibrium constant and for the free energy of micelle formation:

$$K_n = \frac{1}{n} \times \mathrm{CMC}^{1-n} = e^{\frac{-\Delta_r G^\circ_{\mathrm{micelle}}}{RT}} \rightarrow \Delta_r G^\circ_{\mathrm{micelle}} = RT \times \ln(n \times \mathrm{CMC}^{n-1})$$

Here it can be seen that if the aggregation number and free energy remain constant, the temperature will increase with CMC which is under 1 mol/dm^3. So the comparison is as follows:

$$T_1 < T_2$$

Problem 14. The basics of Metal–Organic Frameworks

1. From the figure and model, it is clear that MOF-5 has a simple cubic (SC) unit cell (if we do not take into account linker geometry). The Zn_4O cluster forms corners and the linker forms edges. SC has 8 corners and 12 edges. Knowing that $Z = 1$ for SC we can calculate the empirical formula:

$$Zn_4O + ¼*12*C_8H_4O_4 = C_{24}H_{12}O_{13}Zn_4$$

2. $4Zn(NO_3)_2 + 3H_2C_8H_4O_4 + 8Et_3N + H_2O \rightarrow Zn_4O(C_8H_4O_4)_3 + 8Et_3N{\cdot}HNO_3$

3. Elemental analysis of crystals before activation showed nitrogen content N 7.64%, which originates from DMF molecules. In the mass spectra, the signal at 73 *m/z* indicates its presence. The 112 and 114 signals with relative intensity 3:1 are characteristic of a chlorine containing molecule C_6H_5Cl. The 77 *m/z* signal indicates $C_6H_5^+$ ions. From the nitrogen content and knowing the empirical formula of activated MOF-5 we can calculate the unactivated material formula: $C_{24}H_{12}O_{13}Zn_4 \cdot xC_3H_7NO \cdot yC_6H_5Cl$

We have two proportions:

$14.01x$ ---------- 7.64	$14.01x$ ---------- 7.64
$1.008{\cdot}(12+7x+5y)$ ------ 5.02	$12.01{\cdot}(24+3x+6y)$ ------ 44.2

Solving this system of equations we get $x = 8$ and $y = 1$.

Empirical formula $C_{24}H_{12}O_{13}Zn_4 \cdot 8C_3H_7NO \cdot C_6H_5Cl$

4. Volume of unit cell $V = a^3 = (12.94\ Å)^3 = 2166.72\ Å^3$

Volume of the atom $v = 4/3\pi r^3$

$$v = 4/3\pi\ (1.5^3{\times}4 + 13{\times}1.5^3 + 24{\times}1.7^3 + 12{\times}1.2^3) = 820.68\ Å^3$$

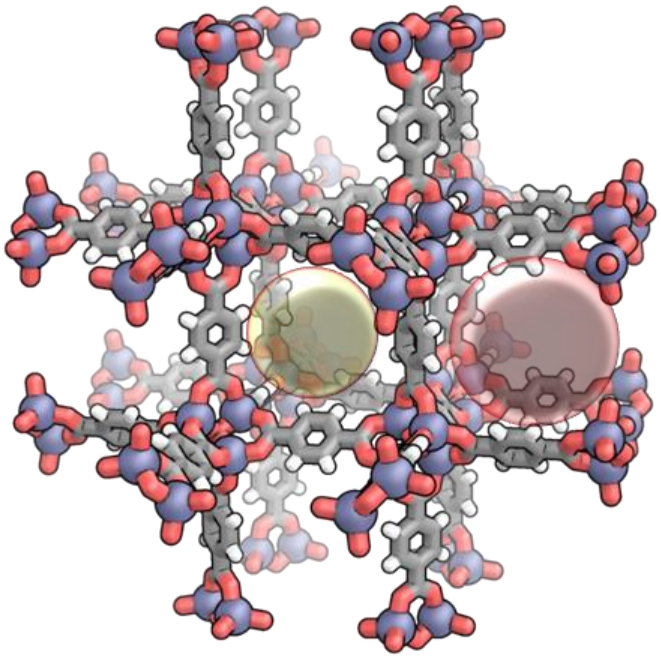
Free volume equals to $2166.72\ Å^3 - 820.68\ Å^3 = 1346.04\ Å^3$ or 62.1%.

The mass of one unit cell $m = \frac{M}{N_A} = \frac{769.856\ \text{g/mol}}{6.02{\times}10^{23}\ \text{mol}^{-1}} = 1.28 \times 10^{-21}$ g

Experimental surface area of one unit cell:

$S_{\text{cell}} = S \times m = 2900\ \text{m}^2/\text{g} \times 1.28 \times 10^{-21}\ \text{g} = 3.7 \times 10^{-18}\ \text{m}^2$ or $370.86\ Å^2$

5. As can be seen from the constructed model, the BDC units are placed at 45^o to the cube faces. This along with not matching orientation of Zn_4O clusters in a row differentiates two neighbour cubes with different pore sizes. Therefore, we should increase the unit cell to eight cubes, doubling unit cell parameter to 25.88 Å.



6. First, we should find the volume occupied by 10 g of material. This can be done by finding the density: $\rho = \frac{m}{V} = \frac{1.28\times10^{-21}\ \text{g}}{2166.72\times10^{-24}\ \text{cm}^3} = 0.59\ \text{g/cm}^3$, then 10.0 g would occupy 10/0.59 = 16.95 cm^3. The mass of hydrogen can be found through the ideal gas law

$$m = \frac{M\times P\times V}{R\times T} = \frac{2\ \frac{\text{g}}{\text{mol}}\times 20\times10^5\ \text{Pa}\times 16.95\times10^{-6}\ \text{m}^3}{8.314\ \frac{\text{m}^3\times\text{Pa}}{\text{K}\times\text{mol}}\times 298\ \text{K}} = 0.027\ \text{g}.$$

MOF-5 uptakes 0.1 g hydrogen. Excess uptake equals to $\frac{0.1-0.027}{0.027} * 100 = 270\%$

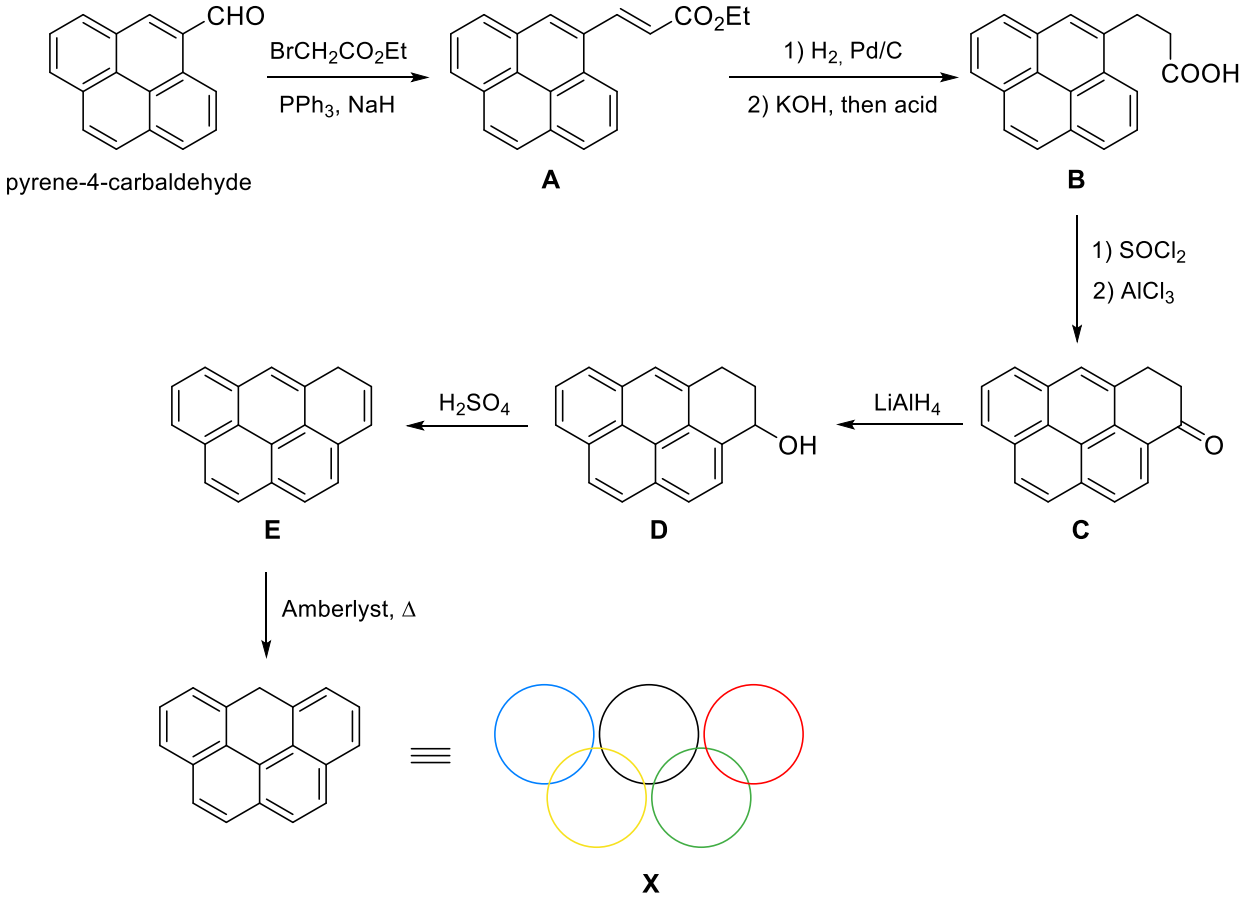
7. We know the distance *a* = 12.94 Å. Now we should find the length of the linker:

$$\frac{V_1}{V_0} = \left(\frac{12.94-2.69+x}{12.94}\right)^3 = 0.693; \quad \frac{10.25+x}{12.94} = \sqrt[3]{0.693}; \qquad x = 1.20\ \text{Å}.$$

The linker is acetylenedicarboxylic acid *(students are not expected to memorise bond lengths in exam).*

Problem 15. Microscopy in chemistry

1.



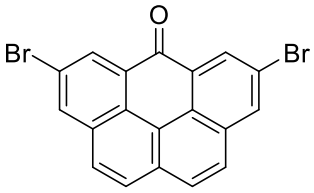
2. Because of its similarity to the Olympic flag, it was called *Olympicene*.

3. Let's find F first. It's similar to olympicene X and contains 2 bromine atoms. Moreover, we are given it's $m/z = 411.89 \approx 412$ in the mass spectrum. Considering that 2 hydrogen atoms of the original olympicene are substituted with bromine and subtracting M(X) and 2M(Br) we get:

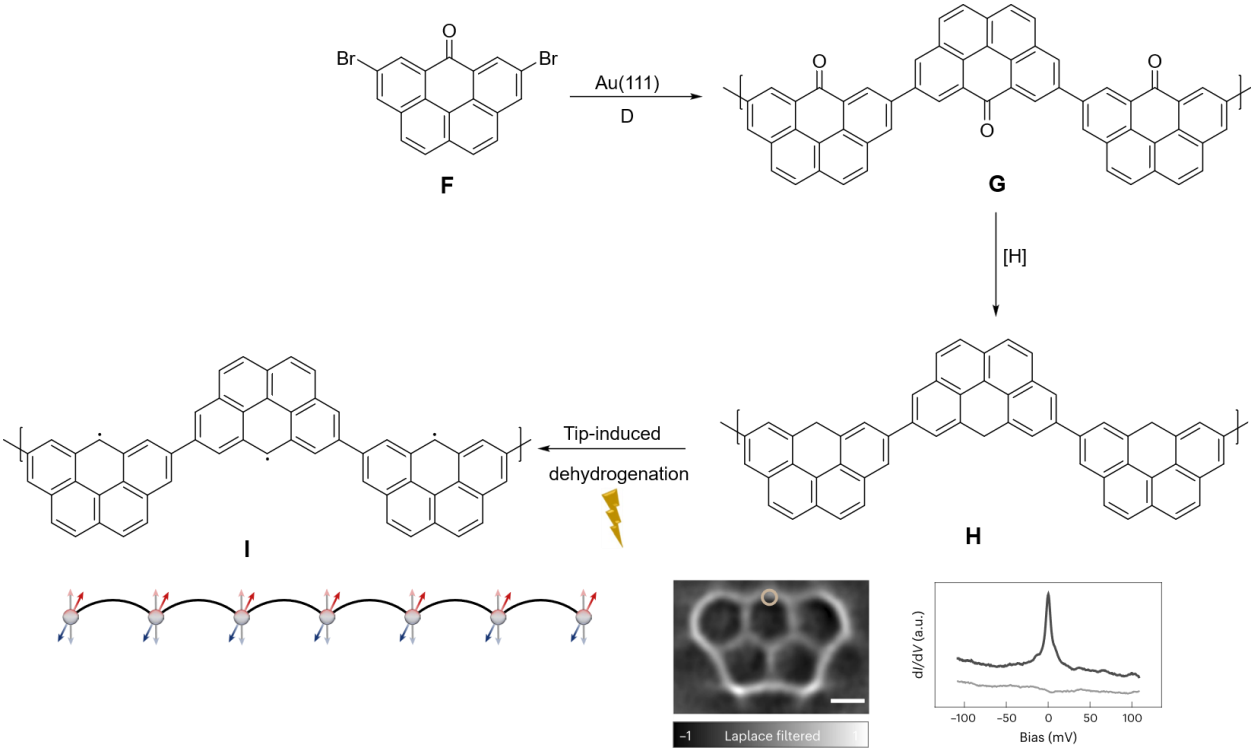
$$\Delta M = 412 - M(\mathrm{X}) - 2M(\mathrm{Br}) + 2M(\mathrm{H}) = 412 - 240 - 160 + 2 = 14\ \mathrm{g/mol}$$

As we know that the central carbon atom of olympicene is fully substituted (because there are only 4 types of protons according to problem), we thus can add 2 missing H to get:

$$\Delta M = M(\square) = 14 + 2 = 16\ \mathrm{g/mol} \Rightarrow \mathrm{C{=}O\ group}$$



Then proceed with the scheme:

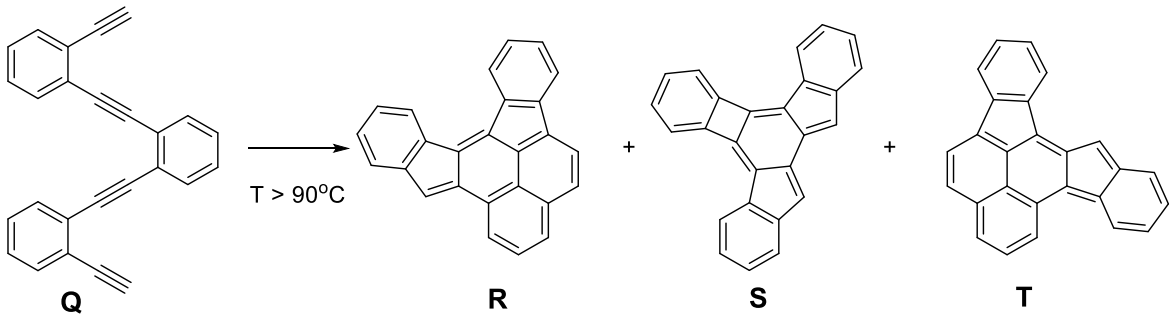


4.

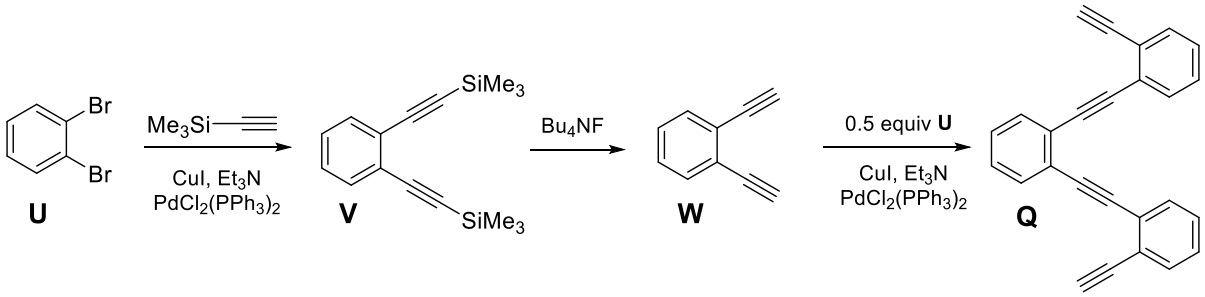
$$38 \cdot 10^{-3} \cdot \frac{2\pi^2}{L} \cdot 1.602 \cdot 10^{-19} = 1.38 \cdot 10^{-23} \cdot 298$$

So, the minimal $L = 29.22 \approx 30$

5.

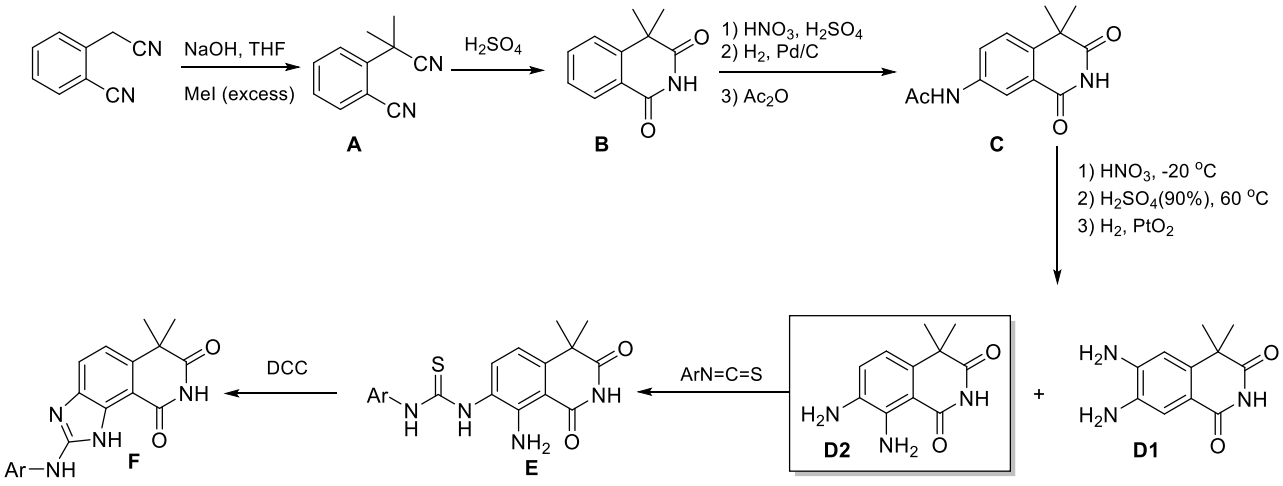


6.

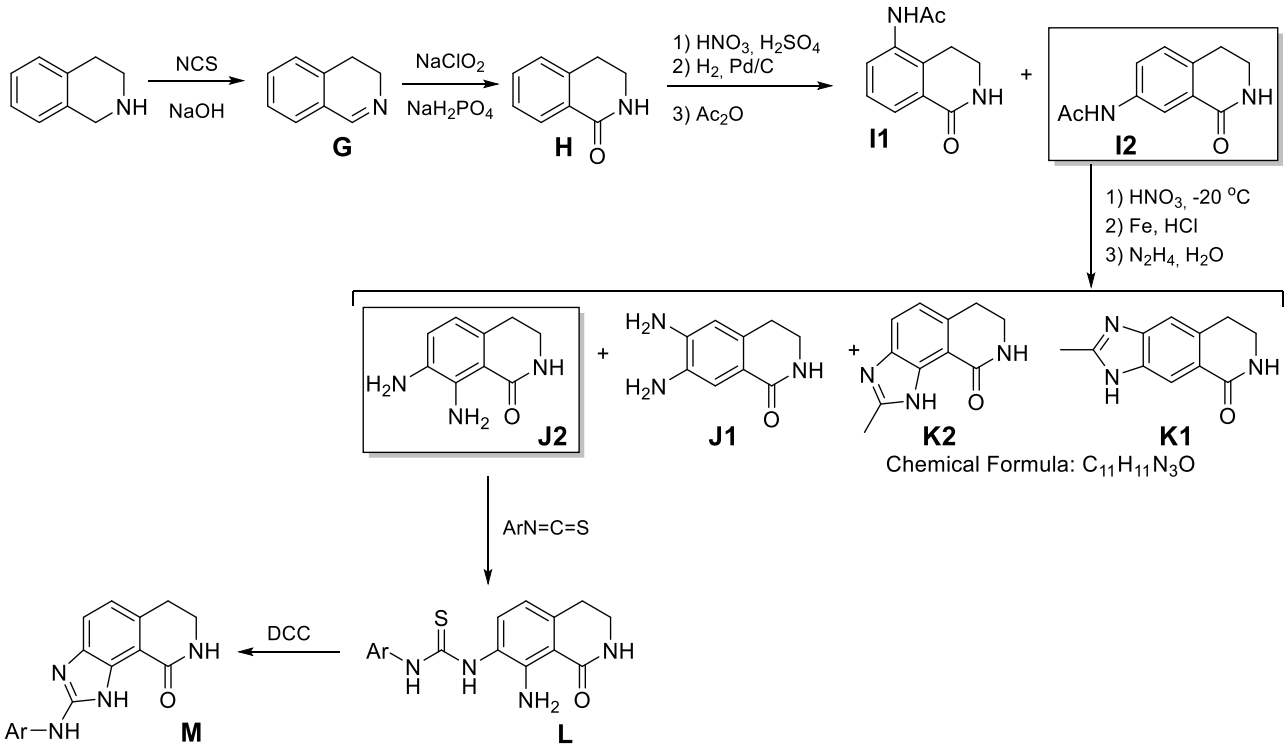

Br
Br
U
Me3Si
CuI, Et3N
PdCl2(PPh3)2
SiMe3
SiMe3
V
Bu4NF
W
0.5 equiv U
CuI, Et3N
PdCl2(PPh3)2
Q

Problem 16. Kinase inhibitors

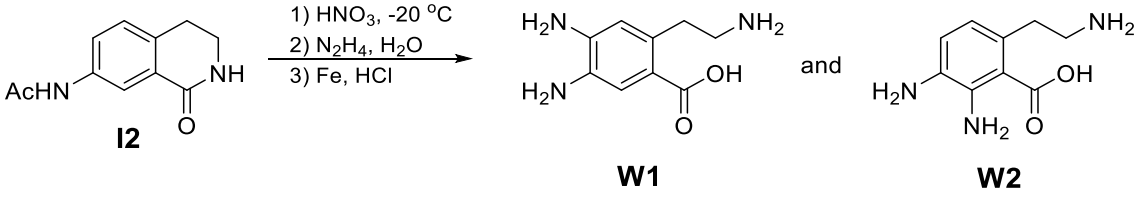
1. The formation of the two isomers of compound **D** occurs because of non-selective nitration of **C**.



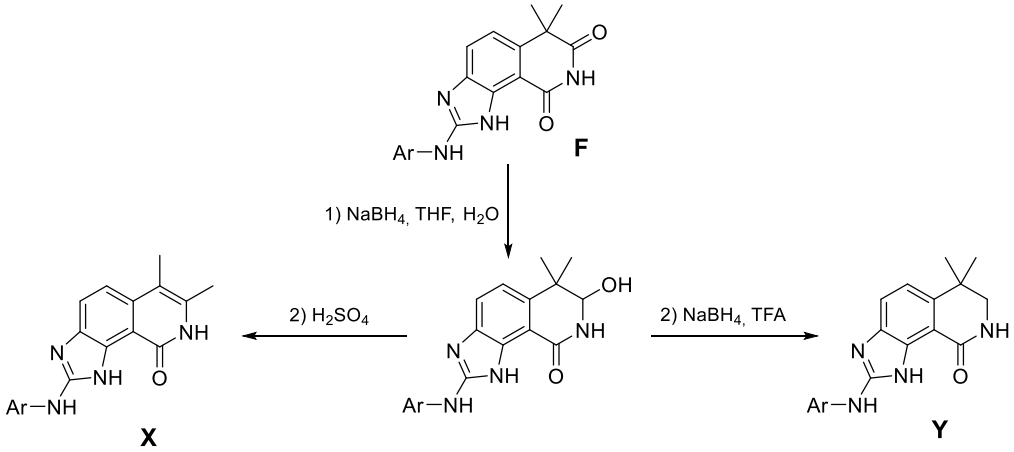
2. During the hydrolysis of the amide bond in diamine **I2** the acyl group tends to form a conjugated benzimidazole ring.



3. The *ortho*-position to the alkyl substituent in **B** is sterically hindered.
4. The order was changed due to hydrolysis of the lactam. Reduction of the nitro group to the amino group increases the electron donating effect of the aromatic ring and reduces the electrophilicity of lactam. If the order is left unchanged, two isomeric structures can be proposed for **W**:



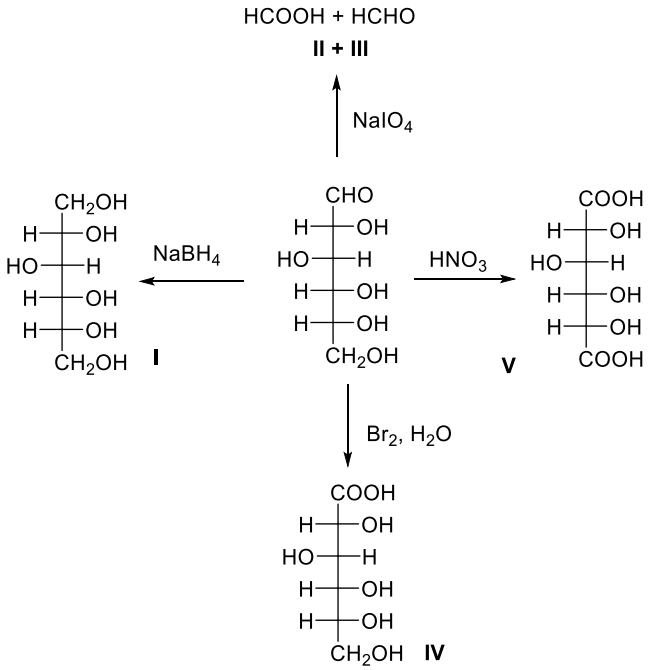
5. The reduction of the imide with $NaBH_4$ as the first step leads to the hemi-aminal formation. Since **Y** is a homologous compound to **M**, subsequent elimination and then reduction leads to **Y**. If only acid is used as the second step, the 1,2-alkyl shift product **X** is formed via tertiary carbocation formation.



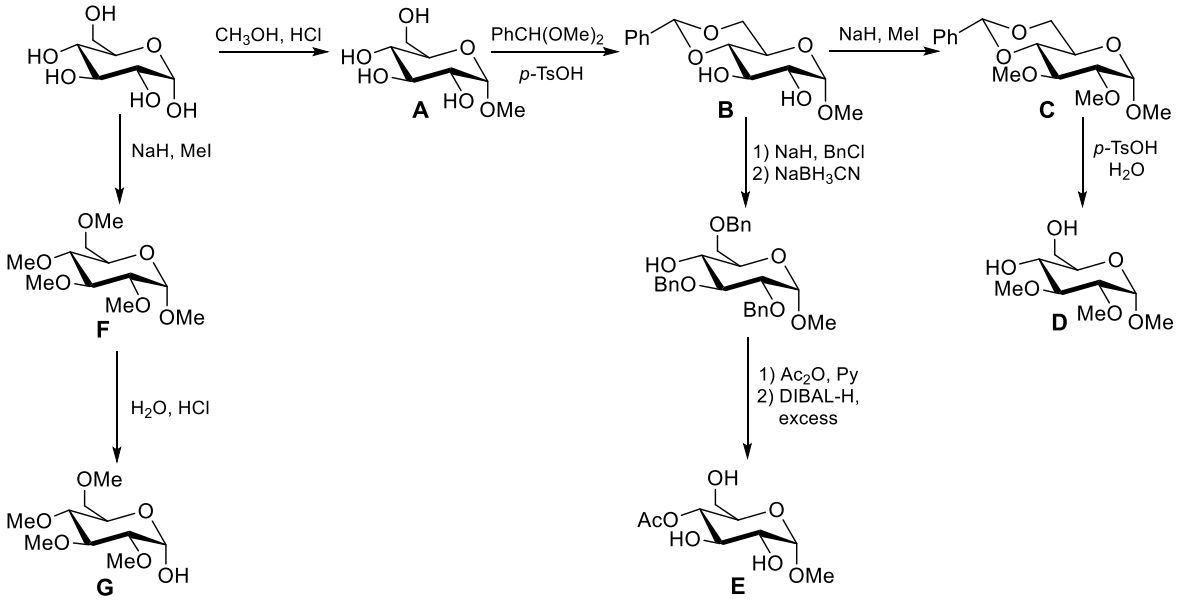
6. The phenyl-substituent has a coplanar orientation with the core structure. The 2,6-dichlorophenyl-substituent, due to steric repulsion of the *ortho*-substituents, limits any rotation, which enhances packing in hydrophobic pocket of the enzyme binding site. If one of the substituents is replaced by smaller (H) or bigger group (Et), coordination becomes weaker.

Problem 17. Carbohydrate chemistry

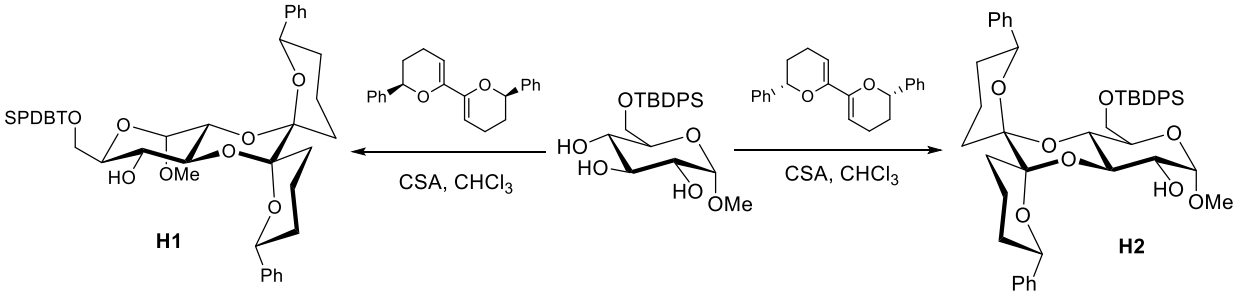
1.



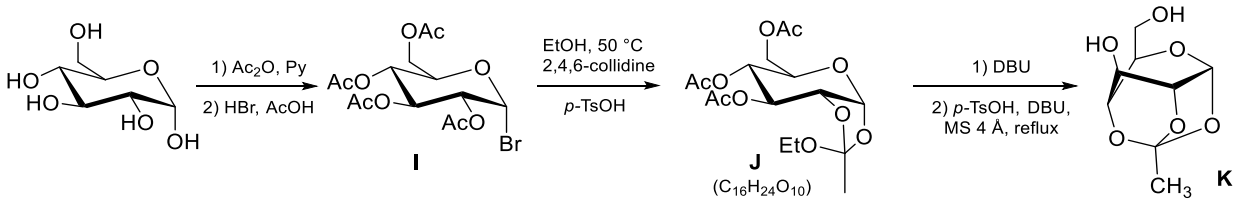
2. Acetal formation and hydrolysis occur under acidic conditions due to the reactivity of the anomeric positions (**A** and **G**). However, in the presence of TsOH, hydrolysis of the anomeric group can be avoided. According to the hint, **D** has fewer methylene groups than **G**. The formation of a second six-membered ring clearly indicates the position of acetal formation. The acetal group then is reduced to a benzyl group. DIBAL-H removes the benzylic protecting group.



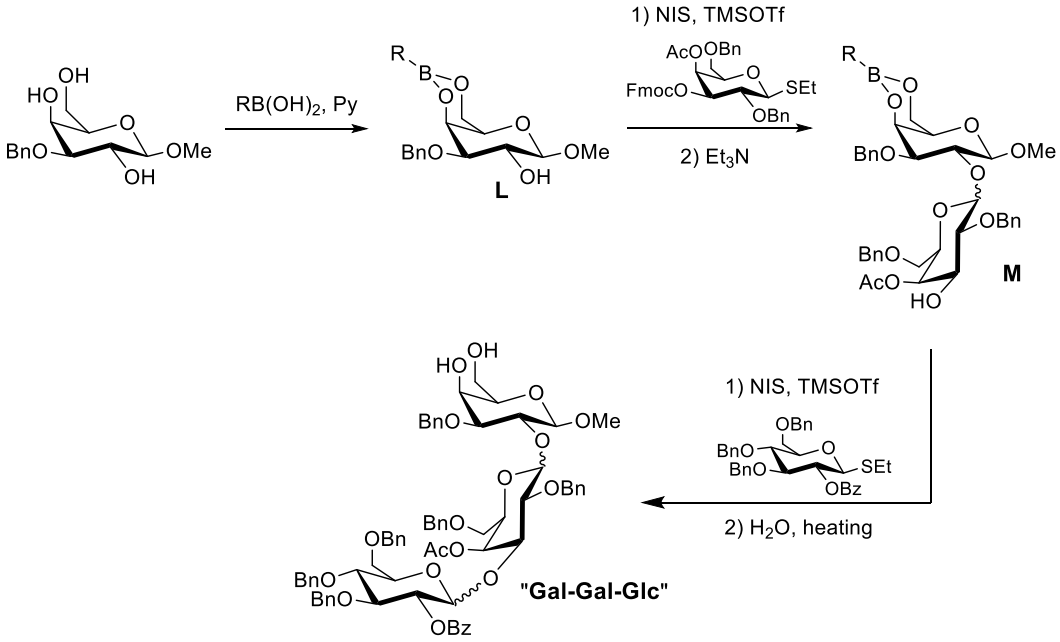
3. The phenyl groups in the THP rings sit in equatorial positions in chair conformations. This defines which of the two OH groups gets protected. For the correct stereochemical result in **H1**, the glucose derivative should be redrawn:



4. In the first step, all hydroxy groups are transformed to esters. Then the anomeric ester group is substituted by bromide. According to the structure of **K**, we can predict a bicyclic structure of **J**. **J** is further deprotected in basic conditions which do not affect the orthoester. Transorthoesterification in weak acidic conditions yields **K**.



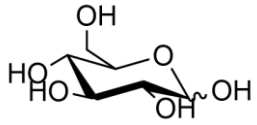
5. The reaction of a galactose derivative with the boronic acid derivative yields boronic ester **L**. NIS/TMSOTf-promoted glycosylation and following cleavage of the Fmoc group in basic conditions leads to formation of disaccharide **M** (which group leaves can be deduced from the chemical formula). Another glycosylation and hydrolysis of the boronic ester in hot water yields the Gal-Gal-Glc derivative.



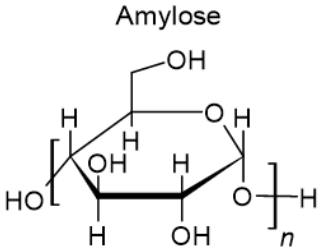
Problem 18. Inositol

1.

a)



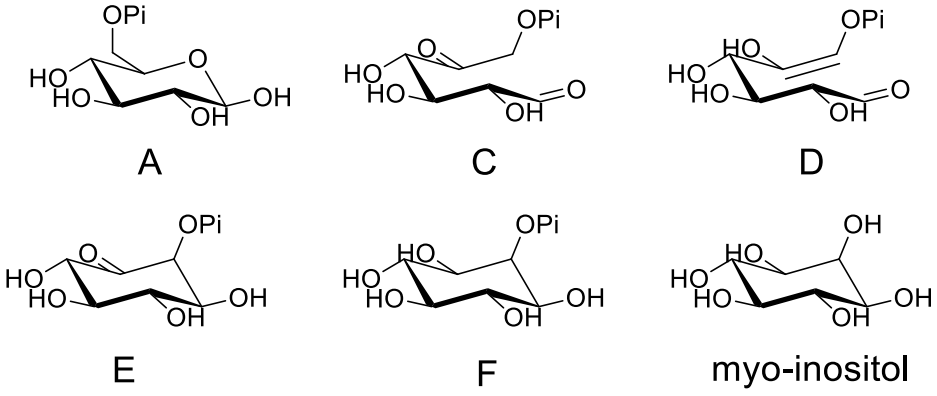
b)



2. $243\,000 = 162 \times n$;

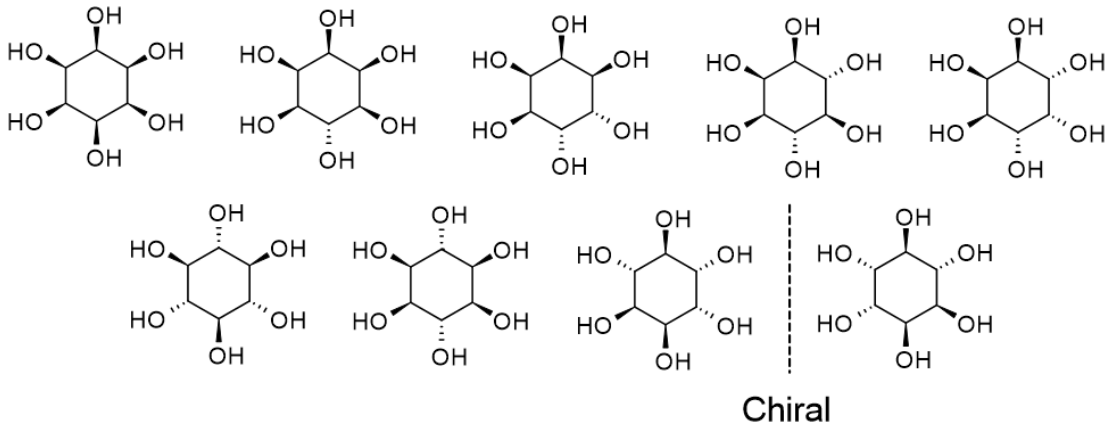
$\boldsymbol{n = 1500}$

3.

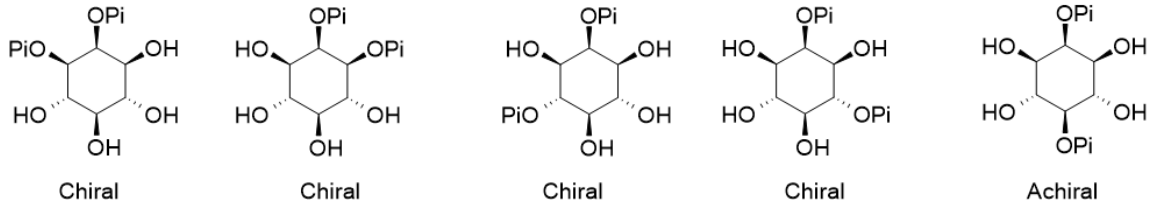


Another geometric isomer of **D** is also correct.

4.

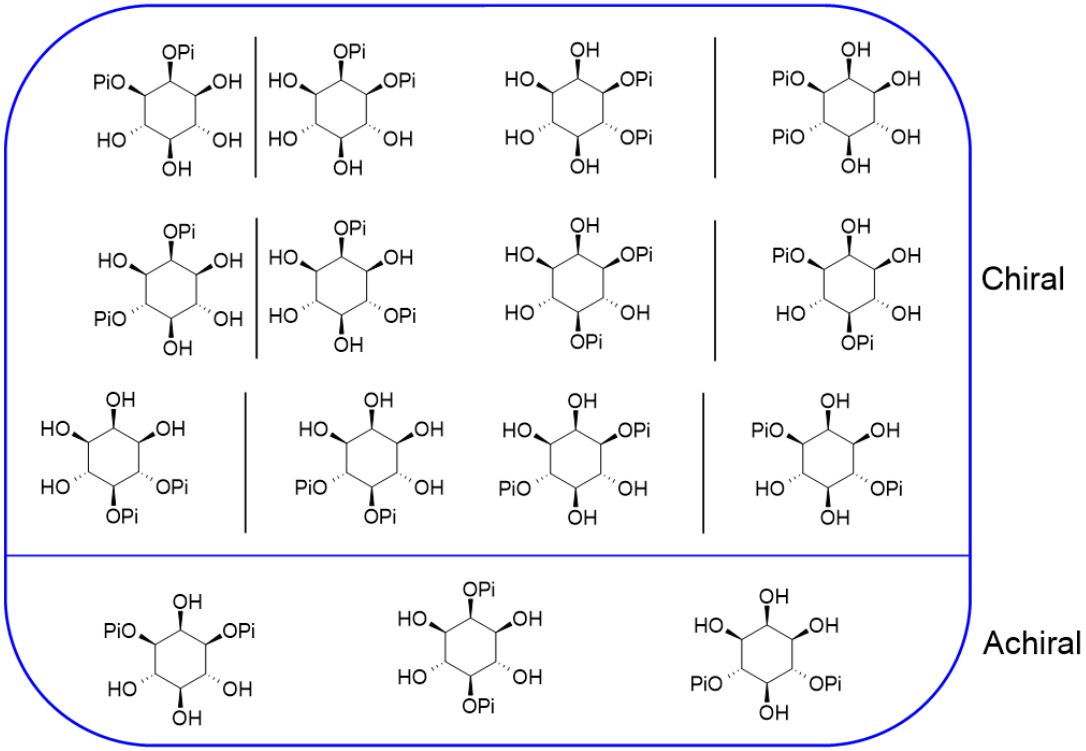


5.



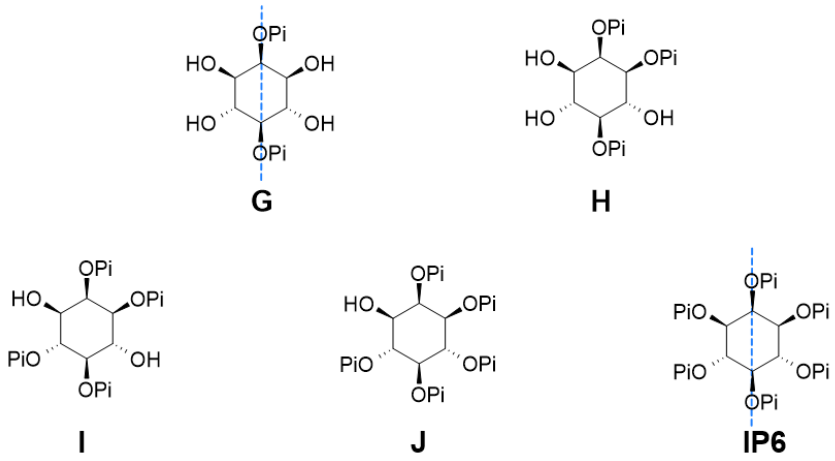
5 isomers of **G** can be formed in this pathway.

6.



15 isomers

7.

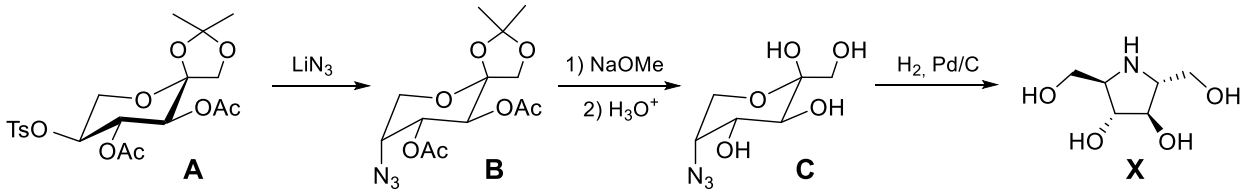


Enantiomers of **H**, **I** and **J** are also correct**.**

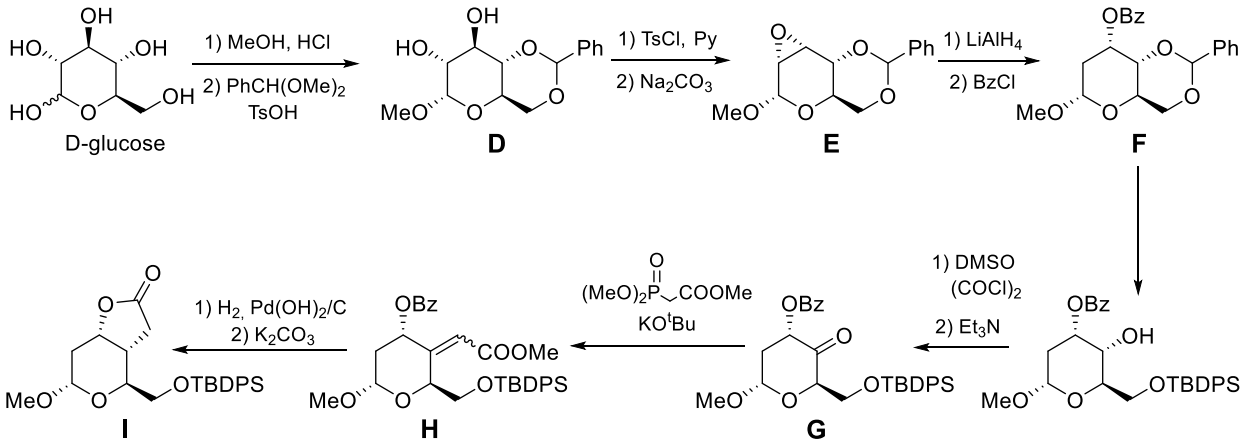
G and **IP6** contain a plane of symmetry.

Problem 19. Carbohydrates as chiral pools

1. When azide acts on **A**, a nucleophilic substitution reaction occurs with an inversion of configuration, and **B** is formed. When the protecting groups are removed, **C** is obtained, and then upon reduction, **X** is obtained.

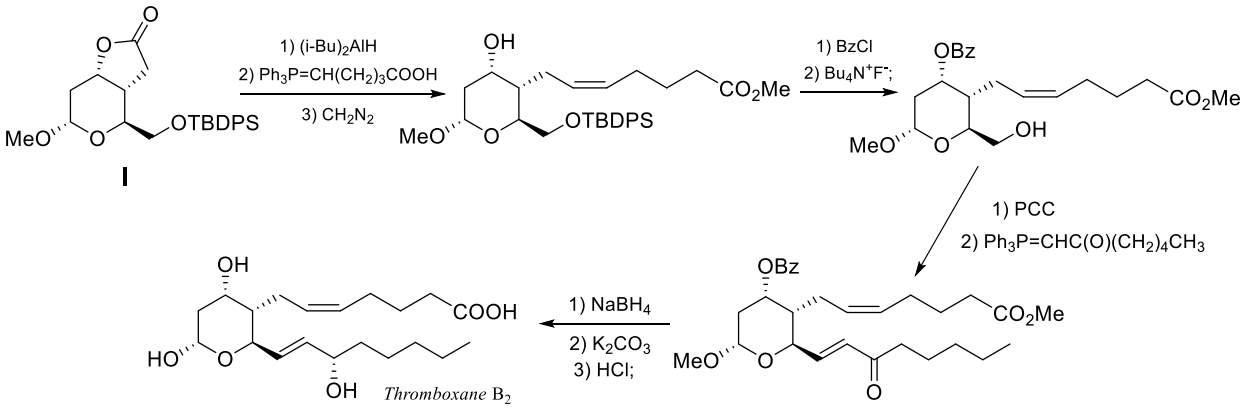


2. Reactions: MeOH, HCl with D-glucose leads to the formation of methyl glucoside, which then yields benzylidene acetal **D**. Reaction of **D** with TsCl: OH group (3-OH) reacts with TsCl and then in basic conditions, the other OH group (2-OH) attacks the tosylate and forms the epoxide **E**. Next, **E** is reduced with $LiAlH_4$, forming an OH, which upon benzoylation leads to **F**. The reaction of **F** with H_2, $Pd(OH)_2/C$ leads to hydrogenolysis of the benzylidene group, which is followed by protection of the OH with TBDPSCl. Further oxidation in the Swern reaction yields **G**, which through the Wittig reaction yields **H**. After reduction and hydrolysis, **H** yields **I**.

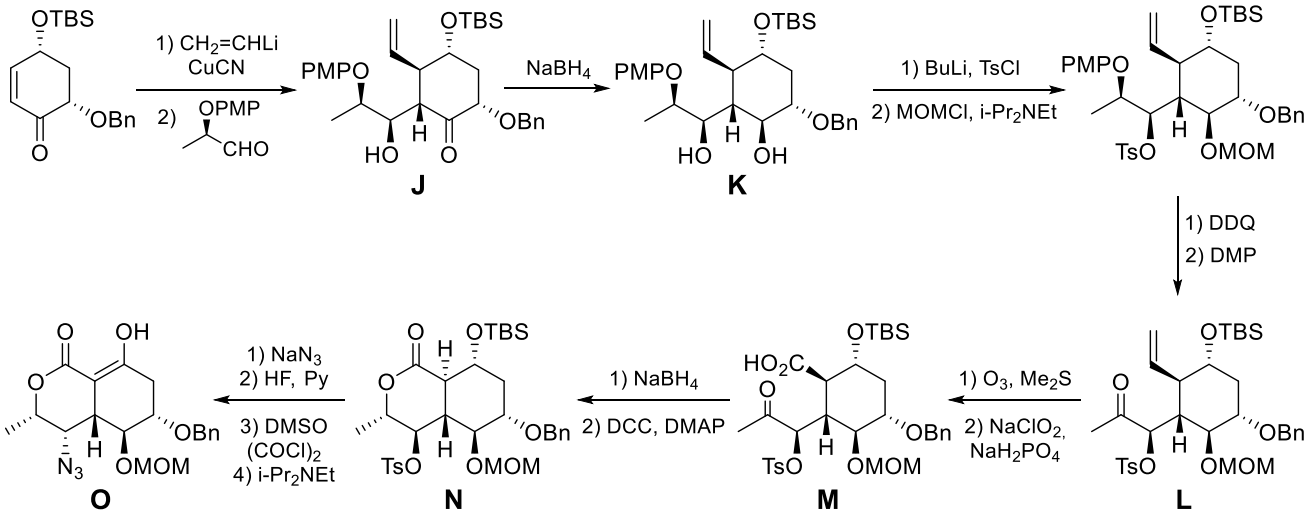


3. To obtain *Thromboxane* B_2, the reagents must be added in the following sequence:

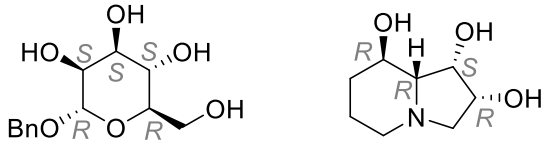
$Z_4 \rightarrow Z_2 \rightarrow Z_1 \rightarrow Z_3$ (the reduction of enone in conditions Z_3 is not diastereoselective)



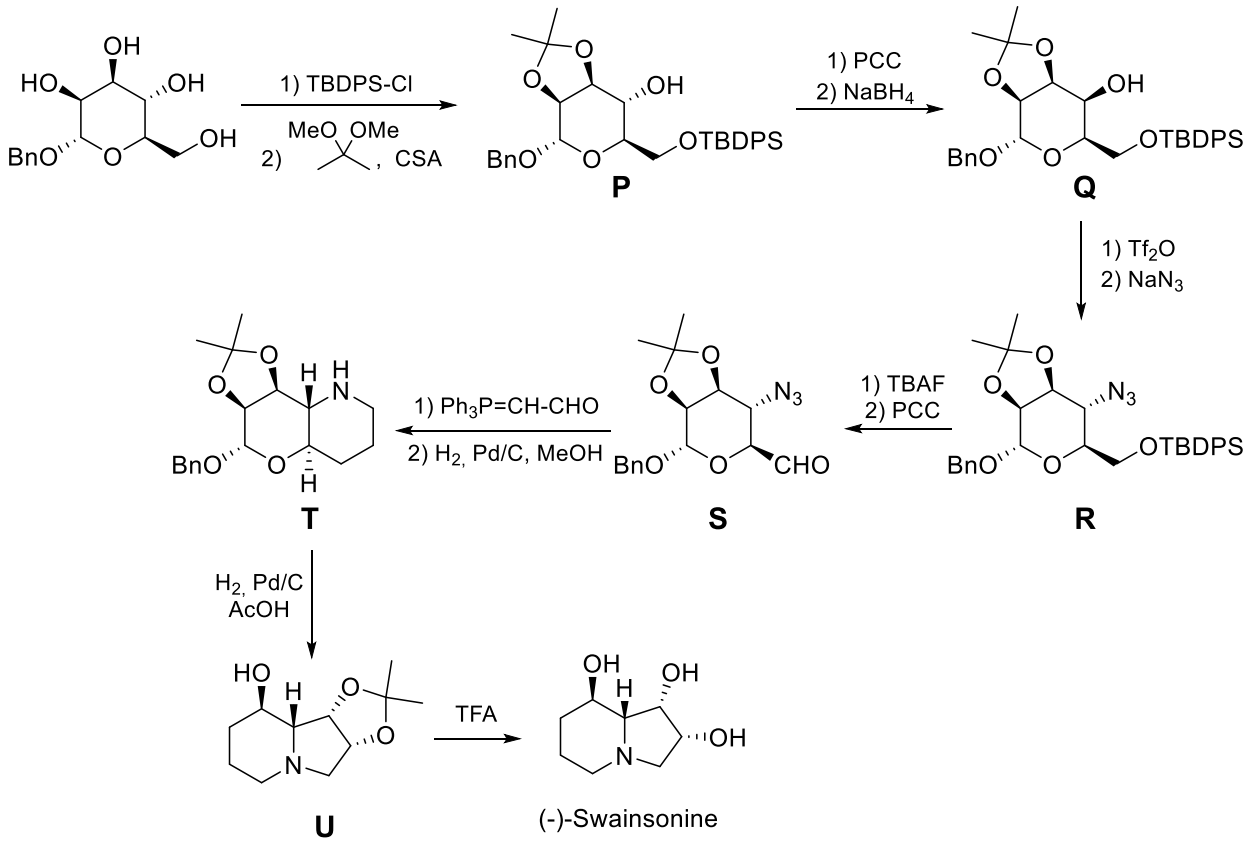
4. The 1,4-addition and coupling reactions yield **J** from the starting compound, which upon reduction yields **K**. The OH is then protected by the MOM group. During oxidation with DDQ, the PMP group is removed, and the resulting OH is then oxidized to form **L**. Ozonolysis of **L** yields an aldehyde, which upon oxidation yields carboxylic acid **M**. Reduction and condensation yield lactone **N**. Nucleophilic substitution, removal of the TBS-protecting group, Swern oxidation followed by basic treatment yields **O**. Hydrogenation in the presence of HCl reduces azide function along with removal of two other protecting groups.



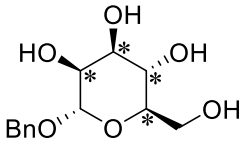
5.



6. The OH groups from the starting compound are subsequently protected, yielding **P**. Oxidation and reduction yield **Q**, which is the epimer of **P**. After the OH group is replaced by an azide, compound **R** is formed. After the protecting group is removed and the newly formed OH group is oxidised, aldehyde **S** is formed. After the Wittig reaction and reduction, **T** is obtained, which further upon reduction yields **U**. After the protecting group is removed, the final product is obtained.



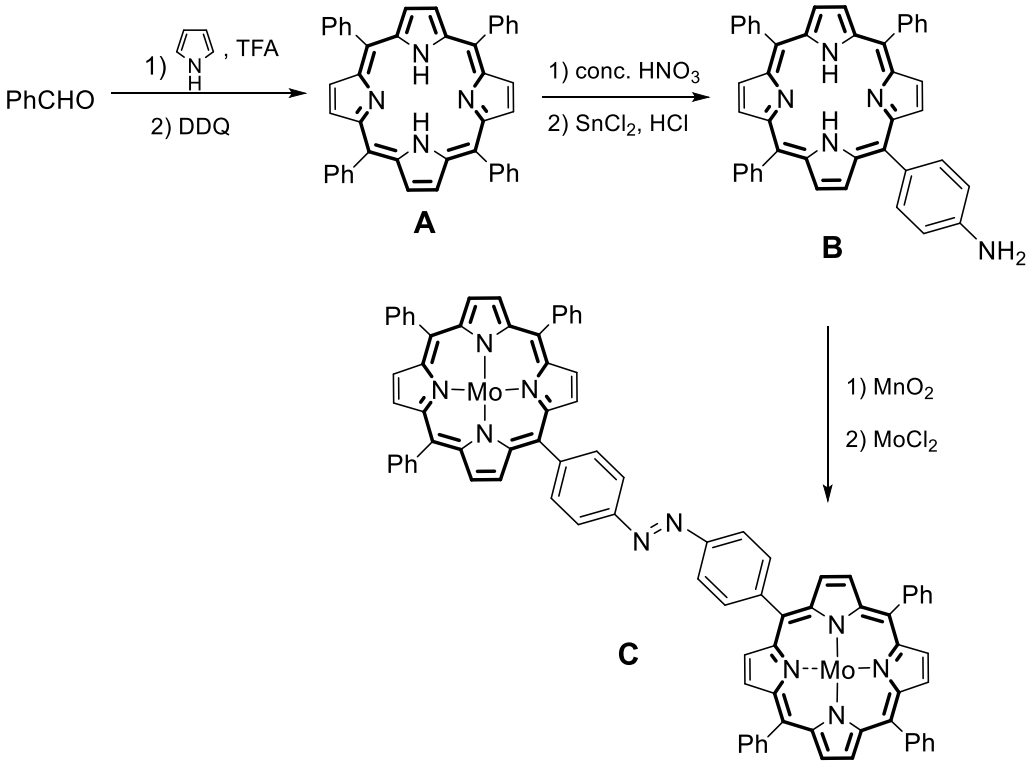
8. Preserved stereocentres:



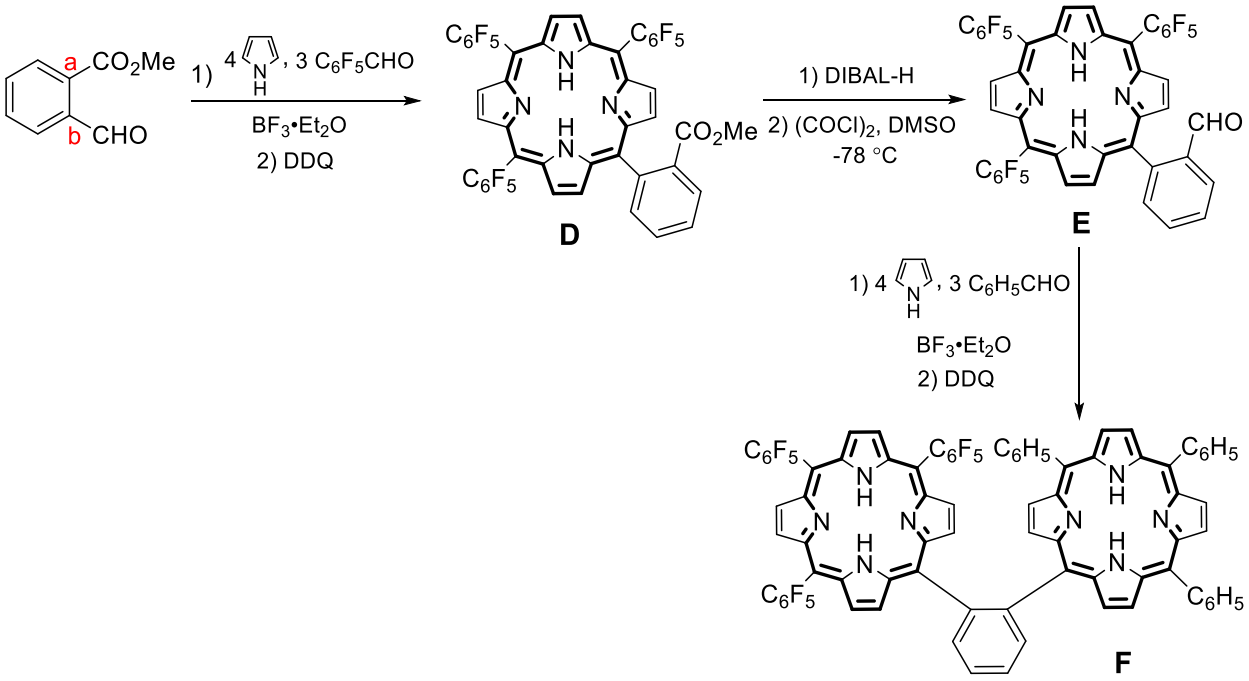
The newly formed stereocentre is absent in the final product.

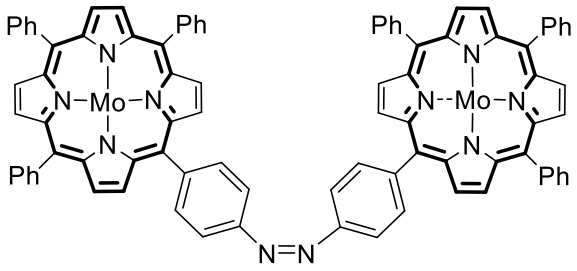
Problem 20. Pacman, Hangman, Cageman

1. According to the nitrogen content, only one nitro group was added to the structure of compound **B**. The next stage is dimerisation and formation of a diazaporphyrin with molybdenum insertion:

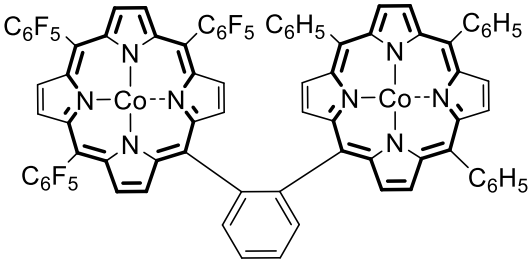


2. In this scheme stepwise formation of porphyrin rings was used.



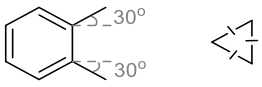


Pacman 1

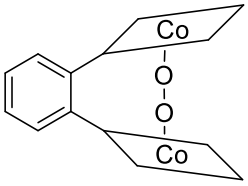


Pacman 2

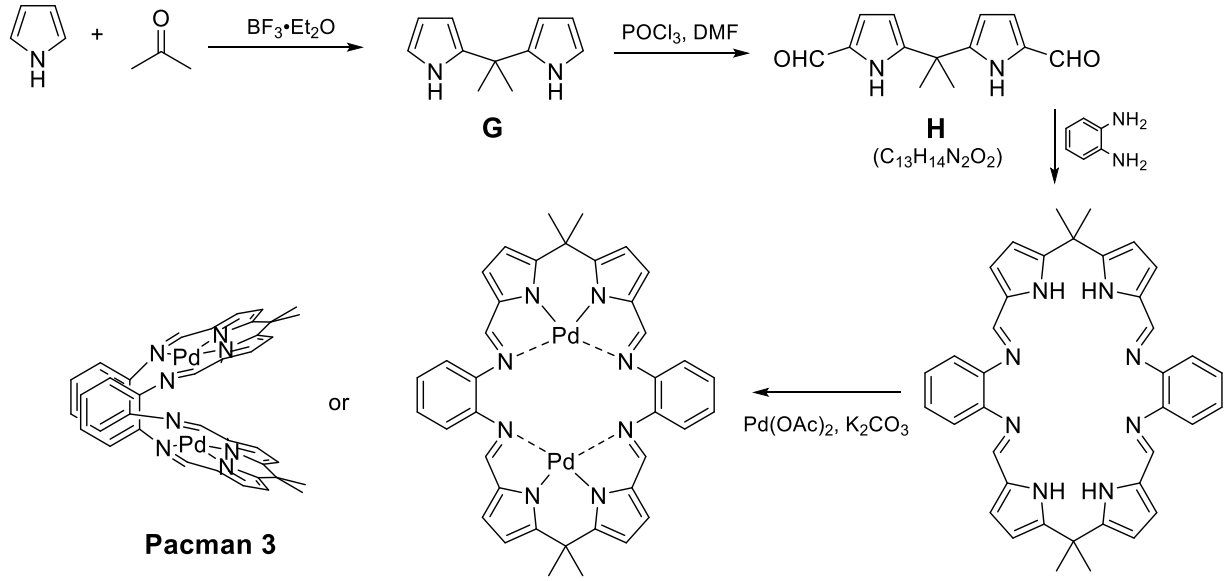
3. The porphyrin rings and metals form an equilateral triangle. Thus, the distance between metals will be equal to 4.7 + 1.39 = 6.09 Å *(students are not expected to memorise bond lengths in exam).*



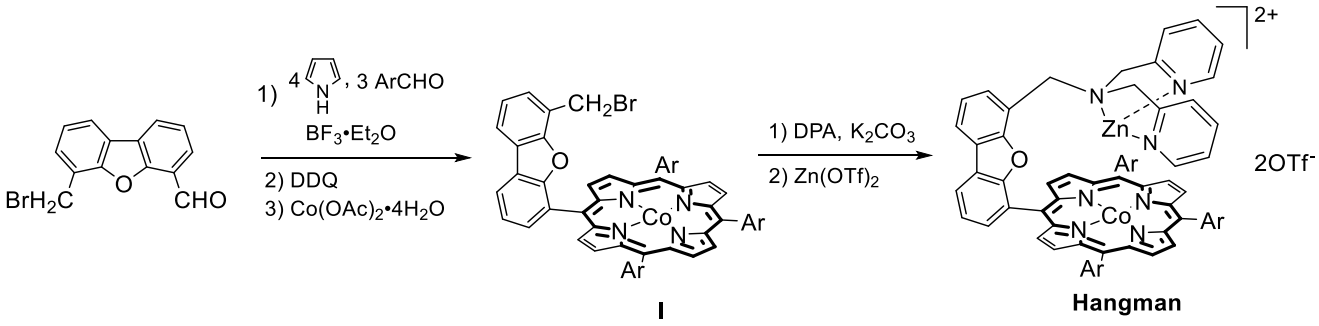
4.



5. In an excess of pyrrole, the reaction of acetone and pyrrole yields bispyrrolic compound **G** which is then formylated through the Vilsmeier–Haack reaction**.** Further macrocyclisation happens through imine condensation. Addition of the metal forces the macrocycle to the Pacman-type conformation.



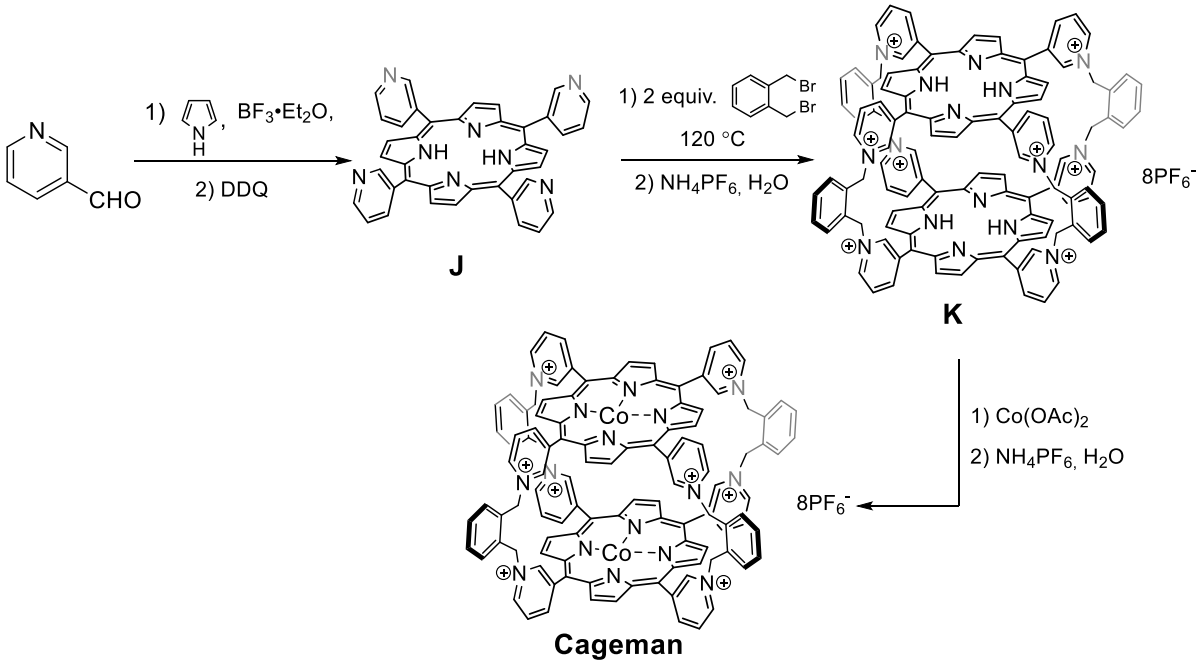
6.



7. In a molecule **J** there are 4 Py groups. In *o*-xylene dibromide, there are two benzylic bromides. These two compounds can react in 1+2, 2+4, 3+6, … manner. From the molecular weight of **J** (618) and the dimethylenebenzene unit (104), it is easy to calculate molecular weight of **K**: $618 \cdot n + (104 \cdot 2n) + (4n \cdot 145) = 1406n$. From the name – **Cageman** and the mass spectra results (for $n = 2$: $1406 \cdot 2 - 145 = 2667$), it is clear that **K** forms in a 2+4 manner.

In the mass spectrum signal is assumed as mass to charge ratio, signals from ions of higher charge (and thus lower *m/z*) can be expected. The easiest way of ionisation in case of **K** is losing anions. Therefore, the following ion signals could be expected: $[cage*7PF_6]^+$ – 2667.0, $[cage*6PF_6]^{2+}$ – 1261.0, $[cage*5PF_6]^{3+}$ – 792.3, $[cage*4PF_6]^{4+}$ –

558.0, $[cage*3PF_6]^{5+}$ – 417.4, $[cage*2PF_6]^{6+}$ – 323.7. Therefore, the molecular weight of **Cageman** is 2930 g/mol.



Problem 21. Catalyst C

1. Carbon and hydrogen in **B** indicate that it might be an organic compound. The threefold symmetry suggests that it has formulae PR_3, in which R is an organic radical. Hence, M(R) equals:

$$M(R) = \left(\frac{30.97}{0.118} - 30.97\right) \div 3 = 77 \text{ g/mol}$$

This corresponds to a phenyl group C_6H_5, hence **B = PPh_3**, a well-known ligand for large transition metals.

The overall reaction equation of compound **A** can be written like:

$$n\text{A} + m\text{PPh}_3 \rightarrow \mathbf{C} + \ldots$$

By varying the values of m and n (which are less than 5), one can determine the molar mass of compound **A**, as well as the metal **M** contained in it:

$$M(\mathbf{A}) = \frac{2.0 \cdot (77 \cdot 3 + 31) \cdot m}{7.96 \cdot n} \rightarrow A_r(\mathbf{M}) = \frac{2.0 \cdot (77 \cdot 3 + 31) \cdot m}{7.96 \cdot n} \cdot 0.3909$$

$$A_r(\mathbf{M}) = \frac{m}{n} \cdot 25.73$$

For metals in the 5th period, A_r is more than 85(Rb) and less than 120(Sn) (85-120).

(m:n)=(1:1)	(m:n)=(2:1)	(m:n)=(3:1)	(m:n)=(4:1)	(m:n)=(5:1)
25.73(<85)	51.46(<85)	77.2(<85)	102.93 **(Rh)**	128.66(>120)

(as we can see trying numbers more than 1 for n would lead to values less than 85)

From the table, when (m:n) = (4:1), the mass corresponds to the molar mass of **Rh**, which fits the conditions in the task. Other values of A_r(**M**) don't fit any metal from the 5th period. Furthermore, the molar mass of **A** is 263.3 g/mol. By subtracting the mass of the metal **M** and three H_2O molecules, we get the mass of 106.4, which corresponds to three Cl atoms.

Hence, **A = $RhCl_3 \cdot 3H_2O$**.

There are many ways to find **C**. One of them is with the molar mass of **C**, considering that **C** contains rhodium:

$$M(\mathbf{C}) = \frac{6.18}{0.88 \times \frac{2.00}{263.3}} = 924.54 \text{ g/mol}$$

By subtracting the molar mass of Rh and $x\text{PPh}_3$ molecules, we can find other species in compound **C**.

When x = 1, the remainder is 559.61;

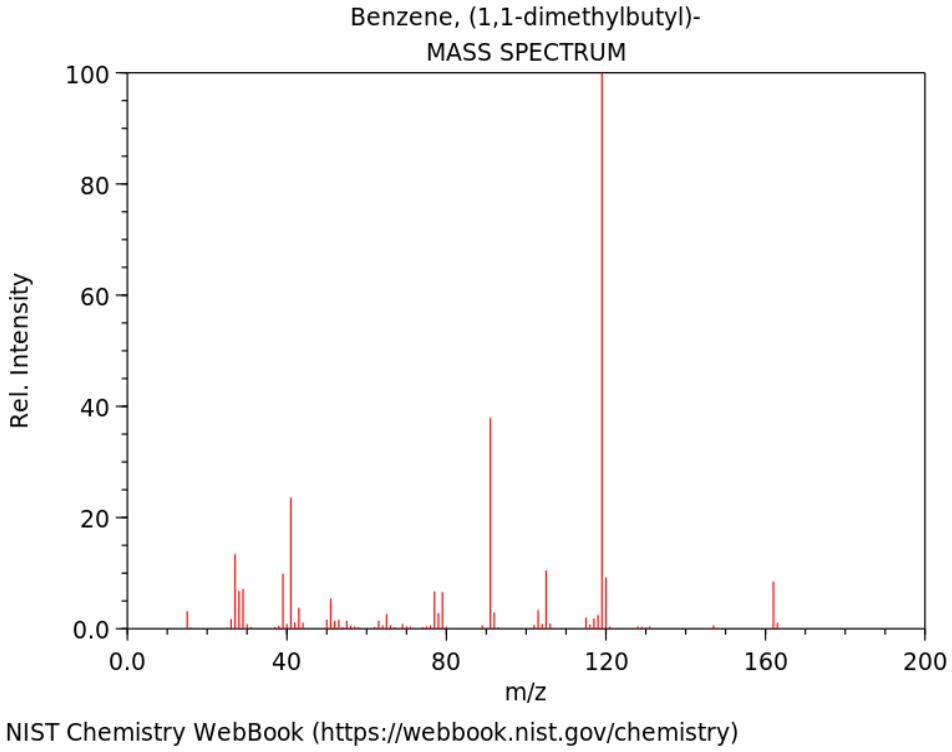
when x = 2, the remainder is 297.61;

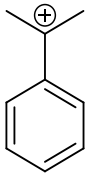
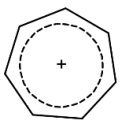
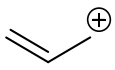
when x = 3, the remainder is 35.6, which is approximately equal to one Cl atom.

M	A	B	C
Rh	$RhCl_3 \cdot 3H_2O$	PPh_3	$RhCl(PPh_3)_3$

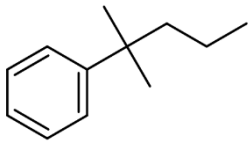
2. $RhCl_3 \cdot 3H_2O + 4PPh_3 \rightarrow RhCl(PPh_3)_3 + Ph_3PO + 2HCl + 2H_2O$

3. Mass spectrum below is copied from 'NIST' database.



m/z	119	91	41
Structure			

4.



Product

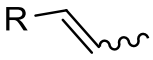
5. Natural carbon consists of ^{12}C and ^{13}C isotopes:

$12.011 = 12 \cdot x + (1 - x) \cdot 13$; x – abundance of ^{12}C

$$x = 0.989$$

$$[M]^+ : [M+1]^+ = \frac{0.989^{12}}{0.989^{11} \cdot (1 - 0.989) \cdot \frac{12!}{(12-1)!}} = \frac{0.989}{0.011 \times 12} = \mathbf{7.5}$$

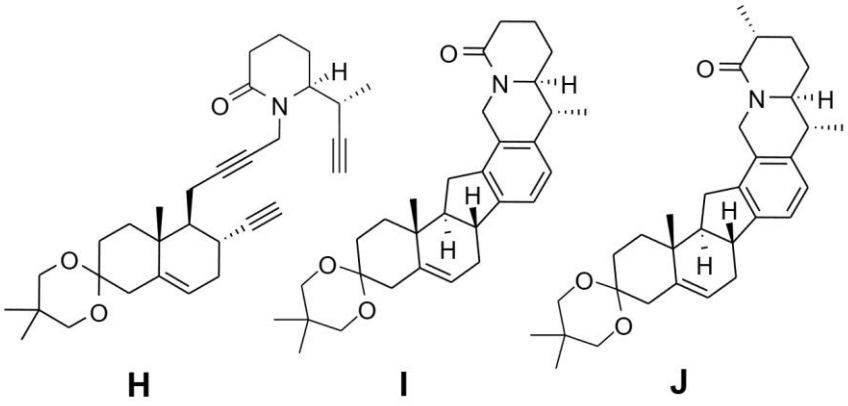
6. D



7. F is a product of oxidizing **B** with oxygen, triphenylphosphine oxide, $Ph_3P{=}O$.

$2PPh_3 + O_2 \rightarrow 2Ph_3P{=}O$

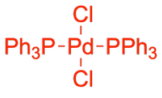
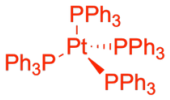
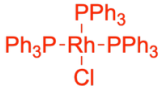
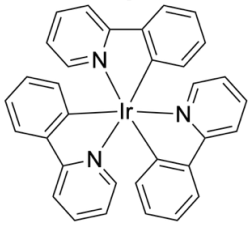
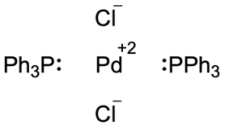
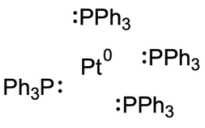
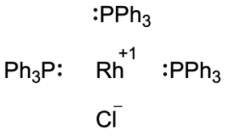
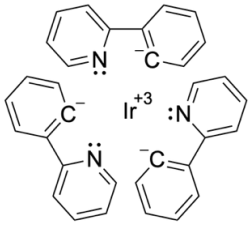
8.



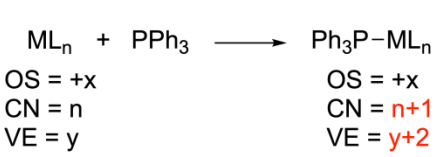
H **I** **J**

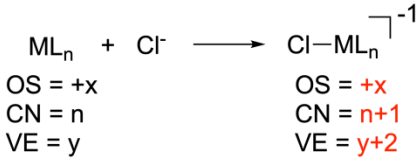
Problem 22. Transition metals in organometallic catalysis

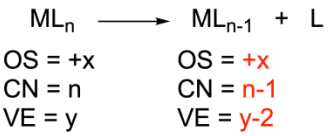
1.

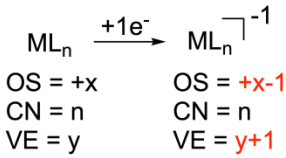
1	2	3	4
$Pd(PPh_3)_2Cl_2$	$Pt(PPh_3)_4$	$Rh(PPh_3)_3Cl$	
			
			
OS = +2 CN = 4 VE = 8 + 4*2 = 16	OS = 0 CN = 4 VE = 10 + 4*2 = 18	OS = +1 CN = 4 VE = 8 + 4*2 = 16	OS = +3 CN = 6 VE = 6 + 6*2 = 18

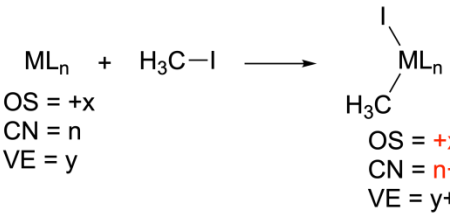
2.

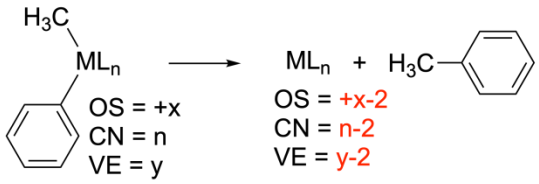


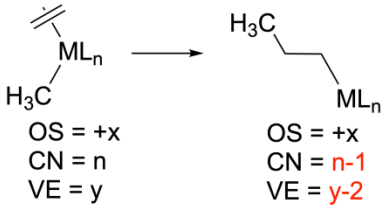




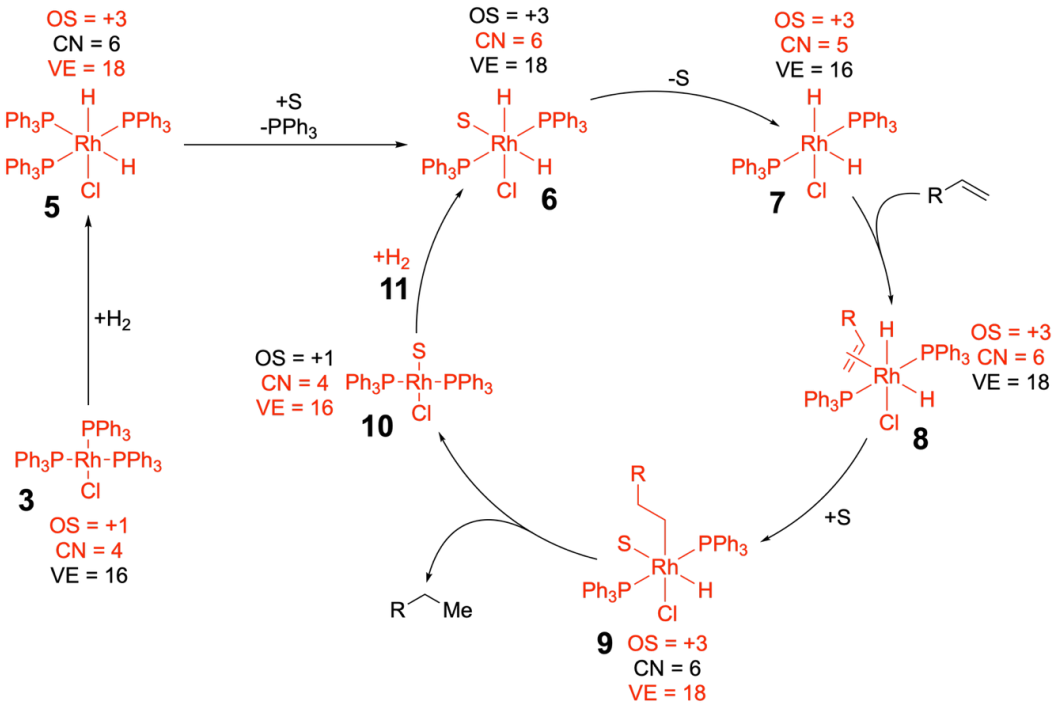






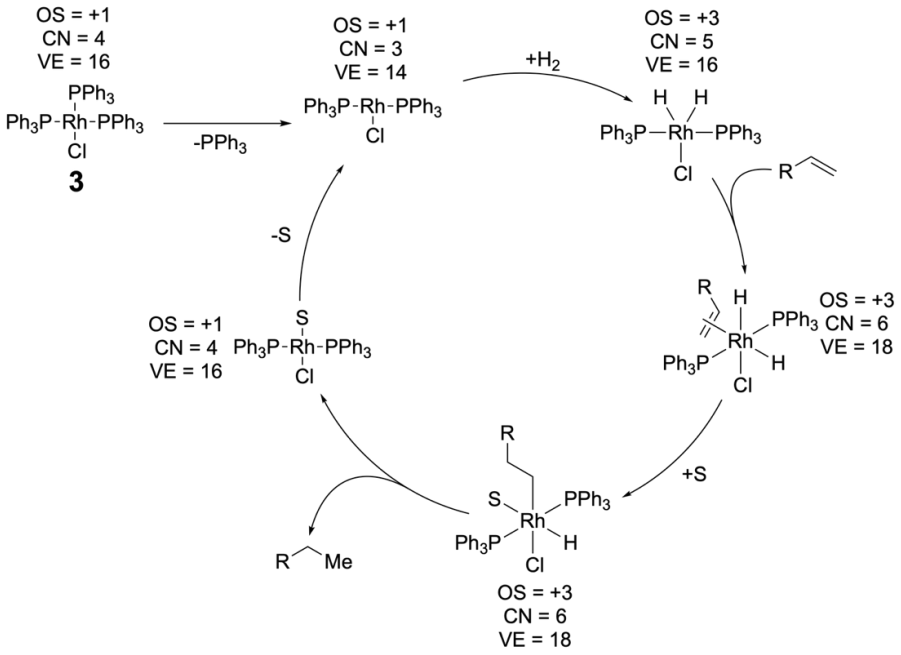


3. Any reasonable alternative stereoisomers of complexes are accepted as a correct answer since this is not a field of advanced difficulty, and it is beyond the purpose and scope of this task.

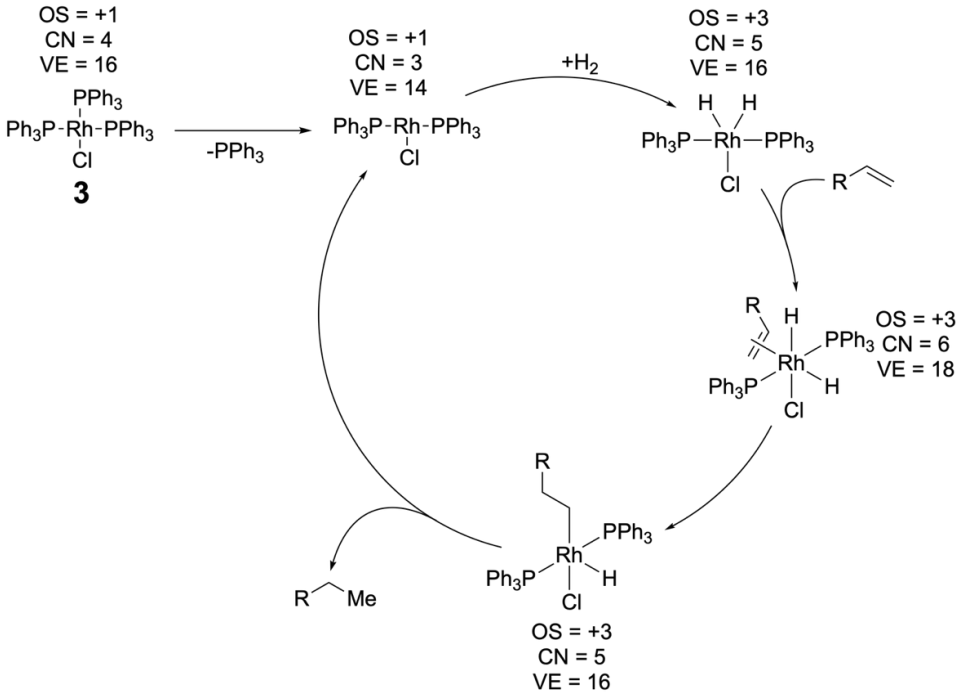


4. 11 = H_2

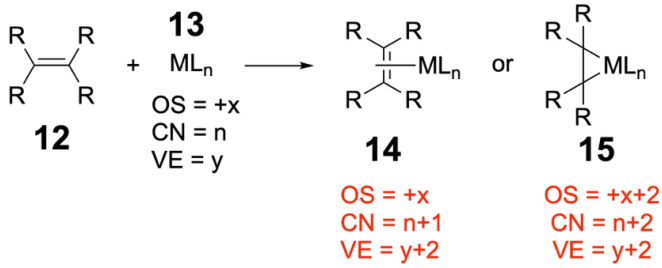
5. An alternative catalytic cycle is presented below.



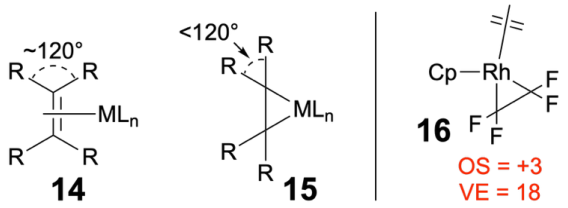
Any other reasonable catalytic cycles which fit the conditions in the task are graded fully. For instance, see an alternative option below.



6. When an alkene coordinates via dative bonding, the oxidation state of the metal is not changed, the coordination number increases (+1) and the total amount valence electrons is also increased (+2). If the metallocyclopropane mode takes place, all the parameters are increased (+2).



7. Ethylene did not significantly change its C–C bond length upon metal binding, whereas tetrafluoroethylene did. Moreover, metallocyclopropane coordination implies that substituents on the C=C bond will become non-coplanar with the C–C bond. This should decrease the angle R–C–R in the alkene upon binding (compare complex **14** where the R–C–R angle is about 120°, to an R–C–R angle of less than 120° for **15**). Therefore, ethylene has dative bond coordination, while tetrafluoroethylene has coordination closer to metallocyclopropane.

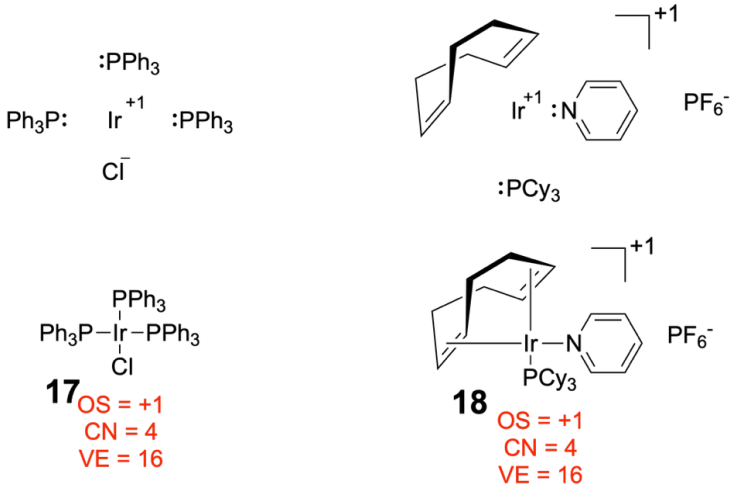


8. The geometry of complex **16** was discussed above.

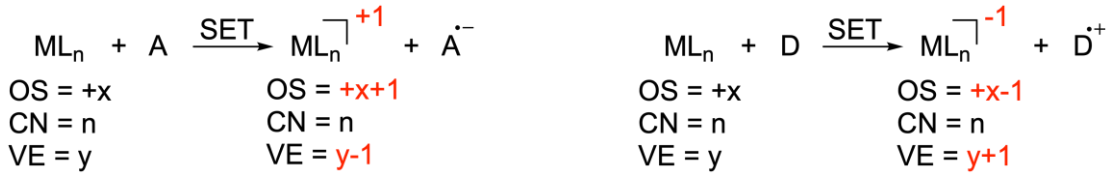
ethylene = a. alkene coordination

tetrafluoroethylene = b. metallocyclopropane coordination

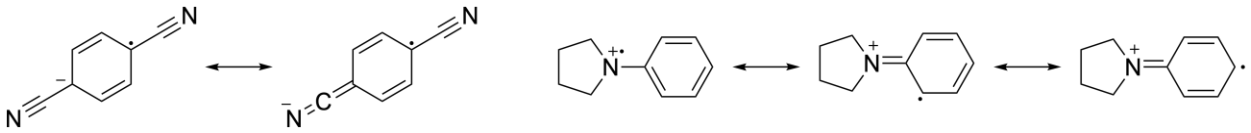
9. Complex **17** $IrCl(PPh_3)_3$ should be identical to that of rhodium ($RhCl(PPh_3)_3$), thus, it is reasonable to start with its parameters. After finding them, we know that **18** has exactly the same parameters, therefore, it should be a cationic complex (which is reasonable, since PF_6^- is a non-coordinating anion).



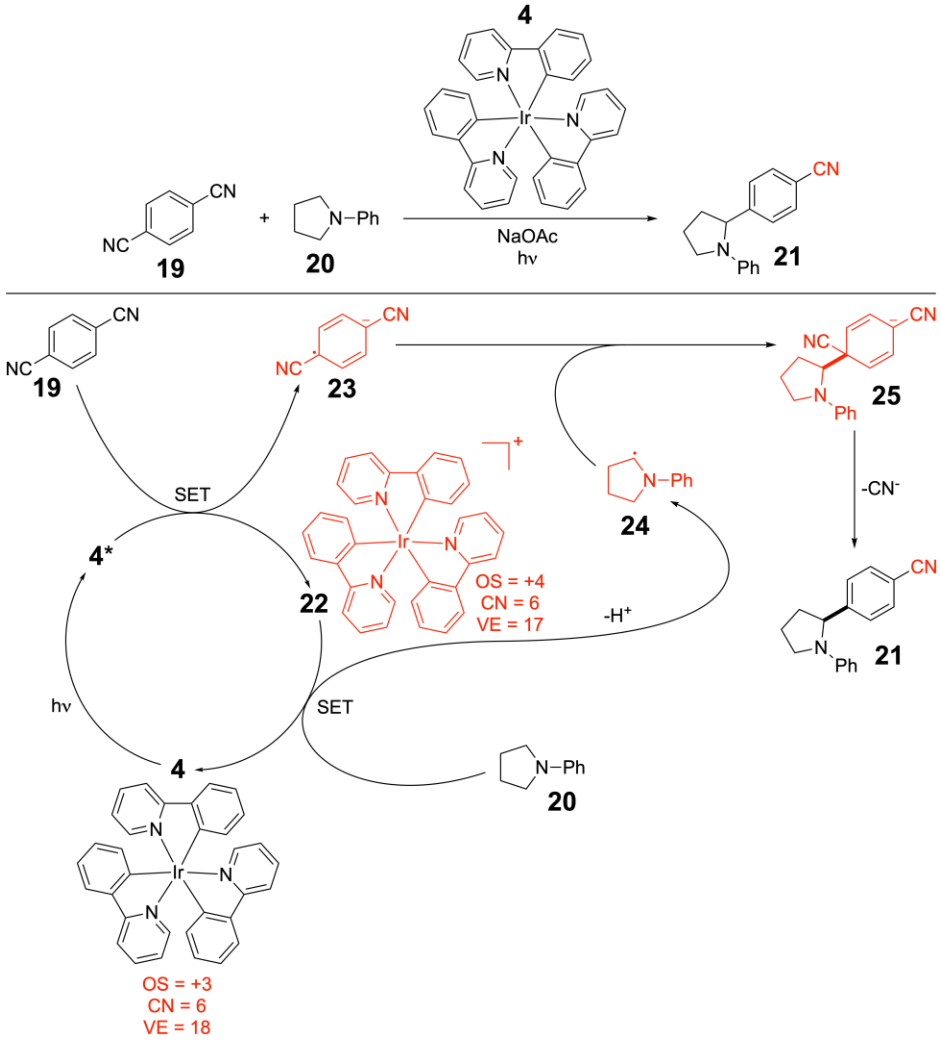
10.



11. Reasonably, electron poor arene **19** should be an electron acceptor, since it has two electron withdrawing groups to stabilize negative charge which can be demonstrated by the resonance structures shown below for the product of single electron reduction of **19**. Similarly, compound **20** will give away one electron more readily than **19**, because of the stabilization of radical cation obtained from it (see the resonance structures below).



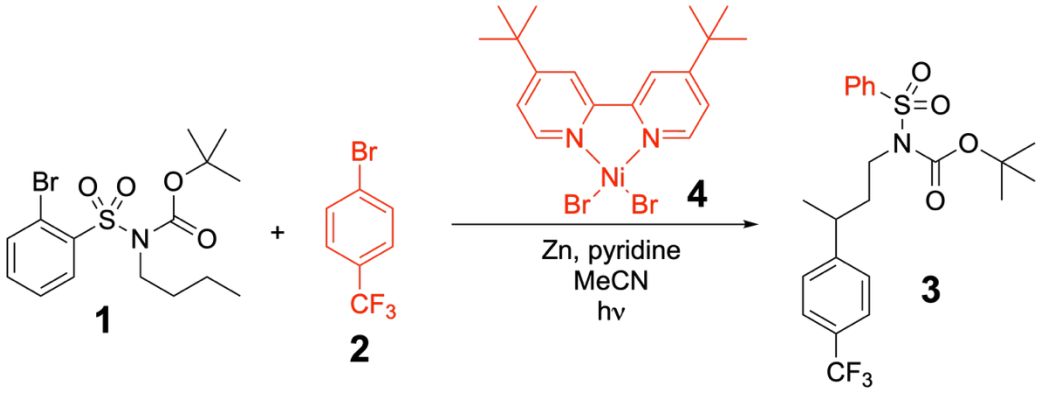
12. As discussed, **19** will be an electron acceptor, thus, excited complex **4*** should donate one electron during SET to **19**, providing cationic complex **22** and radical anion **23**. Complex **22** has iridium in a +4 oxidation state and it has only 17 VE, thus, during the SET it will be an electron acceptor to produce complex **4** again. Upon SET and deprotonation, radical **24** is generated which can be coupled with **23** creating the highlighted C–C bond in **25**.



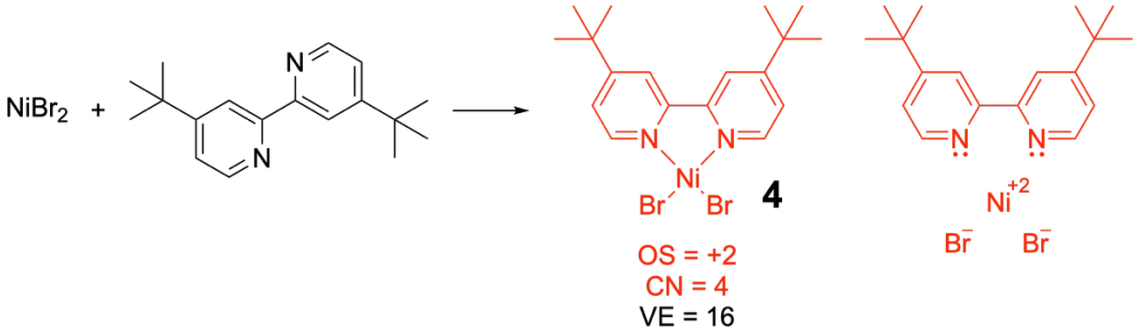
13. R = CN

Problem 23. HAT and XAT (HalAT)

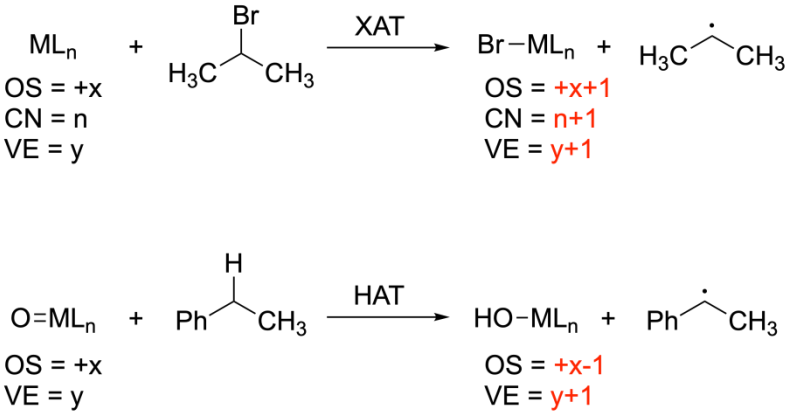
1. According to MS of **2**, there is a bromine atom there, while in **3** there is no bromine and R^2 = Ph.



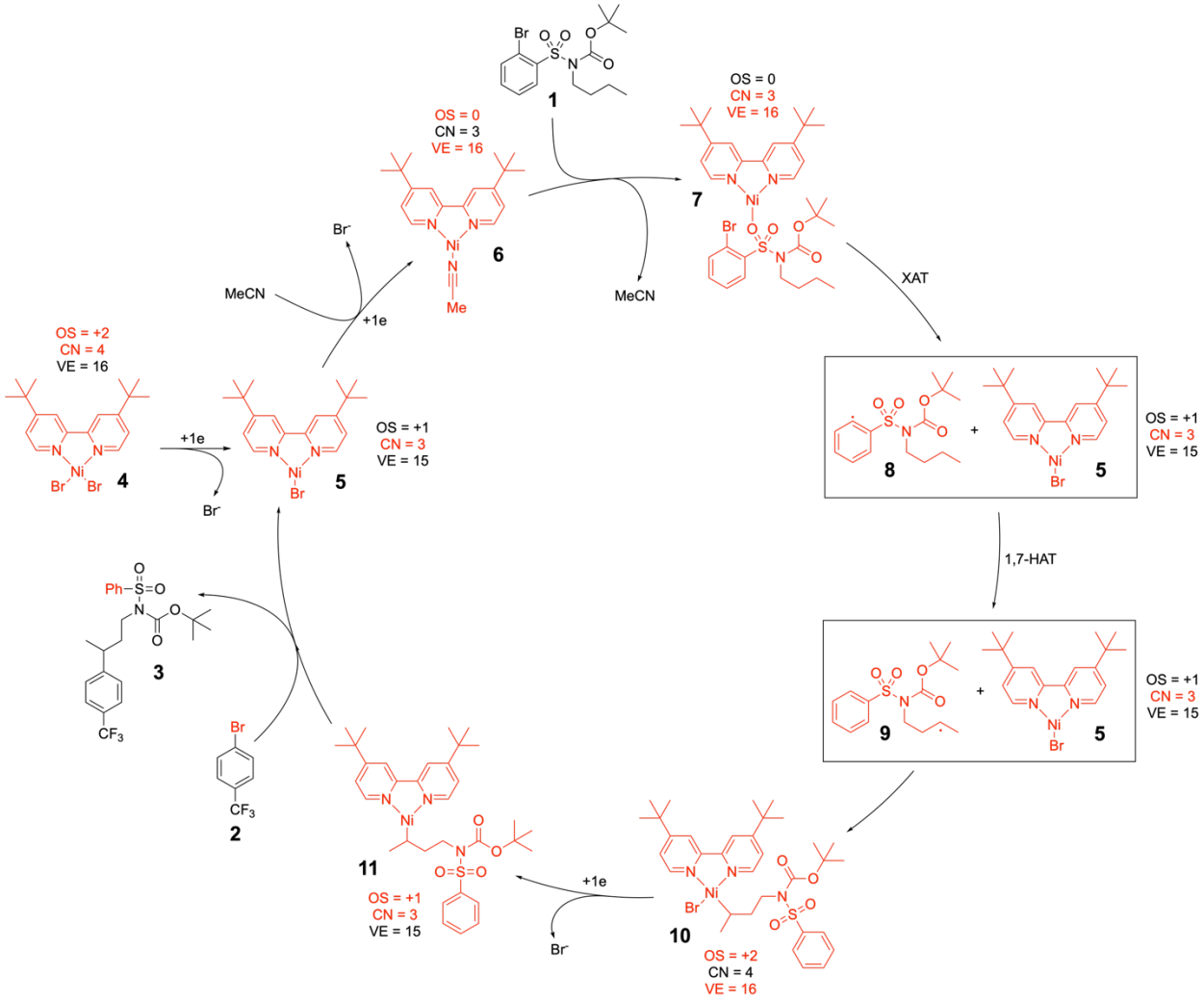
2.



3.

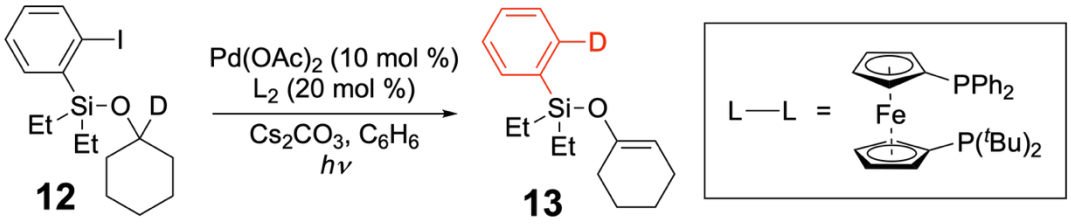


4. Two coordination modes are accepted for **7** (coordination via S=O or coordination via C=O; the latter is less favorable due to bulky ^{t}Bu-group); coordination from a nitrogen atom or another oxygen (O-^{t}Bu) are considered as incorrect answers for structure **7**.

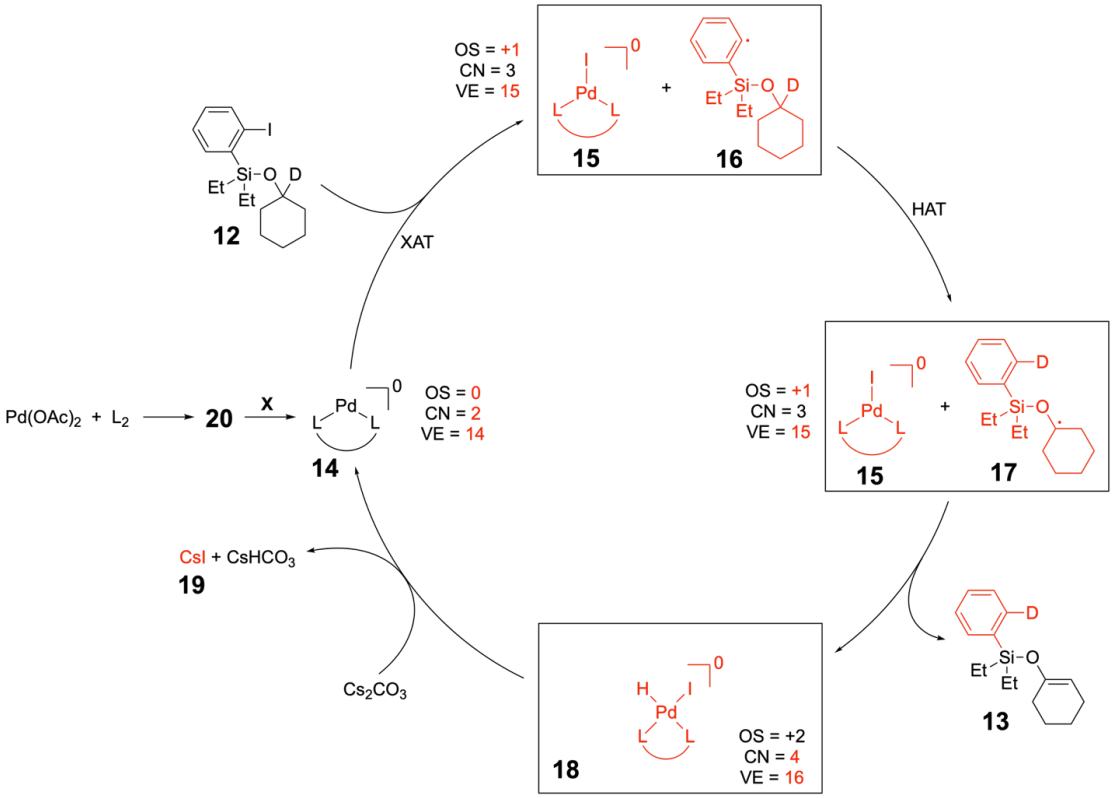


5. Zn

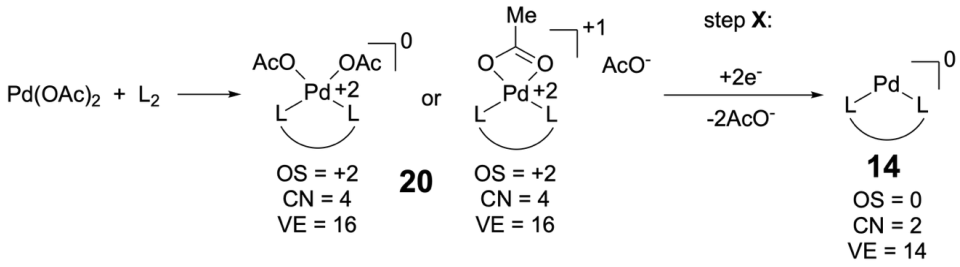
6. MS of **13** suggests that HI was eliminated from molecule **12**. Since there is a C=C bond in **13**, it is reasonable to suggest a deuterium transferred to the aromatic ring (it corresponds to the catalytic cycle of this transformation and it is conceptually similar with previous catalytic cycle).



7.



8. Evidently, a reduction should occur with acetate groups leaving the metal. In principle, excess of phosphine ligand L_2 can reduce the complex **20** to **14**.



20

9.

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

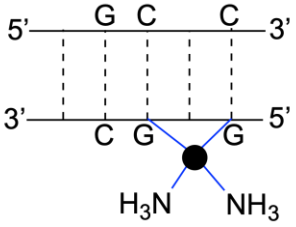
$BDE(C–H)_{\mathbf{21}} > BDE(C–H)_{\mathbf{23}}$

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

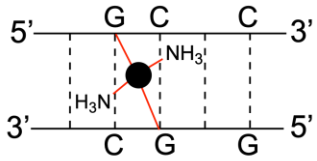
Problem 24. Anticancer complexes

1. a)

Intrastrand crosslink



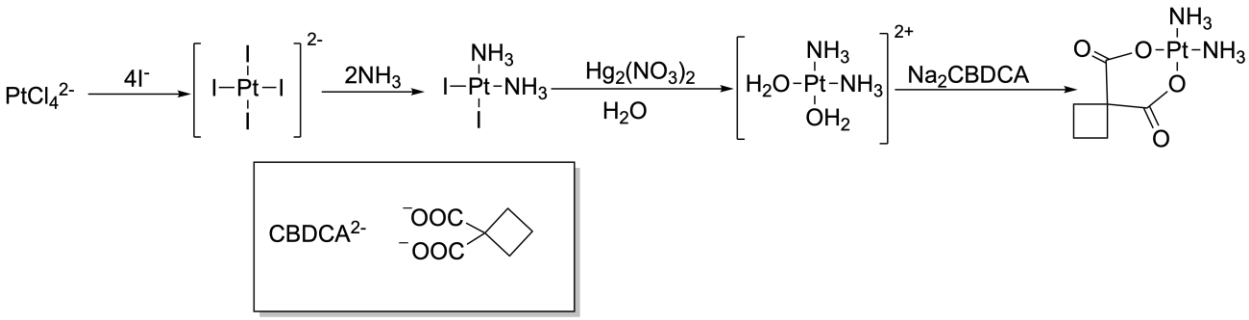
Interstrand crosslink



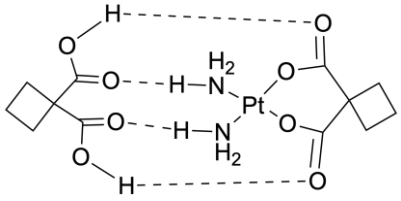
b) Cisplatin favours formation of intrastrand crosslinks, while transplatin, due to its geometry, forms a higher proportion of interstrand crosslinks.

2. Intrastrand crosslinks decrease the melting temperature as they destroy the helix structure of DNA. Interstrand crosslinks increase the duplex melting temperature as they physically tie the two strands together and prevent separation.

3.

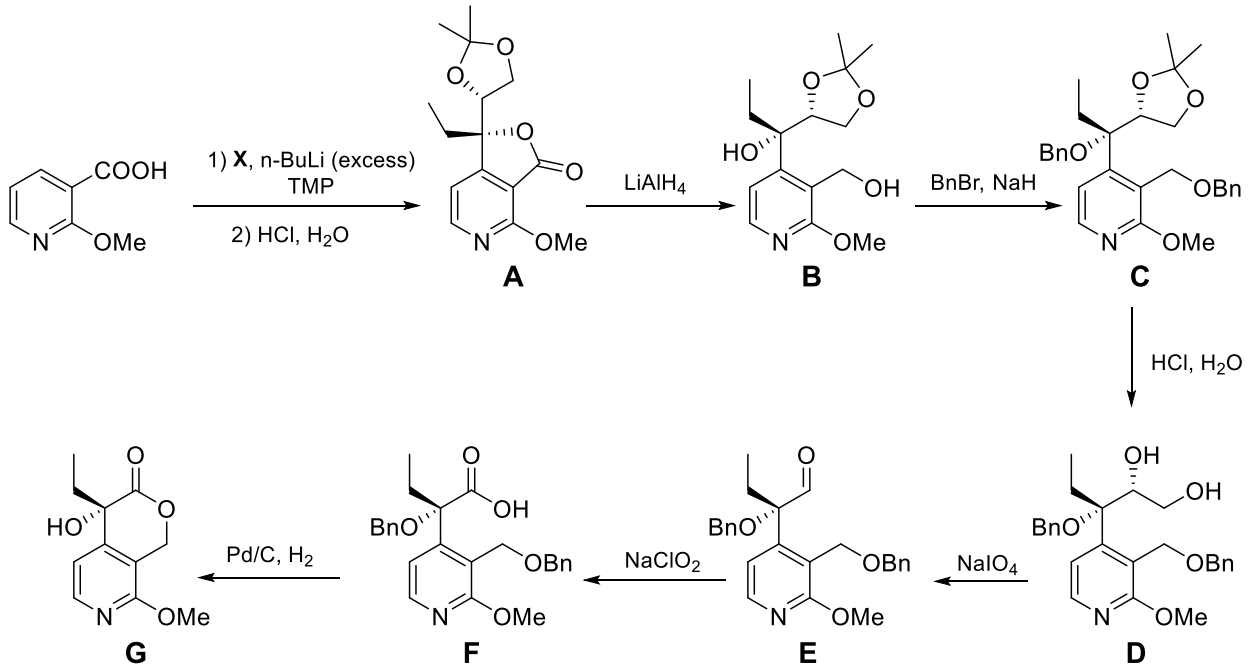


4.

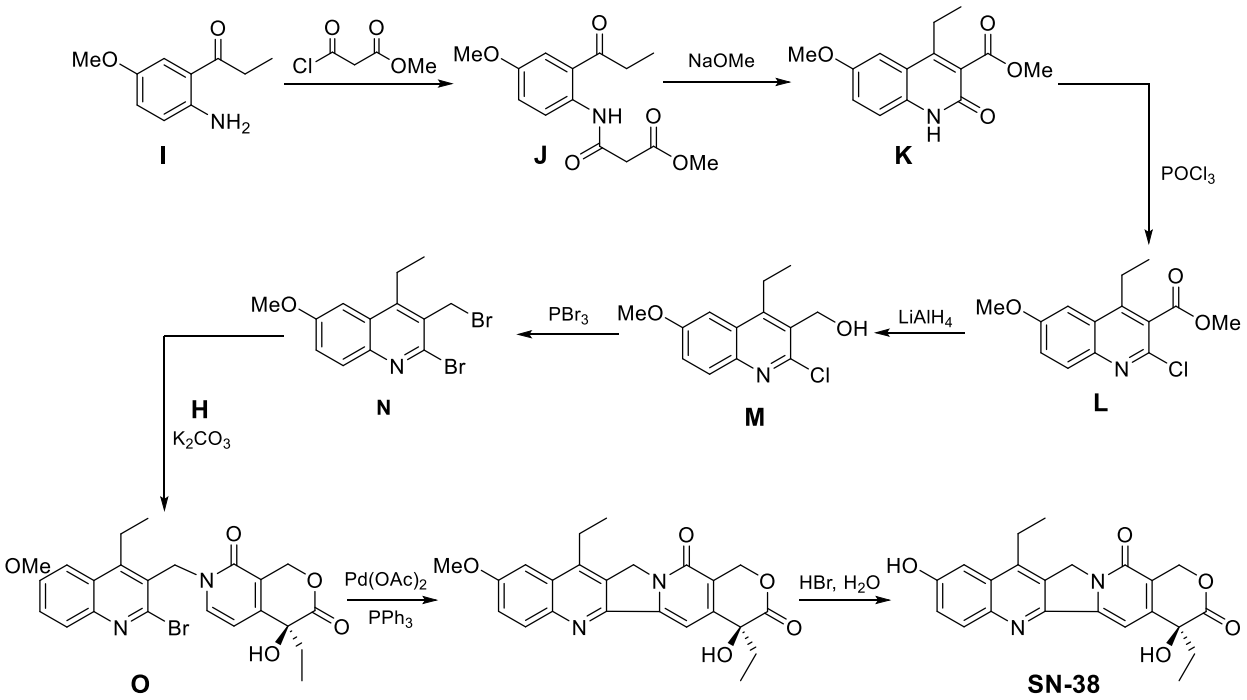


5. Dicycloplatin is more stable in aqueous solution than carboplatin because it forms a supramolecular complex with 1,1-cyclobutanedicarboxylic acid, which enhances stability through hydrogen-bonding interactions.

6.


COOH
N
OMe
1) X, n-BuLi (excess)
TMP
2) HCl, H2O
A
LiAlH4
HO
OH
B
BnBr, NaH
BnO
OBn
C
HCl, H2O
D
NaIO4
E
NaClO2
F
Pd/C, H2
G

7.


MeO
NH2
I
Cl
OMe
J
NaOMe
K
POCl3
L
LiAlH4
M
PBr3
Br
N
H
K2CO3
O
Pd(OAc)2
PPh3
HBr, H2O
SN-38

Problem 25. The equilibria in DNA

1. $n = 4 \cdot 4 \cdot 1 \cdot 1 = 16$. They are: 5'-AATT-3', 5'-TTAA-3', 5'-ATAT-3', 5'-TATA-3', 5'-GGCC-3', 5'-CCGG-3', 5'-GCGC-3', 5'-CGCG-3', 5'-AGCT-3', 5'-GATC-3', 5'-ACGT-3', 5'-CATG-3', 5'-GTAC-3', 5'-TGCA-3', 5'-CTAG-3', 5'-TCGA-3'.
2. At 10 °C:

$$\begin{cases} [\text{ss}-\text{DNA}] + 2[\text{ds}-\text{DNA}] = 1 \times 10^{-5}\ \text{M} \\ 2.44 \times 10^4 \text{cm}^{-1}\text{M}^{-1} \times 5\ \text{cm} \times [\text{ss}-\text{DNA}] + 2.93 \times 10^4 \text{cm}^{-1}\text{M}^{-1} \times 5\ \text{cm} \times [\text{ds}-\text{DNA}] = 0.89 \end{cases} \rightarrow$$
$$\begin{cases} [\text{ss}-\text{DNA}] = 3.23 \times 10^{-6}\ \text{M} \\ [\text{ds}-\text{DNA}] = 3.38 \times 10^{-6}\ \text{M} \end{cases} \rightarrow K(10\ ^\circ\text{C}) = \frac{3.38 \times 10^{-6}}{(3.23 \times 10^{-6})^2} = 3.24 \times 10^5$$

At 30°C:

$$\begin{cases} [\text{ss}-\text{DNA}] + 2[\text{ds}-\text{DNA}] = 1 \times 10^{-5} \\ 2.44 \times 10^4 \text{cm}^{-1}\text{M}^{-1} \times 5\ \text{cm} \times [\text{ss}-\text{DNA}] + 2.93 \times 10^4 \text{cm}^{-1}\text{M}^{-1} \times 5\ \text{cm} \times [\text{ds}-\text{DNA}] = 1.17 \end{cases} \rightarrow$$
$$\begin{cases} [\text{ss}-\text{DNA}] = 8.97 \times 10^{-6} M \\ [\text{ds}-\text{DNA}] = 5.13 \times 10^{-7} M \end{cases} \rightarrow K(30\ ^\circ\text{C}) = \frac{5.13 \times 10^{-7}}{(8.97 \times 10^{-6})^2} = 6.38 \times 10^3$$

Combining the results at different temperatures we obtain the following:

$$\begin{cases} \Delta_\text{r}H^\circ - 283.15 \times \Delta_\text{r}S^\circ = -8.314 \times 283.15 \times \ln(3.24 \times 10^5) = -29870.1\ \text{J mol}^{-1} \\ \Delta_\text{r}H^\circ - 303.15 \times \Delta_\text{r}S^\circ = -8.314 \times 303.15 \times \ln(6.38 \times 10^3) = -22080.9\ \text{J mol}^{-1} \end{cases} \rightarrow$$
$$\begin{cases} \Delta_\text{r}H^\circ = -140157.0\ \text{J mol}^{-1} = -140.2\ \text{kJ mol}^{-1}, \\ \Delta_\text{r}S^\circ = -389.5\ \text{J mol}^{-1}\ \text{K}^{-1}. \end{cases}$$

3. At T_m:

$$\begin{cases} [\text{ds}-\text{DNA}] = 2.5 \times 10^{-6}\text{M} \\ [\text{ss}-\text{DNA}] = 5.0 \times 10^{-6}\text{M} \end{cases} \rightarrow K(T_\text{m}) = \frac{2.5 \times 10^{-6}}{(5.0 \times 10^{-6})^2} = 1.0 \times 10^5$$

$$-140157.0 - T_\text{m} \times (-389.5) = -8.314 \times T_\text{m} \times \ln(1.0 \times 10^5)$$

$T_\text{m} = 288.85\ \text{K} = 15.70\ ^\circ\text{C}$.

4. Dimerisation with pairing four pairs of complementary nucleotides:

```
5'-ACCGG-3'          5'-ACCGG-3'
     +          ⇌         ||||
3'-GGCCA-5'          3'-GGCCA-5'
```

```
5'-CCGGT-3'          5'-  CCGGT-3'
     +          ⇌         ||||
3'-TGGCC-5'          3'-TGGCC-5'
```

Dimerisation with pairing six pairs of complementary nucleotides:

5’-ACCGGT-3’ + 3’-TGGCCA-5’ ⇌ 5’-ACCGGT-3’ | | | | | | 3’-TGGCCA-5’

5. At T_m:

$$\begin{cases}[\text{ss}-\text{DNA}] = \frac{1}{2}[ss-DNA]_0 \\ [\text{ds}-\text{DNA}] = \frac{1}{4}[\text{ss}-\text{DNA}]_0\end{cases} \rightarrow K(T_m) = \frac{\frac{1}{4}[\text{ss}-\text{DNA}]_0}{(\frac{1}{2}[\text{ss}-\text{DNA}]_0)^2} = \frac{1}{[\text{ss}-\text{DNA}]_0}$$

$$\Delta_r H° - T_m \times \Delta_r S° = -R \times T_m \times \ln\frac{1}{[ss-DNA]_0},$$

$$\frac{1}{T_m} - \frac{\Delta_r S°}{\Delta_r H°} = \frac{R}{\Delta_r H°}\ln[\text{ss}-\text{DNA}]_0,$$

$$\frac{1}{T_m} = \frac{R}{\Delta_r H°}\ln[\text{ss}-\text{DNA}]_0 + \frac{\Delta_r S°}{\Delta_r H°} \rightarrow k = \frac{R}{\Delta_r H°}, b = \frac{\Delta_r S°}{\Delta_r H°}.$$

$\Delta_r H°$(5’-ACCGG-3’) = 8.314 / (-5.15×10^{-5}) = −161436.9 J mol^{-1} = −161.4 kJ mol^{-1}.

$\Delta_r H°$(5’-CCGGT-3’) = 8.314 / (-4.80×10^{-5}) = −173,208.3 J mol^{-1} = −173.2 kJ mol^{-1}.

6. $E_{A\text{-}CG}$ = [−161436.9 – (−140157.0)] / 2 = −10639.9 J mol^{-1} = −10.6 kJ mol^{-1}.

$E_{T\text{-}GC}$ = [−173208.3 – (−140157.0)] / 2 = −16525.6 J mol^{-1} = −16.5 kJ mol^{-1}.

7. $\Delta_r H°$(5’-ACCGGT-3’) = $\Delta_r H°$(5’-CCGG-3’) + 2$E_{A\text{-}CG}$ + 2$E_{T\text{-}GC}$ + 2$E_{A=T}$ =
= −140157.0 + 2×(−10639.9) + 2×(−16525.6) + 2×(−28842.0) = −252172.0 J mol^{-1} =
= −252.2 kJ mol^{-1}.

Problem 26. Sensing of acetaldehyde

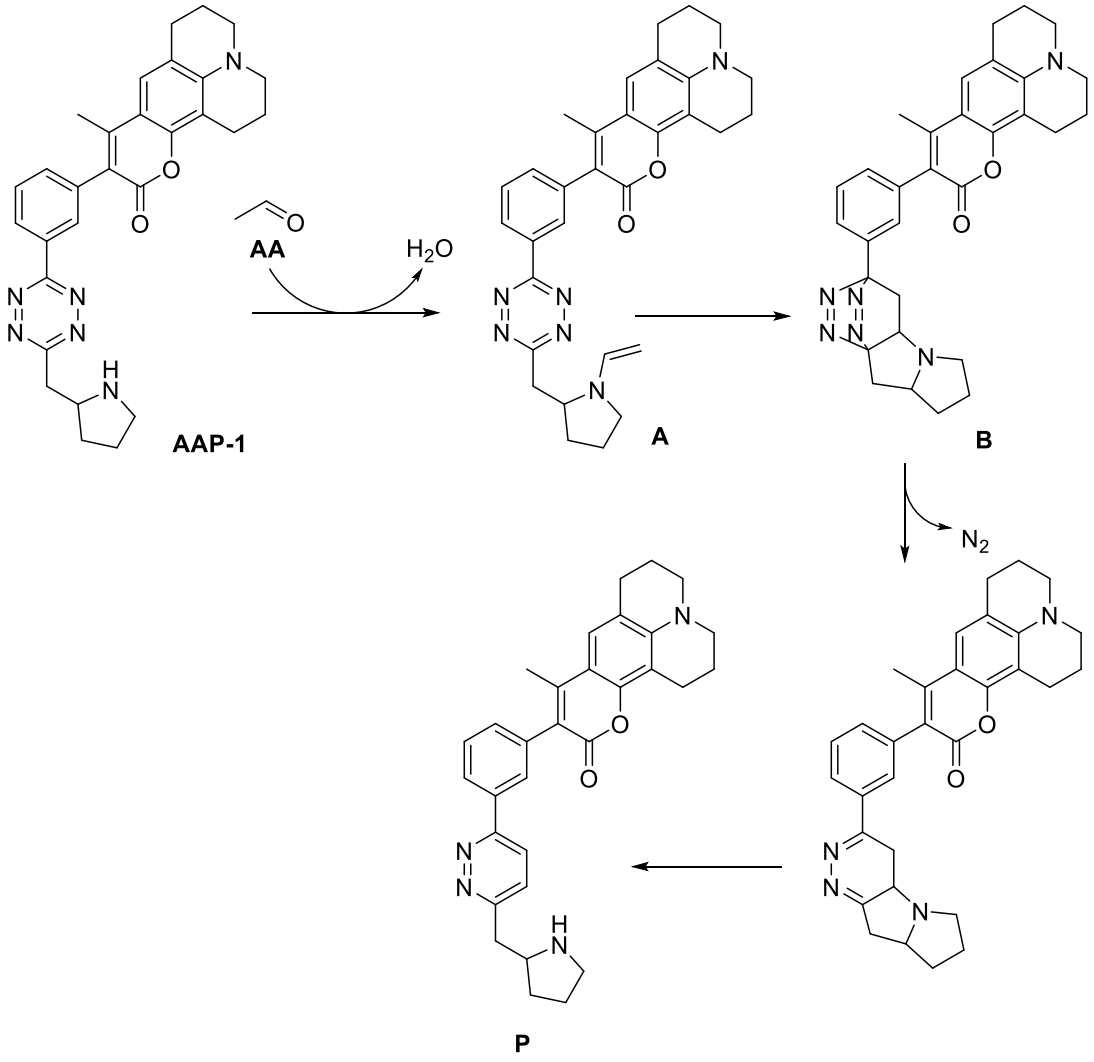
1.

$$M_{\mathbf{AAP-1}} - M_{\boldsymbol{P}} = 494 - 492 = M_{\mathbf{AAP-1}} - \left(M_{\mathbf{AAP-1}} + M_{\mathbf{AA}} - M_{\mathrm{H_2O}} - M_{\mathbf{C}}\right) = M_{\mathbf{C}} - 26$$

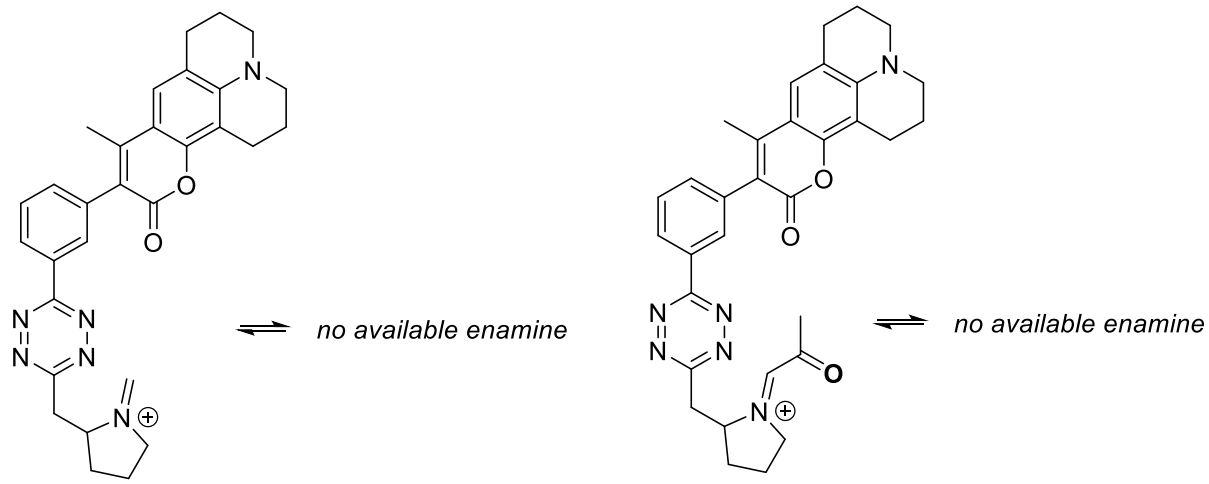
$$M_{\mathbf{C}} = 28$$

Therefore this is nitrogen (N_2). This means that the tetrazine ring of **AAP-1** is involved in the reaction.

By analysing the mechanism, we can understand that **B** is a product of intramolecular [4+2] cycloaddition of **A**, and **P** is formed as a result of retro [4+2] cycloaddition followed by aromatisation. The intermediate product of the retro [4+2] cycloaddition of **B** does not count as **P** given that **P** contains the same number of rings as **AAP-1.** The complete deciphered mechanism looks like this:



2. In case of formaldehyde and methylglyoxal, iminium-enamine tautomerisation cannot occur:



In case of methylglyoxal, keton group will not be involved in the reaction due to less electrophility.

3. The Gibbs energy changes are as follows:

ΔG (**A** → **B**) = −21.74 – (−8.36) = −13.38 kJ·mol^{-1}.

ΔG (**A'** → **B'**) = −19.23 – (−12.96) = −6.27 kJ·mol^{-1}.

Both processes are thermodynamically favorable.

4. By using an Arrhenius equation, we obtain the following ratio of rate constants:

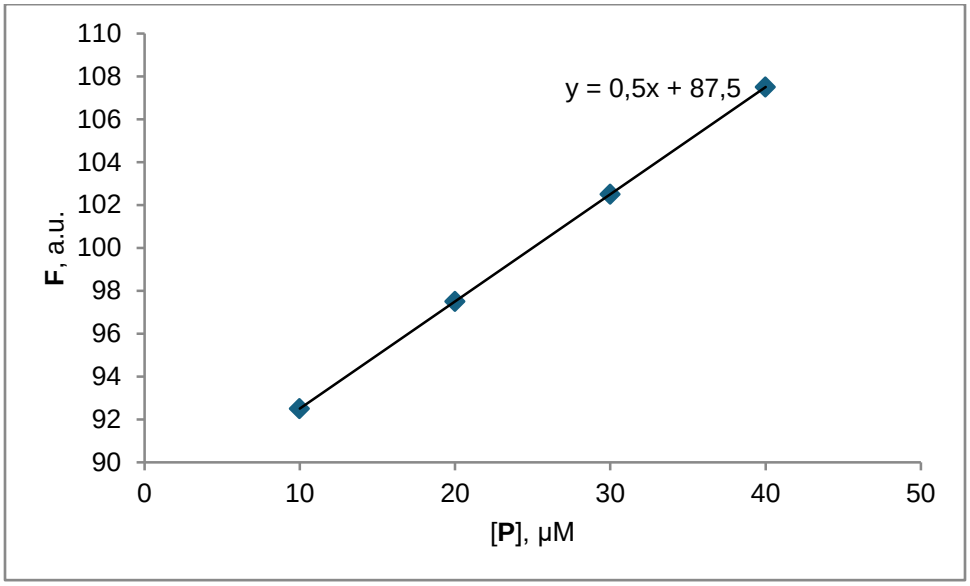
$$\frac{k_{\mathbf{A}\rightarrow\mathbf{B}}}{k_{\mathbf{A'}\rightarrow\mathbf{B'}}} = e^{\frac{(68.13-(-12.96))-(63.95-(-8.36))\cdot 10^3}{8.314\cdot 310}} = 30$$

This means that hexanal, due to steric effects, will react very slowly with **AAP-1**.

5. Molecules of **P** must emit photons with lower energy than the photons they absorb during excitation. This means that wavelength of emitted photons must be higher than wavelength of absorbed photons. So, **X** – excitation, **Y** – emission.

6. From the data in the table in the condition, we construct the following graph and find the corresponding parameters of the linear function:

$$F = 87.5 + 0.5[\mathbf{P}]$$



So, F_0 = 87.50 arb.u., k = 0.5 arb.u.·μM^{-1}.

7. Let's calculate the concentration of **AA** at the beginning and at the end of the enzymatic reaction:

$$[\mathbf{AA}]_0 = \frac{137.50 - 87.50}{0.5} = 100.0\ \mu\text{M}; \quad [\mathbf{AA}]_t = \frac{90.65 - 87.50}{0.5} = 6.3\ \mu\text{M}.$$

Due to $[\text{S}]_0 \gg K_\text{M} = 0.2\ \mu\text{M}$ during this enzymatic reaction, we simplify Michaelis-Menten kinetics as follows:

$$r = \frac{r_{\max}[\text{S}]_0}{K_\text{M} + [\text{S}]_0} \approx \frac{r_{\max}[\text{S}]_0}{[\text{S}]_0} = r_{\max} = 4.1\ \text{mM} \cdot \text{min}^{-1} = 6.83 \cdot 10^{-5}\ \text{M} \cdot \text{s}^{-1}$$

$$t = \frac{(100.0 - 6.3) \cdot 10^{-6}}{6.83 \cdot 10^{-5}} = 1.37\ \text{s}$$

8. Co-treatment of the cell cultures with ethanol and Fomepizole, which inhibits ADH, markedly suppresses **AAP-1** activation, with fluorescence signals comparable to basal levels ($F_1 \approx F_3 < F_2$), because there is no conversion of ethanol to **AA**. Conversely, inhibition of ALDH with CVT-10216 in the presence of ethanol, which leads to the intracellular accumulation of **AA**, further enhances the fluorescence signal ($F_2 < F_4$). So, $F_1 \approx F_3 < F_2 < F_4$.

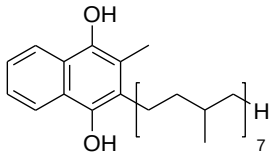
Problem 27. Archaeal metabolism

1. In the 3.5-4 billion years since the emergence of archaea, the Earth's appearance has changed dramatically. Moreover, a number of physical parameters, such as temperature, have fluctuated cyclically due to alternating glacial and interglacial periods. Radioactivity has gradually decreased over this time, as long-lived isotopes gradually decayed, reducing natural background radiation. According to the standard model of stellar evolution, the Sun's luminosity increases continuously throughout the main sequence.

The chemical composition of the environment has also undergone significant changes: initially, there was virtually no oxygen on Earth, the planet's early atmosphere was denser than it is today (and therefore had higher atmospheric pressure), the salt composition and pH of the water were largely determined by violent volcanic activity, etc. Thus, it is impossible to propose any chemical parameter that has been constant for as long as there is life on Earth.

In fact, the only physical parameter that has remained approximately constant throughout the entire era of life on Earth is gravitational force. Changes in gravity would only be possible as a result of catastrophic events, such as a collision between Earth and a large celestial body, ejecting a significant portion of the planet's mass into outer space, but archaea would not have survived such a catastrophe. It should be noted that studies in microgravity conditions have demonstrated the influence of gravity on the metabolic rate of archaea, including through changes in gene expression. Moreover, since gravitational forces have an indirect effect on archaea (changes in sedimentation, convection, and nutrient/waste distribution in aquatic environments), archaea can easily adapt to changes in these forces, unlike other factors: for example, a sharp increase in radiation is more likely to destroy the archaea population.

2. The two-electron reduction of menaquinone essentially corresponds to the conversion of the quinone to hydroquinone:



MKH_2-7

3. Menaquinone is found predominantly in anaerobic organisms (although it is not an obligate signature) because its low standard redox potential corresponds to the low standard potential terminal electron acceptors (*e.g.*, fumarate, $E^{\circ\prime} = +0.03$ V) typical of anaerobic respiration. This ensures efficient electron transfer under energy-limited conditions. Ubiquinone is characteristic of aerobes (which include the vast majority of eukaryotes), in which oxygen ($E^{\circ\prime} = +0.82$ V) serves as the terminal electron acceptor in the respiratory chain. Ubiquinone's higher potential allows for efficient electron transfer in the aerobic respiratory chain. Thus, menaquinone is suitable for the low-potential systems of anaerobes, while ubiquinone is suitable for the high-potential systems characteristic of aerobes.

4. The first thing to consider is a way of oxidation of ubiquinol (a hydroquinone derivative) with a bromine-containing reagent to ubiquinone. It should also be noted that the benzene ring in ubiquinol is completely devoid of hydrogen atoms, which prevents electrophilic substitution reactions. The reaction with Br_2 should be carried out strictly in the dark to avoid side processes, such as benzylic bromination.

Second, ubiquinol, which is predominantly found in the mitochondrial membrane of eukaryotes, is a mixture of several compounds differing in the number of isoprenyl groups (from 6 to 10). Note how carefully the formulae and names of the three coenzymes are reflected in the problem statement, with the number of C_5H_7 units always indicated as a subscript. Therefore, coenzyme Q_{10}, whose structure is given in the problem statement, is only one representative of the ubiquinol family. Therefore, the substance isolated from cell culture will behave experimentally as a mixture of compounds with varying numbers of unsaturated bonds.

5. Using mass balance for reaction equation (1), we write the mathematical equation:

$$M_r(\mathbf{X}) = M_r(\mathbf{X1}) + M_r(\mathbf{X2}) - 2$$

Taking into account the inequality $M_r(\mathbf{X}) < 250$, we obtain:

$$M_r(\mathbf{X1}) + M_r(\mathbf{X2}) < 252$$

Since the inequality $M_r(\mathbf{X1}) > M_r(\mathbf{X2})$ is valid, the upper estimate of $M_r(\mathbf{X2})$ is $M_r(\mathbf{X2}) \leq 125$, where $M_r(\mathbf{X1})$, $M_r(\mathbf{X2}) \in \mathbb{N}$.

6. In accordance with the answer to the previous question, the element **E2** in **X2** accounts for no more than 1.3 g/mol of the molar mass. Therefore, **E2** is hydrogen, and in the formula unit of both **X1** and **X2** it is represented by only one atom (since even in the limit, the inequality $250 \times 0.010 < 3$ is valid). Hence, **X** contains no hydrogen (0% by mass), since both hydrogens in the products of reaction (1) originate from ubiquinol.

7. Knowing the mass fraction of oxygen in **X**, we first find three possible values for the sum $M_r(\mathbf{X1}) + M_r(\mathbf{X2})$ with high precision: 70.35, 138.68, and 207.01. However, only for the third option is it possible to obtain the relative mass corresponding to two hydrogen atoms by multiplication. Based on the number of significant digits in the mass fraction of hydrogen, we note the closeness of the values of the terms in the expression $M_r(\mathbf{X1}) + M_r(\mathbf{X2})$. Accordingly, it is necessary to examine the sums 103 + 104, 102 + 105, 101 + 106, etc., to determine whether the correctly rounded mass fraction of hydrogen for $M_r(\mathbf{X1})$ exceeds the lower limit of the range, equal to 0.950%. In this case, $A_r(H)$ should be taken as 1.008 in the calculations. At values of $M_r(\mathbf{X1}) > 106$, the mass fraction of hydrogen becomes less than 0.95. Hence, the range of possible molar masses of **X1** is {104, 105, 106}, meaning for **X2** the range will be {101, 102, 103}.

8. The equality of the element mass fractions in two compounds despite the difference in their whole-number molecular masses can be explained in at least two ways.

First, the small number of significant digits in the element mass fractions.

Second, the composition of **X1** can be represented as multiple formulae $(\mathbf{X1})_2$ or $(\mathbf{X1})_3$. This is impossible due to the small mass fraction of **E2** and the small molar mass of **X** itself.

9. Let's try to determine element **E1**. Since **X1** and **X2** each contain one hydrogen atom, the atomic mass A_r(**E1**) obeys the following equation: $A_r(\mathbf{E1}) \approx 22/n$, where n is the number of **E1** atoms in **X1** (or **X2**).

It is clear that **X** also contains approximately 22% of element **E1** (while not being oxygen, since 23.41% is too far from 22%). Then, by a quick search of four elements (Na, C, B, and Li), we find that **X** contains Na (**E2**) and O in a molar ratio of 2:3, respectively.

Let's rewrite reaction equation (1) taking into account the data obtained:

$$Na_2\ldots O_3 + MKH_2 \rightarrow NaH\ldots + NaH\ldots + MK$$

Note that the three oxygen atoms could be redistributed between **X1** and **X2** in only two ways: 3+0 or 2+1. Thus:

$$Na_2\ldots O_3 + \underline{MK}H_2 \rightarrow NaH\ldots O_2 + NaH\ldots O + MK \text{ (3) or}$$

$$Na_2\ldots O_3 + MKH_2 \rightarrow NaH\ldots O_3 + NaH\ldots + MK \text{ (4)}$$

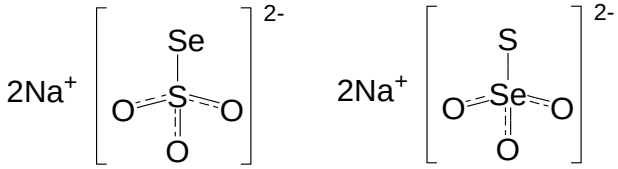
It's clear that we're dealing with neutral (reagents) and acidic salts (one subclass of chemical compounds) of oxygen-containing and (possibly) anoxic acids. For the NaH…O salt, we can't come up with an adequate solution, so reaction equation (4) is correct.

The molar mass of residue **X** (taking into account its constraint in the conditions) accounts for 111 g/mol. However, given the similarity of the molar masses of **X1** and **X2**, we can find the unknown elemental residue in the anoxic acid salt: (111 – 48)/2 = 31.5, which is close to the atomic mass of sulfur. Thus the sodium salt of the oxyacid contains selenium: the sum of 32 + 48 is, in turn, close to 79. From this, we can fully reconstruct the equation for reaction (1) involving sodium selenosulfate (**X**):

$$Na_2SeSO_3 + MKH_2 \rightarrow NaHSO_3 + NaHSe + MK$$

The molar masses of sodium hydrosulfite $NaHSO_3$ (**X1**) and sodium hydroselenide NaHS (**X2**), rounded to the nearest whole number, are close, amounting to 104 and 103 g/mol, respectively. This, given the low accuracy of the element mass fraction calculations, leads to the coincidence of the obtained values.

Now we can depict the structural formulae of both isomeric salts with tetrahedral anions of similar shape:



X **Y**

Based on the formula of **Y**, it is easy to determine the products of the reduction reaction: sodium hydroselenite $NaHSeO_3$ (**Y1**) and sodium hydrosulfide NaHS (**Y2**), simultaneously revealing reaction (2):

$$Na_2SSeO_3 + MKH_2 \rightarrow NaHSeO_3 + NaHS + MK \quad (2).$$

Note that all reactions must occur at a physiological pH close to 7.5, which clearly indicates the preferential formation of acidic salts of the corresponding acids.

10. Based on the condition, neither **X** nor **Y** are natural substrates for any enzymes. A group of compounds found in nature that are structurally similar to them is, of course, the thiosulfates (**Z**), for example, sodium thiosulfate $Na_2S_2O_3$. The active site of the enzyme thiosulfate reductase cannot fundamentally distinguish between metabolites that differ by the replacement of one sulfur atom with selenium, due to the similar geometric parameters of the anions. The sulfate ion also has a similar structure, but it exhibits chemical properties different from those of **X** and **Y**.

11. The known substrates **X**, **Y**, and **Z** must be modified in some way to satisfy the conditional criterion. The only possible option is to use isotopically labeled sodium salts to increase M_r(**X2**) by one unit ($NaH^{80}Se$; ^{80}Se is the most common stable isotope of selenium; **W** is $Na_2{}^{80}SeSO_3$) or decrease M_r(**X1**) by one unit ($NaH^{31}SO_3$ or $NaHS^{15}OO_2$; ^{31}S and ^{15}O are radioactive isotopes with short half-lives). The second option is obviously unsuitable, so the only other suitable candidate is a combined compound – $Na_2{}^{82}Se^{34}SO_3$ –where the ^{34}S isotope is stable, and the ^{82}Se isotope has a long half-life and extremely low radioactivity, which does not interfere with the biochemical experiment. The correct answer is: $Na_2{}^{80}SeSO_3$ and $Na_2{}^{82}Se^{34}SO_3$.

12. It is logical to assume that thiosulfate is integrated into the metabolism of sulfur-containing compounds in archaea. The interconversion of inorganic sulfur and nitrogen compounds is an example of energy metabolism, established at the initial stages of evolution, before the advent of photosynthesis and requiring no oxygen (the process of sulfate reduction with the simultaneous oxidation of organic compounds during anaerobic respiration). Therefore, these archaea, being anaerobes, inhabit bodies of water rich in sulfur compounds (*e.g.*, hydrothermal vents or hypersaline environments).

13. The first hint: indicate that all unknown products of the two reactions belong to the same subclass of chemical compounds (referring to acidic salts), preferably without directly naming the subclass itself in the problem.

The second clue is much more elegant: the structure of any of the six compounds with the letters **X** and **Y** contains at least two elements that are located in the same group of the Periodic Table.

Perhaps you could come up with other, equally creative clues.

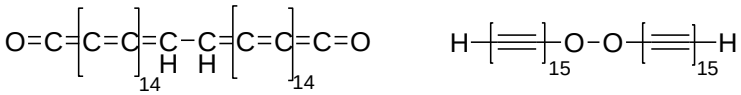
Problem 28. Non-Archea-typal form of life

1. Since cuts occur along the strands and not on the nodes, **A** can be represented as a dimer $(\mathbf{A'})_2$. In other words, the axis of symmetry in **A** passes through the middle of the central bond, not through any specific atom. This, in turn, means that the molecular formula of **A'** is $C_{x/2}H_{y/2}O_{z/2}$, and due to the integer nature of the indices, the values of *x*, *y*, and *z* must be divisible by two without a remainder.

Since each cut breaks exactly one bond, the total number of bonds in molecule **A** is equal to the total number of cuts (123). Note that carbon in organic compounds is tetravalent, while the valences of hydrogen and oxygen, in turn, are 1 and 2, respectively. Considering that two atoms are involved in the formation of each covalent bond, we can write the following equation:

$$\boldsymbol{4x + y + 2z = 246}$$

To estimate the upper limit on the number of carbon atoms, we minimise the number of hydrogen and oxygen atoms to two each (remember the parity!), then *x* equals 60. Theoretically, symmetrical structures corresponding to the formula $C_{60}H_2O_2$ might be:



2. a) Let's try to prove the impossibility of such a molecule in general. It is clear that the number of multiple bonds in such molecule is odd (1, 3, 5, etc.). In fact, it does not matter how many of them are triple or double; the important thing is that they occur in pairs, and the only double bond without repetition is located in the centre of the molecule. Then we can re-write the expression as follows (4*k* is the number of hydrogens lost with the appearance of each pair of double bonds):

$$\boldsymbol{4x + 2x - 4k + 2z = 246, where\ k \in \mathbb{Z}_{\geq 0}}$$

Rearranging, we obtain:

$$\boldsymbol{3(x - 41) = 2k - z}$$

According to the fundamental theorem of arithmetic, with the even number of oxygen atoms *z*, the left-hand side of the equation must be a multiple of two. However, for $\boldsymbol{x \vdots 2}$, both factors (3 and *x* – 41) are odd. Therefore, with an even number of carbon atoms, the equation has no integer solution.

This statement can be also proven by understanding that the removal of the two central bonds would lead to the identical fragments having 121 bonds in total, which is impossible due to the odd number.

b) In the case of a single central carbon-carbon triple bond, the equation transforms as follows:

$$\boldsymbol{4x + 2x + 2 - 4 + 2z = 246}$$

and after rearrangement: $\boldsymbol{3x + z = 124}$

For example, for $z = 4$, the value of x is even, and the formula is $C_{40}H_{78}O_4$.

At the crossroads of chemistry and math: In fact, the formulations in the problem are a veiled application of graph theory in chemistry. An alternative solution to the first two questions is as follows.

The molecule **A** is a connected graph $\boldsymbol{G = (V, E)}$, where vertices are atoms and edges are covalent bonds. The graph has x vertices $\boldsymbol{v_c}$, such that $\boldsymbol{\deg(v_c) = 4}$, y vertices $\boldsymbol{v_h}$ $\boldsymbol{\deg(v_h) = 1}$, and z vertices $\boldsymbol{v_o}$ $\boldsymbol{\deg(v_o) = 2}$. E (the number of edges) is equal to 123 by assumption.

1a) An edge can be removed from graph G such that two identical connected graphs $\boldsymbol{G = (V, E)}$, remain. For G', $\boldsymbol{v'_c, v'_h, v'_o}$ are called vertices that were inherited from $\boldsymbol{v_c, v_h, v_o}$ of graph G. x', y', z' are the numbers of vertices of the form $\boldsymbol{v'_c, v'_h, v'_o}$, respectively.

$$2x' = x; 2y' = y; 2z' = z \text{ (by construction)} => x, y, z - even$$

1b) By the well-known handshaking lemma:

$$\sum \deg(v) = 2 \cdot E => 4x + 2z + y = 246.$$

Suppose the contrary. Let $x > 60$. Then, considering the parity, $x >= 62$.

$$4x + 2z + y \geq 248 + 2z + y \geq 246$$

We arrive at a contradiction, since z and y are non-negative.

2a) We define the graph $G'' = (V'', E'')$ similarly to the graph G', but obtained by deleting two edges. Then:

$$2 * E'' = E + 2 => E'' = 60.5$$

E'' is a natural number. We arrive at a contradiction. Accordingly, this is impossible.

3. The odd molecular mass of a compound composed of carbon, oxygen, and hydrogen can be explained only in one way: the calculations used precise atomic masses of the elements, which have a small deviation from the whole number (0.01 for carbon and 0.008 for hydrogen). With a total number of these atoms close to 100, a surplus of one unit is added to the molecular mass of **A**, which is cleared of tenths and hundredths when rounding. Taking this into account, we can write a system of equations:

$$\begin{cases} \mathbf{4}\boldsymbol{x} + \boldsymbol{y} + \mathbf{2}\boldsymbol{z} = \mathbf{246} \\ \mathbf{12}\boldsymbol{x} + \boldsymbol{y} + \mathbf{16}\boldsymbol{z} = \mathbf{594} \end{cases}$$

Removing *z*, we obtain:

$$\boldsymbol{y} = \frac{\mathbf{1374} - \mathbf{20}\boldsymbol{x}}{\mathbf{7}}$$

Chemical considerations (remember the alkanes!) yield the inequality:

$$\boldsymbol{y} \leq \mathbf{2}\boldsymbol{x} + \mathbf{2}$$

Substituting one into the other, we obtain $\boldsymbol{x} \geq \mathbf{40}$.

Now we express *x* in terms of *z* from the system of equations:

$$\boldsymbol{x} = \frac{\mathbf{348} - \mathbf{14}\boldsymbol{z}}{\mathbf{8}}$$

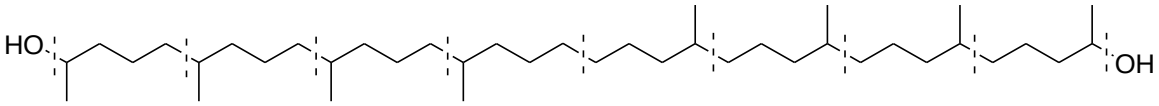
Substituting this into the inequality we just obtained, we find that $\boldsymbol{z} \leq \mathbf{2}$. But, as we discovered earlier, *z* is an even natural number, so *z* = 2 is certain.

Now it is easy to find the values of *x* (40) and *y* (82), as well as the molecular formula of **A** as $C_{40}H_{82}O_2$.

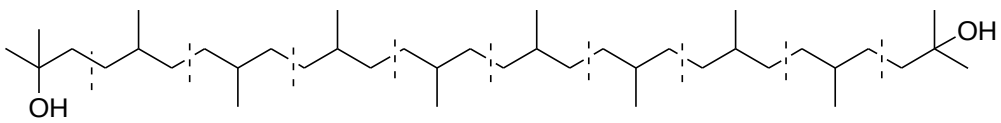
4. Note that each of the terminal *d* fragments must contain one hydroxyl group, since the number of *d* fragments coincides with that of oxygen atoms in **A**. Incidentally, the possibility that **A** is an ether is discriminated by the wording of the question and the small number of oxygen atoms in the molecule.

Since molecule **A** contains 8 stereocentres, there are 4 branching points of the hydrocarbon chain on each side of the symmetry axis, one for each *c* fragment.

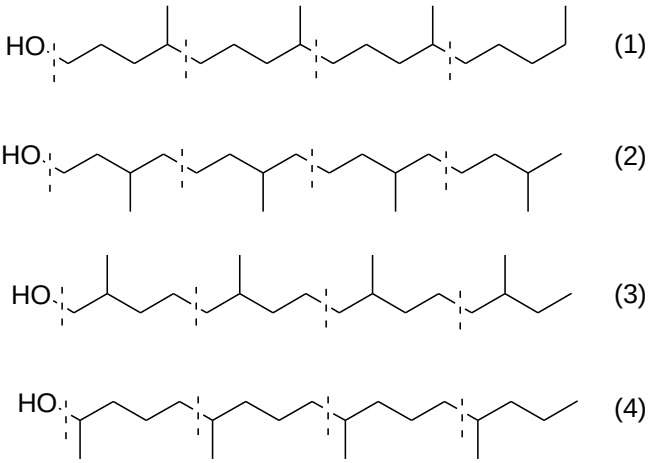
The hydroxyl group can be secondary only if the *d* fragment is represented by a hydroxyl group, and the stereocentre of the adjacent c fragment is represented by a carbon atom directly bonded to the hydroxy group (the C–O bond is cut):



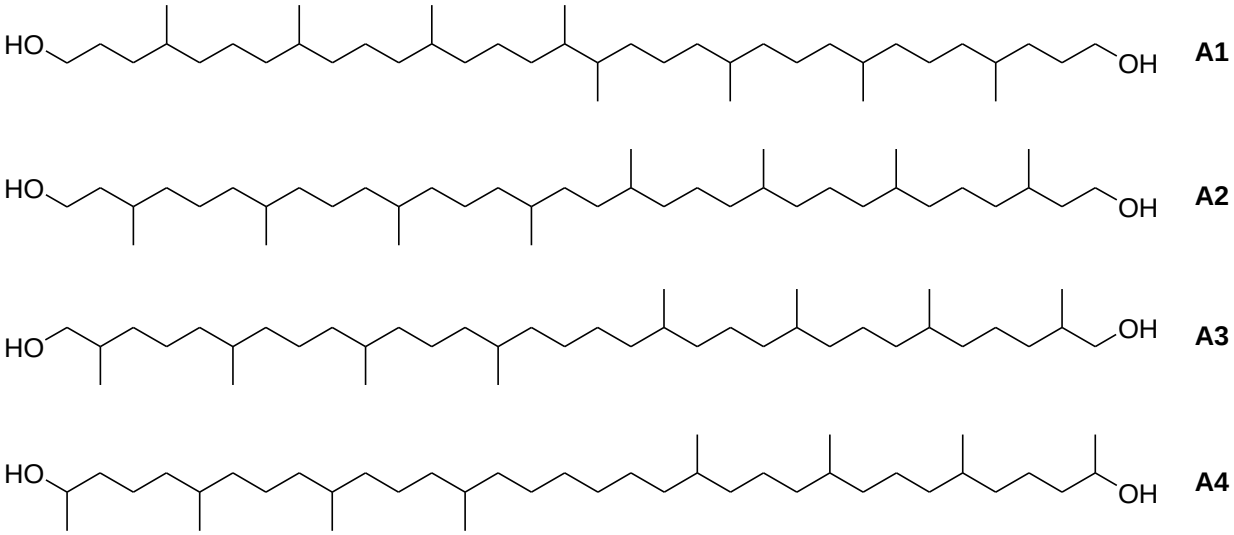
There are many more variants with a tertiary hydroxyl group, for example, one shown below (with two methyl groups bonded to the carbon atom to avoid the formation of an additional stereocentre):



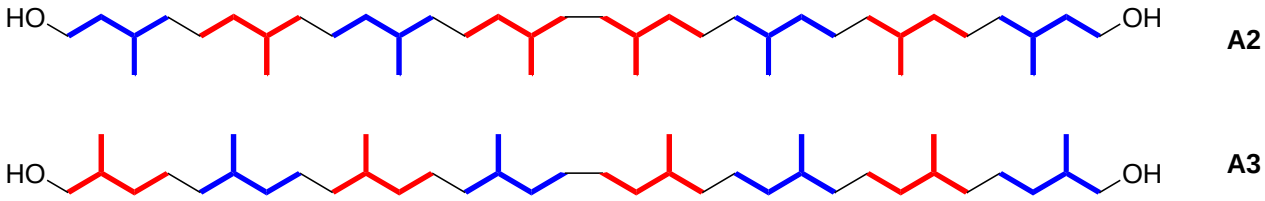
5. Since two symmetrical half-molecules with 20 carbon atoms each can be found in **A**, each c fragment have at least 5 carbon atoms (three C-C bonds formed by four atoms plus a methyl group as a branch), and *d* is a hydroxy group (a complete fragment consisting of two linked nodes). Clearly, the length of **A** is not sufficient to obtain eight *c* fragments containing six or more carbon atoms. 4 possible half-chains of **A** with corresponding cuts are:



All of these satisfy the condition. Therefore, 4 structures are possible for **A** (**A1**-**A4**) are:



6. Structures based on isoprenyl units (shown in colours) are extremely common in nature, so two intermediate variants remain for consideration. Note that structures **A2** and **A3** differ in the type of attachment of the two symmetrical fragments being head-to-head in the former case, and tail-to-tail in the latter case:



7. Note that the two **A** molecules contain four hydroxyl groups, and four water molecules reacted. Therefore, due to symmetry, **C** must contain at least two identical oxygen-containing functional groups, and **B** is a single large macrocycle, regardless the order in which the **A** and **C** residues are joined:

In any case, **B** will contain not more than 2*k* hydrogen atoms.

8. Since the hydrolysis reaction of **B** is not a redox reaction, we can apply the assumption that the total number of bonds in the reactants is equal to the total number of bonds in the products. It is easy to see that in this case, the number of bonds in **B** is an even natural number (262 or 264), since all the remaining substances in the reaction equation have even coefficients. It is also important to understand that the elemental composition of **C** is similar to that of **B**.

Let's consider both options (x_1, y_1, and z_1 are the numbers of carbon, hydrogen, and oxygen atoms, respectively).

First:

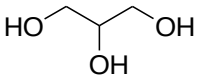
$$\mathbf{264 + 4 * 2 = 2 * 123 + 2}\boldsymbol{x}\mathbf{,\ where\ }\boldsymbol{x}\mathbf{\ is\ the\ number\ of\ bonds\ in\ C}$$

Hence, $\boldsymbol{x = 13}$. Then, the following equation is valid:

$$\mathbf{4}\boldsymbol{x_1} + \boldsymbol{y_1} + \mathbf{2}\boldsymbol{z_1} = \mathbf{26}\text{, where } \boldsymbol{z_1} \geq \mathbf{2}$$

The presence of an axis of symmetry implies that the number of carbon atoms is odd, while chemical considerations imply the inequality: $\mathbf{1} < \boldsymbol{x_1} < \mathbf{7}$. Then $\boldsymbol{x_1} \in \{\mathbf{3}, \mathbf{5}\}$.

If **C** is a dicarboxylic acid (**B**, in turn, is a diester), then $C_3H_6O_4$ is the only formula for it, which is impossible. If **C** contains hydroxyl groups (**B** is an ether), then for C_5 molecular formulae, the hydroxyl groups must be located at unsaturated bonds, so $C_3H_8O_3$ is the only formula left for consideration. It corresponds to glycerol.

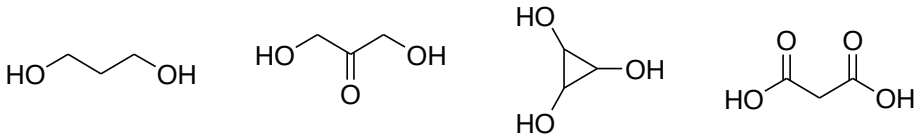


Second:

$$\mathbf{262 + 4 * 2 = 2 * 123 + 2}\boldsymbol{x}\mathbf{,\ where\ }\boldsymbol{x}\mathbf{\ is\ the\ number\ of\ bonds\ in\ C}$$

$$\mathbf{4}\boldsymbol{x_1} + \boldsymbol{y_1} + \mathbf{2}\boldsymbol{z_1} = \mathbf{24}\text{, where } \boldsymbol{z_1} \geq \mathbf{2}$$

It is clear that $\boldsymbol{x_1 = 3}$, so three options are possible: $C_3H_8O_2$, $C_3H_6O_3$ and $C_3H_4O_4$. One structure can be proposed for each of the molecular formulae $C_3H_8O_2$ and $C_3H_4O_4$, while two structures can be proposed for $C_3H_6O_3$:

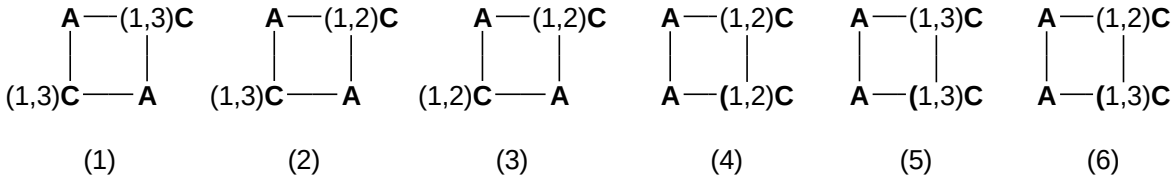


However, cyclopropane-1,2,3-triol is an extremely unstable compound. It should also be noted that the cyclic dimer of dihydroxyacetone does not correspond to the material balance of the equation of the hydrolysis of **B**.

All other symmetrical substances (glycerol, 1,3-propanediol, dihydroxyacetone, and malonic acid) are widely found in nature.

9. The presence of three hydroxy groups in glycerol greatly increases the range of possible variants of **B**, while for each of the rest three molecules the number of isomers equals 2, according to the schematic diagram in q7.

The six structural isomers of **B**, as a combination of glycerol and diphytanyl diol, can be represented schematically:

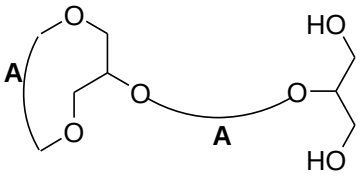


(1) (2) (3) (4) (5) (6)

Each of the variants **B1**, **B2**, and **B5** corresponds to one isomer. The remaining three will yield a greater number of isomers due to the order of formation of ester bonds by the hydroxy groups of glycerol (the numbers indicate their positions):

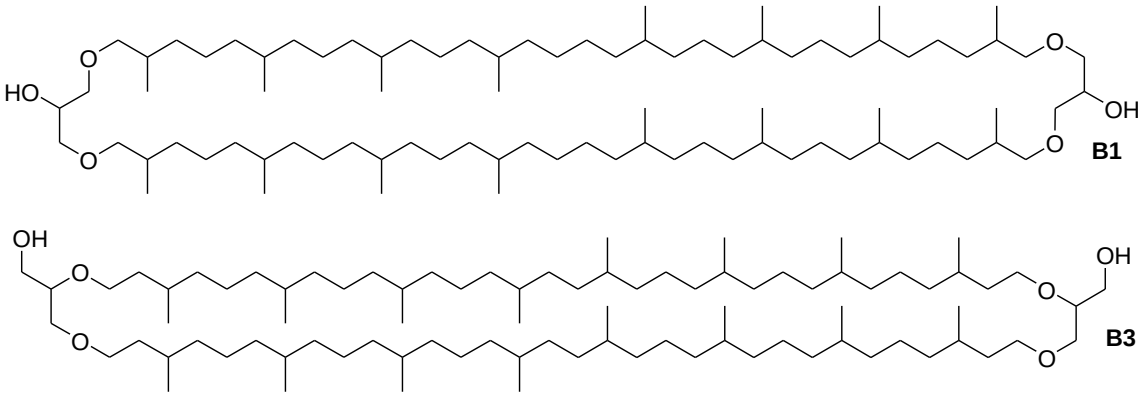
Option	Order of formation of ether bonds by glycerol residues
B3	-**A**-1-2-**A**-1-2-; -**A**-1-2-**A**-2-1-
B4	-**A**-1-2-2-1-**A**-; -**A**-2-1-1-2-**A**-; -**A**-2-1-2-1-**A**-
B6	-**A**-1-2-1-3-**A**-; -**A**-2-1-1-3-**A**-

The indication of identical molecular formulae for the glycerol residues excludes the following variants from consideration:



Thus, the total number of isomers satisfying the condition for each type of **A** is 10, and the total number of isomers is 20.

The structural formulae of **B1** based on **A3** and **B3** based on **A2** are given below as examples:

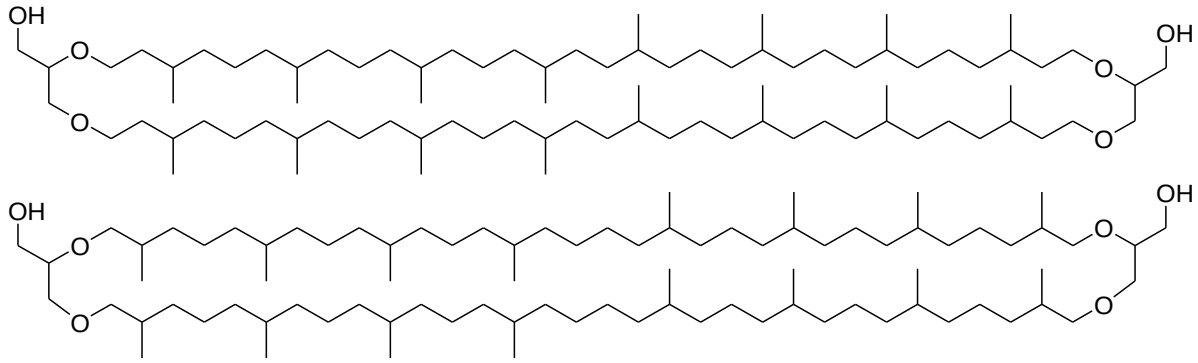


10. The properties of **B** described in the condition correspond to ethers, the hydrolysis of which requires fairly harsh conditions or long time if the conditions are close to standard ones. Enzymes that hydrolyse ester bonds (esterases) are widespread in nature, whereas enzymes that accelerate the hydrolysis of ethers are absent, which is an evolutionary advantage for archaea in competition with bacteria. Furthermore, archaea often live at temperatures that are extremely high for life, which also explains why ethers have an evolutionary advantage over complex membranes that are in direct contact with the aggressive (temperature, pH) aqueous environment of a hydrothermal vent.

11. We will continue to consider derivatives of isomers **B1**, **B2**, and **B3**, as the other ten will behave like classic phospholipids with a hydrophilic head and a lipophilic tail. The presence of two hydrophilic ends, which can be exposed on opposite sides of the membrane (*O*-conformation) or on the same side (*U*-conformation) is a distinctive feature of phospholipids based on dialkylglycerol tetraethers. The monolipid layer that forms the membrane of some archaea has lower fluidity than its bilayer counterpart, which is another factor in archaea's resistance to high ambient temperatures.

12. Polyhydric alcohols with a molecular weight lower than glycerol are either the simplest diol, ethylene glycol, or isomeric propanediols. When creating a phospholipid based on these compounds, one hydroxyl group will participate in the formation of an ester bond with a fatty acid (the lipophilic tail), while the second hydroxyl will be converted to a hydrophilic head after esterification. However, the presence of only one fatty acid residue (the tail) per lipid molecule geometrically leads to the formation of micelles rather than the required liposome bilayer (**the first reason**). Glycerol derivatives (in particular, glycerol-3-phosphate) are integrated into key metabolic pathways (for example, glycolysis), while ethylene glycol and propanediol derivatives are not actually used for living purposes (**the second reason**). Of course, since we are not privy to the evolution's plans, other answers to this question are possible.

13. Continuing our considerations of the correct arrangement of lipids in the archaeal membrane, we can see that the 1,3-arrangement of diphytanyl diol residues would be unfavourable in terms of the dense packing of isoprenyl units into the hydrophobic layer (carefully examine the width of the structures obtained in q9). Therefore, the structure (**A**-(1,2)**C**)$_2$ remains under consideration, which corresponds to two isomers due to the variability of the structure of **A**:



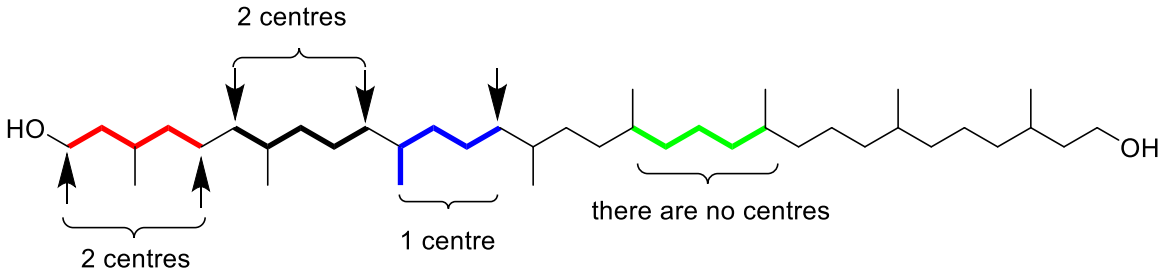
14. Note that any position of the residue **A** in molecule **B** occurs four times due to the symmetry and duplication of **A**. Thus, two enzymes specific to two different positions should form, in the limit, eight rings (in the limit, since **D1** withstands high temperatures best of all in the family). Therefore, the molecular formula of **D1** is calculated as $C_{86}H_{172}O_6 - 8H_2 = C_{86}H_{156}O_6$.

For future reference, we note that theoretically even more rings can be accommodated in molecule **B** (based on the premise of the third and fourth enzymes).

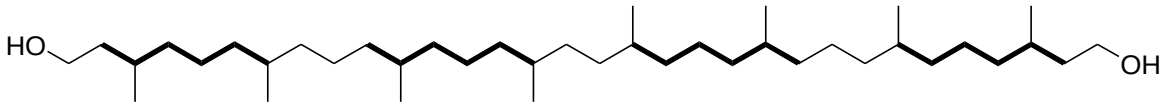
15. B contains 18 stereocentres. Thus, with the formation of eight rings in the **D1** molecule, additional eight stereocentres are formed. In other words, the formation of each ring adds one stereocentre; other variants (e.g., 0+2+0+2+0+2+0+2) can be excluded based on the similarity of the enzymes and the loci they act on. This consideration greatly limits the possibilities for hydrocarbon ring formation from parent isoprenyl units.

Let's consider variants with cyclopentane and cyclohexane fragments separately.

In the case of the formation of cyclopentane residues, 4 possible arrangements of the five carbon atoms relative to the isoprenyl chain are possible. Two of these result in two stereocentres per ring, one variant results in no increase in the number of stereocentres, and only one satisfies the condition:



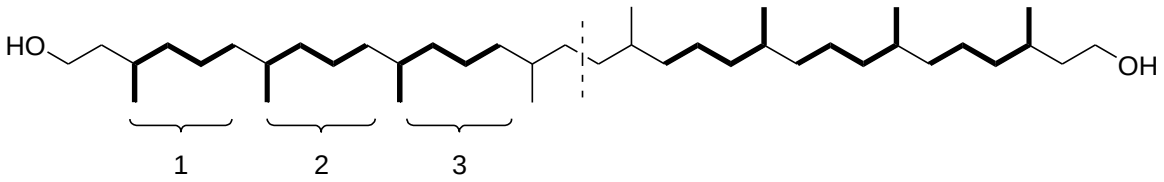
For cyclohexane, several possible arrangements of the six carbon atoms relative to the isoprenyl residues can be proposed, and only one satisfies the "one ring – one stereocentre" condition. However, it is not possible to place more than four such rings in a single **A** molecule:



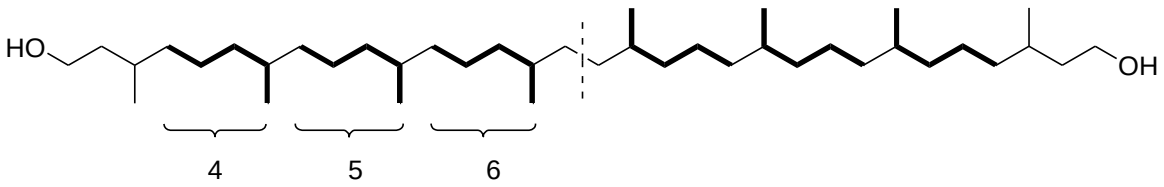
However, cyclohexane rings are found in some archaeal lipids, such as Crenarchaeola.

It should be noted that using **A3** instead of **A2** does not fundamentally change the picture.

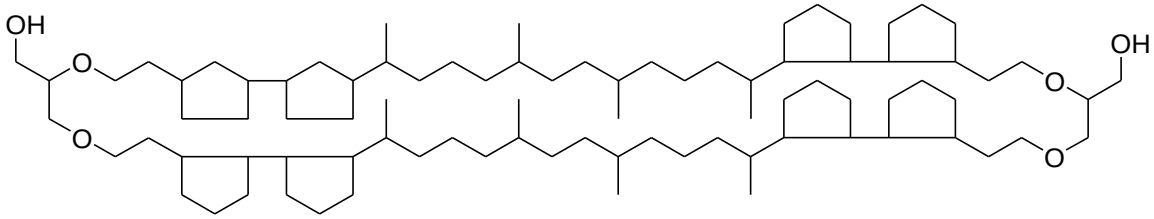
If the rings are formed in a left-to-right direction, the prototype of the future tetracyclopentane derivative looks as follows:



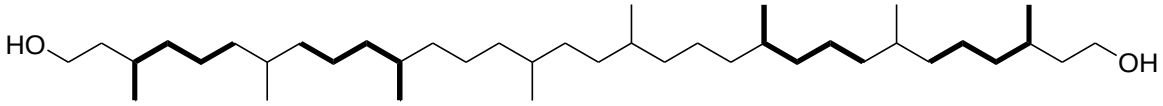
However, nothing prohibits the reverse direction of ring formation; the only difference is that position 1 is lost, but an additional position appears in the center of **A**:



It is clear that for each direction, three binary combinations are possible: 1+2, 1+3, 2+3, and 4+5, 4+6, 5+6. Then, the two different diols in **D1** correspond to 12 isomers. The structure of one that actually occurs in nature is shown below:



Structures that are opposite in the direction of ring formation (an example is given below) should be excluded from consideration given the extreme similarity of the functions of the two enzymes that cyclise the molecule. However, considering these **D1** isomers would not be a significant error.

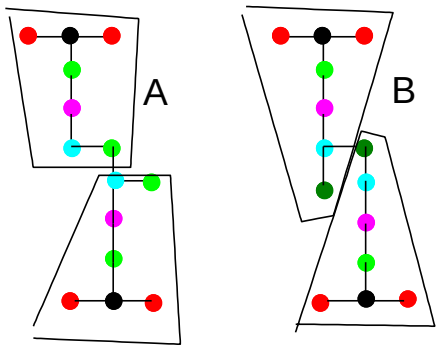


This problem examines members of the glycerol dialkyl glycerol tetraester (GDGT) family, a class of membrane lipids.

Problem 29. Archaea-ology

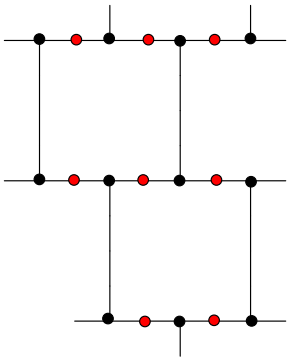
1. The formulae for sequences **A** and **B** can be represented as follows: $(X_1)_4(X_2)_2(X_3)_4(X_4)_2(X_5)_2$ and $(X_1)_4(X_4)_2(X_5)_2(X_6)_2(X_7)_2(X_8)_2$. It is clear that each monomer is present in the biopolymer fragments an even number of times (any index can be represented as 2n), so the contribution of each to the total molecular mass is also an even number. Moreover, the difference between two even numbers is always an even number.

2. To construct the secondary structure of polymers, it is first necessary to correctly schematically depict the I-shaped structure:



The figure shows that the four terminal atoms for each structure are represented by the same monomer (red, X_1); matches in other colours in **A** and **B** are random.

Each I-shaped sequence can be divided into two oppositely oriented T-fragments (as shown above inside the frames). Then, it is necessary to replicate the I-shaped fragments to construct a two-dimensional network. The only feasible way is to match (partially superimpose) the upper and lower crossbars of the "letters I" with the corresponding crossbars of the neighboring ones. Only alteration of $\mathbf{X_1}$ and $\mathbf{X_6}$ in the polymer backbone can provide the ratio of the monomers in **B** close to equimolar. So, the biopolymers **A** and **B** are the networks built of chains of red ($\mathbf{X_1}$) and black ($\mathbf{X_2}$, $\mathbf{X_6}$) monomers linked by palindromic bridges, according to the following diagram:



In other words, to form a uniform two-dimensional network, adjacent I-shaped fragments must face opposite directions.

3. a) Let's note several key features of the sequences presented in the problem statement:

- there are branching points ($\mathbf{X_2}$, $\mathbf{X_5}$, and $\mathbf{X_6}$) present, where a monomer forms covalent bonds with three other ones;
- the sequence composition is quite diverse (8 monomer types at 28 positions), with the primary structure at several positions coinciding between **A** and **B**;
- there is an indication of the key importance of short oligomeric fragments, which apparently repeat multiple times, forming a fairly large polymer (note the estimation of the molecular weights).

Most of the criteria are met by non-ribosomal peptides (without the participation of ribosomes, as the translation process implies the formation of a linear polymer) or proteins subjected to post-translational modification. Polysaccharides do not typically exhibit a wide variety of monomer units, although the structure of monosaccharides implies ample branch points. Therefore, the condition does not exclude heteropolysaccharides.

b) Since the number of monomers is greater than that of molecular formulae describing them, some monomers in fragments **A** and **B** are isomers of each other. This can occur for amino acids (recall the leucine-isoleucine pair) and, to an even greater extent, for monosaccharides (for example, the isomerism of aldohexoses).

c) Some amino acids may be present in proteins in a phosphorylated form (such as phosphoserine), which is quite common in both proteins and glycoproteins. Thus, this option is possible.

d) Aldohexoses and ketohexoses, which can be present in the long main chains of **A** and **B** can be considered as an example of interclass isomerism of monomers forming biopolymers.

e) Since the biopolymers can contain peptide fragments, theoretically, four homologous amino acids (glycine, alanine, valine, and leucine) can be present in the sequences.

Thus, theoretically, none of the five options can be excluded from consideration.

4. Based on the structures of polymers **A** and **B**, we can assume that these are heteropolymers containing heteropolysaccharides (alternation of two monosaccharide residues within long chains) and short peptides linkers. Such structural format is impossible for any other biological homopolymers: even fibrillar proteins with a poor amino acid composition form their 3D structures mostly through regular system of non-covalent bonds rather than regular bridges between chains.

5. The partitioning according to the criterion of fragment pairing is as follows:

for **A**: $(\mathbf{X}_1)_2\mathbf{X}_2 + \mathbf{X}_3\mathbf{—}\mathbf{X}_4 + \mathbf{X}_5\mathbf{—}\mathbf{X}_3$;

for **B**: $(\mathbf{X}_1)_2\mathbf{X}_6 + \mathbf{X}_4\mathbf{—}\mathbf{X}_7 + \mathbf{X}_5\mathbf{—}\mathbf{X}_8$.

The trisaccharide fragments $(\mathbf{X}_1)_2\mathbf{X}_2$ and $(\mathbf{X}_1)_2\mathbf{X}_6$ have the molecular masses of 642 and 700, respectively ($M_r(\mathbf{X_6}) > M_r(\mathbf{X_2})$. $\mathbf{X_1}$ is not an isomer of either $\mathbf{X_2}$ or $\mathbf{X_6}$, since monomers with identical molecular masses do not form covalent bonds with each other. So, $M_r(\mathbf{X_1}) \neq M_r(\mathbf{X_2}) \neq M_r(\mathbf{X_6})$.

Note that three molecular masses correspond to four dipeptide fragments. This means that two dipeptides have the same molecular mass, and they are isomers (if not, four molecular masses would be needed in the case of two similar peptides). Thus, there are four dipeptides built of five amino acids, Moreover, the set includes two pairs or trios of structural isomers (interclass isomerism between monosaccharides and amino acids is impossible!). Given that no dipeptide contains amino acid residues that are isomeric to each other, a system of three equations can be proposed:

$$\begin{cases} M(\text{AA1}) + M(\text{AA2}) = 235 \\ M(\text{AA1}) + M(\text{AA3}) = 236 \\ M(\text{AA2}) + M(\text{AA3}) = 293 \end{cases}$$

The solution leads to the molecular masses of the amino acids of 89, 146, and 147 g/mol.

The molecular mass of 89 g/mol can only correspond to alanine, which naturally occurs in two stereoisomers (D and L) as well as in a structural isomer, β-alanine. Alanine cannot be found at branch points; these positions can be occupied by amino acids with an additional carboxyl group or amino group resides: M(lysine) = 146 g/mol and M(glutamic acid) = 147 g/mol. For these amino acids, possible presence of their stereoisomers in **A** and **B** can also be assumed. Thus. we cannot identify at least some of the monomers at this stage, since there may be both lysine and glutamate at the branching points.

6. Options a) and b) may be correct, since four elements (C, H, N, and O) can easily be combined into various combinations with matching molecular masses (for example, CH_4 and O. As to option c), evolution has no design, purpose, or predetermined plan from the scientific biological perspective.

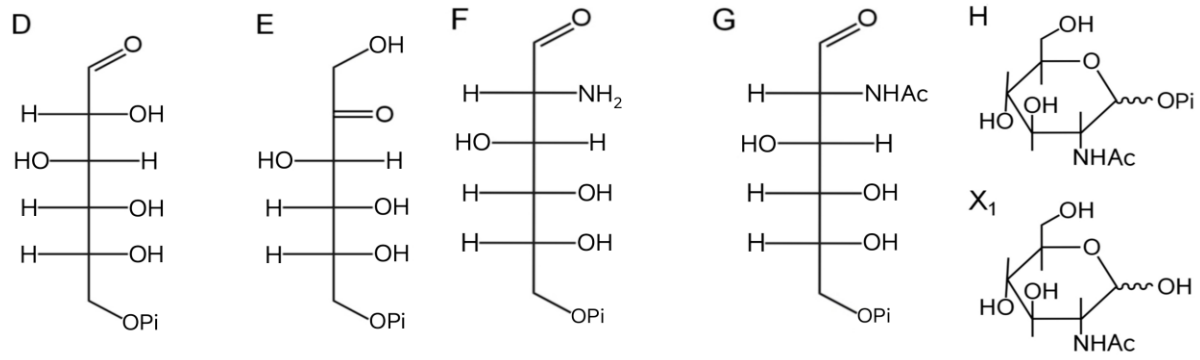
7. Since each sequence is split into three pairs of fragments, the sum of the molecular masses of the three different fragments can be found as 2270/2 = 1135 g/mol. Note that the identical masses of the sums of the fragments can only be explained by the presence of the Glu+Lys pair in **A**. This leaves 860 g/mol of molecular mass for the two remaining pairs, which can be represented as the sum of 218 (Ala+Glu) and 642 **[$(\mathbf{X_1})_2\mathbf{X_2}$]**.

Amino acid **X_3**, repeated four times in **A**, is glutamic acid. **X_3** is not present at branch points, meaning **X_5** is lysine. The remaining amino acid, **X_4**, is alanine.

Similar calculations can be performed for the fragment **B** with the composition of Ala + Glu, Ala + Lys + **$(X_1)_2X_6$**. Based on the positions of **X_4** and **X_5**, we can determine that **X_8** is alanine, and **X_7** is glutamic acid. We were only able to determine the molecular formulae of the amino acid monomers; choosing between D- and L-alanine (β-alanine), D- and L-glutamic acids (as well as the variant of L-isoglutamic acid) is not possible.

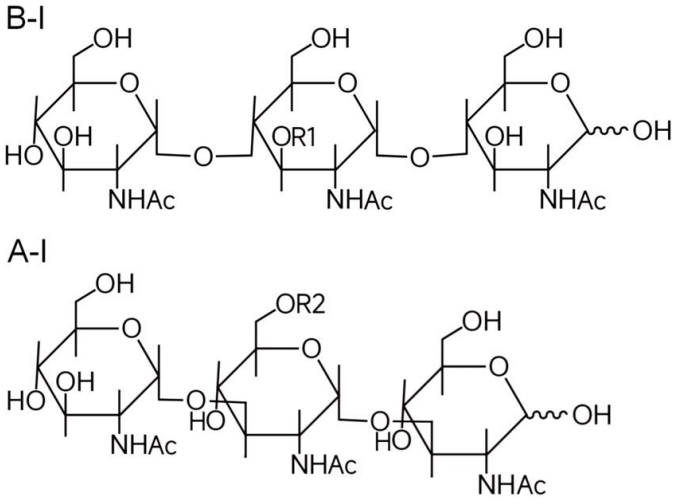
X_3	**X_4**	**X_5**	**X_7**	**X_8**
$C_5H_9NO_4$ (Glu)	$C_3H_7NO_2$ (Ala)	$C_6H_{14}N_2O_2$ (Lys)	$C_5H_9NO_4$ (Glu)	$C_3H_7NO_2$ (Ala)

8. Since **X_1** does not contain phosphorus, there must definitely be one phosphoric acid residue in **D** in order to maintain the balance of phosphate groups. The number of carbon atoms changes only at the stage **F→G** (acetyl group transfer), thus **D** contains six carbon atoms. Taking into account the reference to photosynthesis, **D** and **E** are glucose-6-phosphate and fructose-6-phosphate, only their correlation with a specific letter is required. We can conclude that the transferase transfers an amino group (the molecular masses of glutamine and glutamate differ by 1), while **F** must have a free aldehyde group in order to react with the Tollens reagent:

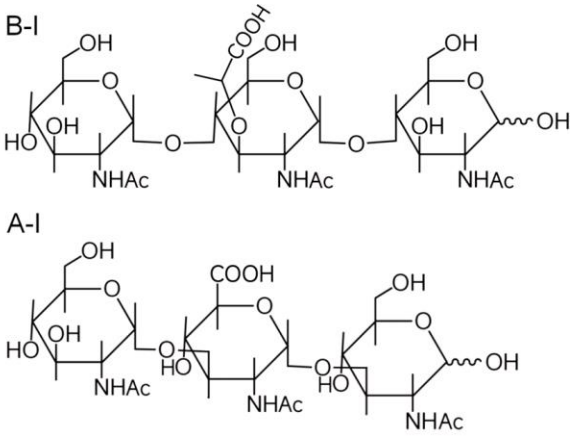


Note that the molecular weights also differ by one in the **D**-**F** and **E**-**F** pairs.

9. It is possible to determine which hydroxy groups are free, as well as the order of the monosaccharide residues as they are connected to each other. It is impossible to determine the stereochemistry of glycoside bonds (α or β):

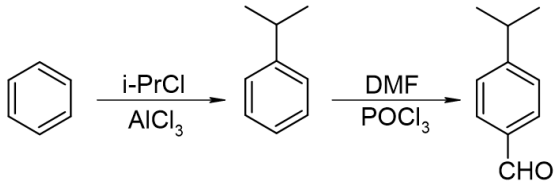


10. Since $\mathbf{X_2}$ and $\mathbf{X_6}$ bind to the peptides, they most likely contain a carboxyl or amino group in the non-methylated fragments. Using the molecular mass of the trisaccharide fragment for **A** (642 g/mol after rounding!), we obtain that the carboxyl group occupies the sixth position in $\mathbf{X_2}$. Acetaldehyde (CH_3CHO) is formed from the fragment -CH(OH)-CH_3 upon oxidation with HIO_4. Such a fragment is present only in lactate (-O-C(O)-CH(OH)-CH_3):

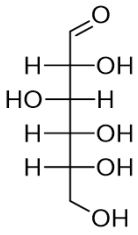


Problem 30. Pilaf ingredients

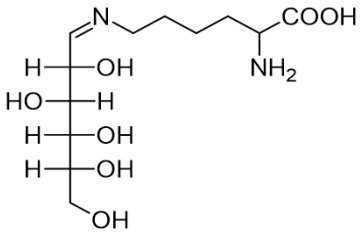
1.



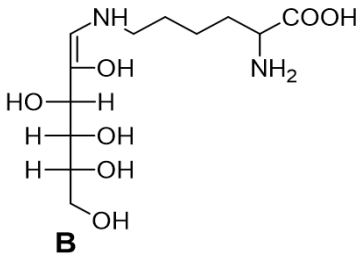
2.



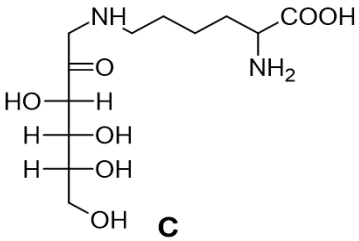
X



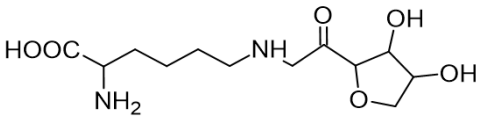
A



B



C



D

3. **a)** Note that the total number of atoms in the $\mathbf{F_2}$ molecule can be represented as 2ε. However, from the condition that $N = 2k$, where $k \in \mathbb{N}$, we can write the following expression:

$$\alpha + \beta + \gamma + 2\varepsilon = 2k$$

From this equation, it follows that the sum of the atoms in compound $\mathbf{F_1}$ is an even number.

From the valences of the elements that form $\mathbf{F_1}$ and $\mathbf{F_2}$, it follows that the coefficients β and ε, corresponding to the hydrogen atoms, are always even.

Let's look at the arithmetic progression formed by α, β, and γ (the order of the terms is naturally unknown). Since the sum of these coefficients, as we just found out, is an even number, two options are possible:

a) α and γ are both odd numbers;

b) α and γ are both even numbers.

If option a) is correct, then there must exist a geometric progression containing two odd numbers and one even number. However, this is impossible, since there are geometric progressions constructed entirely from odd numbers, and geometric progressions constructed from one odd number and the rest even.

Therefore, α and γ are both even numbers. The parity of the sum of the atoms in the compound $\mathbf{F_2}$ (2ε) leads to the conclusion that the last unrecognized coefficient ε is also an even number. Therefore this proves the statement completely.

b) Note that oxidative ozonolysis of the acyclic compound **F** yielded only two products in an equimolar ratio. This means that **F** contains one unsaturated bond, which, depending on its structure, can cleave to form either a ketone or a carboxylic acid.

However, the keto group contains an odd number of oxygen atoms (one), which contradicts the conclusions made previously regarding the coefficients α and γ. Therefore, ozonolysis can only form carboxylic acids. It should also be remembered that hydrolysis of the amide bond in **F** releases a carboxyl group, which is inherited by one of the ozonolysis products (there are no other oxygen-containing groups).

The simplest geometric progression consists entirely of even numbers: 2, 4, and 8. The next progression is 2, 6, and 18. However, for this progression, the sum of the coefficients in $\mathbf{F_1}$ will be 26, and the set of coefficients for $\mathbf{F_2}$ will yield a minimum sum of 4 + 6 + 10 = 20, which together equals 46, exceeding the limit set by the condition.

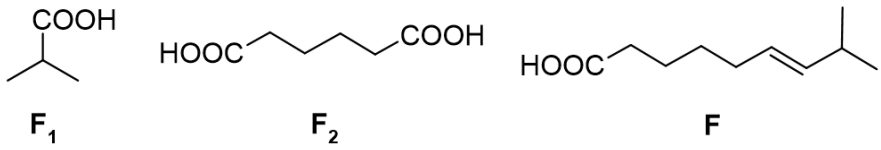
It remains to relate the coefficients α, β, and γ to the values 2, 4, and 8. Note that ozonolysis breaks all unsaturated bonds, and $\mathbf{F_1}$ (the derivative of **F**) contains no rings. Therefore, only one chemical combination is possible: the formula for compound $\mathbf{F_1}$ is $C_4H_8O_2$.

Compound $\mathbf{F_2}$ contains four oxygen atoms. Similar to $\mathbf{F_1}$, we calculate the number of carbon atoms in it (δ):

$$2\delta - 2 = \delta + 4$$

Solving, we find that $\delta = 6$. Thus the molecular formula of $\mathbf{F_2}$ is $C_6H_{10}O_4$.

4. The only possible structures for $\mathbf{F_1}$ and $\mathbf{F_2}$ given the information in the question is as follows:



$\mathbf{F_1}$ $\mathbf{F_2}$ **F**

5. According to the general combustion equation for **E**:

$$(\mathbf{E})\ C_xH_yN_zO_w + (x{+}0.25y{-}0.5w)\ O_2 \rightarrow xCO_2 + 0.5yH_2O + 0.5zN_2$$

we find that the rounded molecular weight of **E** can be expressed in terms of the number of nitrogen atoms it contains:

$$n(N_2) = \frac{0.0915}{28.02} = 3.267 \times 10^{-3}\ \text{mol}$$

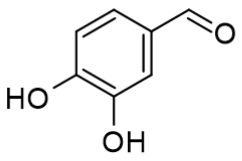
$$M(\mathbf{E}) = \frac{1 \times 0.5z}{3.267 \times 10^{-3}} = 153z$$

Note this means that the dependence of the molecular weight per oxygen on the number of nitrogen atoms is as follows:

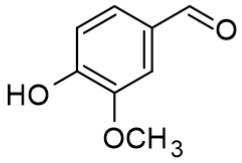
$$M(\mathrm{O}) = 153z \cdot 0.2092 = 32z$$

This means that the empirical formula **E** contains nitrogen and oxygen in a molar ratio of 1:2 and looks like NO_2. Carbon and hydrogen then account for the remainder of the molecular weight, which is equal to 107. This can only correspond to the C_8H_{11} group (9 carbon atoms weigh 108, and for seven carbon atoms we get chemical nonsense in the form of C_7H_{23}). Hence, the empirical formula of **E** is $C_8H_{11}NO_2$.

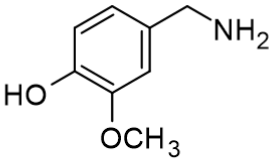
6.



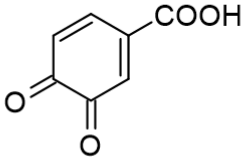
G



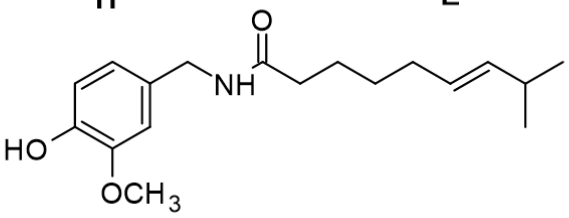
H



E

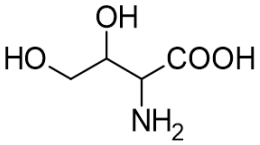


I

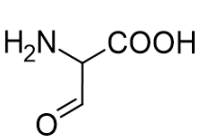


Z

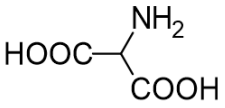
7.



W

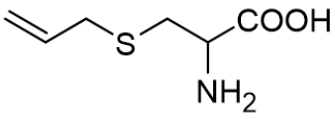


K

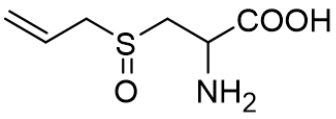


L

8.



N



M

9. General formula for arithmetic progressions:

$$a_n = a_1 + d(n-1)$$

where *d* is common difference

Formula for given progression with molar masses:

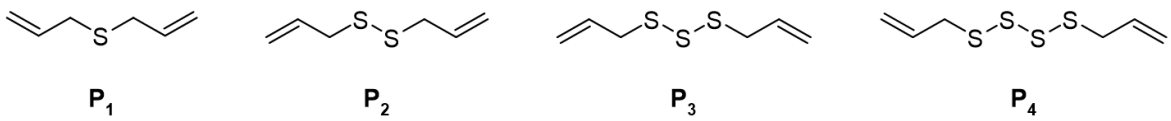
$$M_n = M_1 + d(n-1)$$

$$\begin{cases} M_2 = M_1 + d = 146 \text{ g mol}^{-1} \\ M_3 = M_1 + 2d = 178 \text{ g mol}^{-1} \end{cases}$$

Hence, *d* = 32

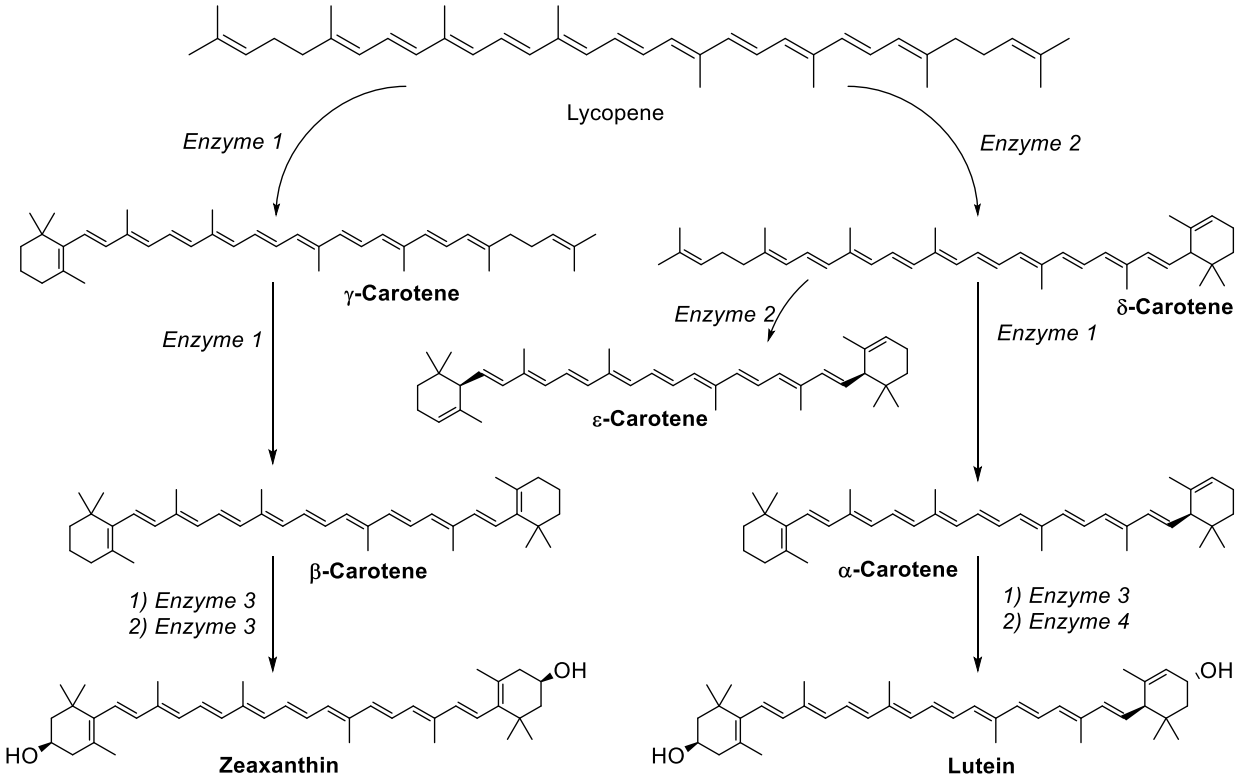
$$M_1 = M_1 + 32(1-1) = 114 \text{ g mol}^{-1}$$
$$M_4 = M_1 + 32(4-1) = 210 \text{ g mol}^{-1}$$

If P_1 is synthesised from alliin, then it should contain a sulfur atom and an allyl substituent. If we check it with molar mass, 114 – 32 – 41 = 41, it indicates that there are two allyl groups attached to the sulfur atom. So, with respect to planes of symmetry, we can propose the following structures:



P_1 **P_2** **P_3** **P_4**

10. Since *Enzyme 1* acts two times to produce β-carotene from lycopene we can conclude that this enzyme catalyzes cyclisation. In the lutein core one double bond is shifted. Therefore *Enzyme 2* catalyses formation of second cycle type with an isomerised double bond. The difference between α-carotene and lutein is in hydroxyl groups, which have different configurations. *Enzymes 3* and *4* are hydroxylases.



11. Since only one type of enzyme is used to obtain zeaxanthin, it formed as either the *R,R*- or *S,S*- form. *R,S*- and *S,R*- are *meso*-forms. Zeaxanthin in this scheme is therefore optically active.

12. *Enzymes 1,2* – **Isomerases** , *Enzymes 3,4* – **Oxidoreductases**

13. Ask the chef in *Besh Qozon* during IChO 2026. (: